



Desafíos

ISSN: 0124-4035

ISSN: 2145-5112

revistadesafios@urosario.edu.co

Universidad del Rosario

Colombia

Isaza Villamizar, Emiliano; González-Casabianca, Felipe; Herrera, Santiago; Rodríguez-Barraquer, Tomás; Ángel, Andrés; Corredor, Vladimir; Feged-Rivadeneira, Alejandro

Políticas públicas, grandes datos, teoría de redes y COVID-19*

Desafíos, vol. 32, núm. 2, 2020, Julio-, pp. 1-19

Universidad del Rosario

Colombia

DOI: <https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.9205>

Disponible en: <http://www.redalyc.org/articulo.oa?id=359663370012>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UAEH  redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Políticas públicas, grandes datos, teoría de redes y COVID-19*

EMILIANO ISAZA VILLAMIZAR**
FELIPE GONZÁLEZ-CASABIANCA***
SANTIAGO HERRERA†
TOMÁS RODRÍGUEZ-BARRAQUER††
ANDRÉS ÁNGEL†††
VLADIMIR CORREDOR§
ALEJANDRO FEGED-RIVADENEIRA§§

Doi: <https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.9205>

Para citar este documento: Isaza Villamizar, E., González-Casabianca, F., Herrera, S., Rodríguez-Barraquer, T., Ángel, A., Corredor, V., & Feged-Rivadeneira, A. (2020). Políticas públicas, grandes datos, teoría de redes y COVID-19. *Desafíos*, 32(2), 1-19. <https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.9205>

* Agradecemos a Ana María Bermúdez y Manuel Riaño, quienes facilitaron el espacio, gestionaron los datos y coordinaron el equipo de trabajo. También, a Mónica Patiño y Servinformación que facilitaron datos y recursos para que esta investigación fuera posible. Por último, al equipo de Salud Pública de la Secretaría Distrital de Salud de Bogotá, quienes participaron en las discusiones sobre la importancia de este tipo de análisis.

** Servinformación. Bogotá, Colombia. Correo electrónico: emiliano.isaza@servinformacion.com.

*** Universidad de los Andes. Bogotá, Colombia. Correo electrónico: f.gonzalez1899@uniandes.edu.co. ORCID: <https://orcid.org/0000-0001-9445-3809>

† Universidad de los Andes. Bogotá, Colombia. Correo electrónico: s.herrerag@uniandes.edu.co. ORCID: <https://orcid.org/0000-0003-2634-1586>

†† Universidad de los Andes. Bogotá, Colombia. Correo electrónico: t.rodriguez@uniandes.edu.co. ORCID: <https://orcid.org/0000-0002-1887-9119>

††† Universidad de los Andes. Bogotá, Colombia. Correo electrónico: ja.angel908@uniandes.edu.co. ORCID: <http://orcid.org/0000-0002-7210-404X>

§ Universidad Nacional de Colombia. Bogotá, Colombia. Correo electrónico: vcorredore@unal.edu.co. ORCID: <https://orcid.org/0000-0002-8597-8125> (Autor de correspondencia)

§§ Universidad del Rosario. Bogotá, Colombia. Correo electrónico: alejandro.feged@urosario.edu.co. ORCID: <https://orcid.org/0000-0001-8649-4213> (Autor de correspondencia)

Resumen

La tecnología aplicada al uso de grandes datos ha sido una herramienta importante para proveer información necesaria para dar respuestas a los retos de salud pública ante la pandemia de COVID-19, y en Colombia no ha sido la excepción. En este documento se analizan dos insumos de analítica de datos (rastreo digital de contactos y análisis agregado de movilidad, ambos basados en datos de telefonía celular) sobre los cuales se han tomado medidas de salud pública en el país latinoamericano como, por ejemplo, determinar zonas de control diferencial dentro de una ciudad, el rastreo de contactos, y la identificación de potenciales superdispersores. Con base en una muestra de los datos utilizados, se hace una reflexión a la luz de los hallazgos reportados hasta el momento, especialmente desde una perspectiva de redes complejas de contactos y superdispersores, las cuales se ha mostrado tienen un papel crítico en el comportamiento de la epidemia. Los análisis expuestos son parte de una compleja interacción entre el contexto político y epidemiológico que hacen diversas sus respectivas implementaciones. Por último, se resalta que Colombia cuenta con diversos ejemplos de políticas públicas informadas por datos de última generación.

Palabras clave: *políticas públicas; teoría de redes; COVID-19; Bogotá; Colombia.*

Public Policy, Big Data, Network Theory and COVID-19

Abstract

Technology applied to the use of big data has been an important tool to provide necessary information for public health responses to the COVID-19 pandemic, and Colombia has not been the exception. In this document, we analyze two data inputs (digital contact tracking and aggregate mobility analysis, both based on cell phone data), based on which public health measures have been taken in the Latin American country, such as, for example, determining differential control zones within a city, contact tracking, and identification of potential super spreaders. Based on a sample of the data used, we reflect upon the findings reported so far, especially from the perspective of complex networks of contacts and super spreaders, which have been shown to have a critical role in the behavior of the epidemic. The analyses shown here are part of a complex interaction of political and epidemiological contexts which have

led to diverse implementations. We highlight that several examples of public policies in Colombia have been informed by this data.

Keywords: *Public policy; big data; network theory; COVID-19; Bogota; Colombia.*

Políticas públicas, grandes datos, teoría de redes e COVID-19

Resumo

A tecnologia aplicada ao uso de grandes dados tem sido uma ferramenta importante para prover informação necessária para dar respostas aos desafios de saúde pública ante a pandemia de COVID-19, e na Colômbia não tem sido exceção. Neste documento se analisam dois insumos de analítica de dados (rastreo digital de contatos e análise agregado de mobilidade, ambos os dois baseados em dados de telefonia celular) sobre os quais se têm tomado medidas de saúde pública no país latino-americano como, por exemplo, determinar zonas de controle diferencial dentro da cidade, o rastreo de contatos, e a identificação de potenciais superdispersores. Com base em uma amostra dos dados utilizados, se faz uma reflexão à luz dos resultados reportados até agora, especialmente desde uma perspectiva de redes complexas de contatos e superdispersores, as quais se tem mostrado têm um papel crítico no comportamento da epidemia. As análises expostas são parte de uma complexa interação entre o contexto político e epidemiológico que fazem diversas suas respectivas implementações. Finalmente, ressalta-se que a Colômbia conta com diversos exemplos de políticas públicas informadas por dados de última geração.

Palavras-chave: *políticas públicas; teoria de redes; COVID-19; Bogotá; Colômbia.*

Introducción

La crisis mundial desatada por un coronavirus de murciélago, rastreado hasta un mercado de comida marina en Wuhan, China, ha puesto de cabeza el mundo de las políticas públicas (Zhu et al., 2020; Zhou et al., 2020). Todos los países, más allá del espectro político y de la aprobación del gobierno, de su momento económico, y de cualquier

otra caracterización posible, han tenido que atender una emergencia humanitaria en medio de una crisis global cuyo par más cercano ocurrió hace más de cien años. La economía global, una tecnología intangible y planetaria, se encuentra en jaque por un agente infeccioso tan pequeño que no es considerado vivo en un sentido estricto.

La respuesta de los países ha sido tan variada como los sistemas de gobierno, las lenguas y los credos donde el virus encuentra terreno fértil. Al igual que en Egipto durante la Segunda Guerra Mundial, cuando los mosquitos y la malaria brotaron en medio de profundas transformaciones en los sistemas de producción y gobierno (Mitchell, 2002), la pandemia de COVID-19 conlleva una parafernalia de transformaciones sobre las cuales, ante la emergencia, se ha tenido poco tiempo para reflexionar. En apenas unos meses, los gobiernos han improvisado —o se han rehusado a dar— las mejores respuestas que sus funcionarios de salud pública y grupos de expertos han podido diseñar. Comparado con el ritmo paquidérmico en el que operan las políticas públicas en términos de innovación, estas transformaciones se han realizado con celeridad. Los países más preparados, sin duda, han sido aquellos afectados por la última amenaza de pandemia, también por coronavirus, en 2003. En los albores de la epidemia, Japón, Corea del Sur, Singapur, Taiwan, Hong Kong —donde hay más casos de COVID-19— han sido un ejemplo para controlar la dispersión del SARS-CoV-2. La tecnología ha cumplido un papel fundamental para el exitoso control de la enfermedad, y quizá otro tanto hayan influido factores como el acceso a salud, la adopción y acatamiento de medidas preventivas por parte de la población civil, y la desigualdad (Ainslie et al., 2020; Chin et al., 2020; Kraemer et al., 2020; Lai et al., 2020; Lee, Chiew & Khong 2020; Parodi & Liu, 2020; Ting, Carin, Dzau & Wong, 2020; Wang, Ng & Brook, 2020; Wang & Tang, 2020). Países como Estados Unidos, Rusia, Brasil, y Reino Unido encabezan la segunda oleada de casos confirmados, después de que brotes vistos Italia y España, o China, en sus inicios, parecieran un límite superior. En el momento de redactar este documento, Perú supera a China en casos confirmados acumulados y la Organización Mundial de la Salud (OMS) declara que Latinoamérica es el nuevo epicentro de la pandemia.

Uno de los misterios de las primeras etapas del fenómeno global está marcado por paradojas que nadie puede explicar bien: Perú parecía haber respondido rápido y con medidas preventivas efectivas, pero ha visto un crecimiento de casos que lo ubica como el segundo país en una región donde no hay muchos ejemplos de éxito en la contención del virus. Países ricos (Suecia) y pobres (Nicaragua) han visto rebotados sus sistemas de salud. Desde los países con gobiernos populistas de izquierda (Nicaragua, México) y derecha (Estados Unidos, Brasil) o con alto índice democrático (Italia, Francia, Alemania) hasta aquellos que gozan de menos libertades civiles (Rusia) han sufrido por igual a causa de los brotes que se han salido de control.

En este contexto, Colombia no ha tenido —hasta ahora— brotes del virus que desborden el sistema de salud, pero estos tampoco están bajo control. El país entró en cuarentena de manera relativamente temprana (comparado con la respuesta mundial y gracias al ejemplo de países como Estados Unidos, España e Italia) y, a pesar de esto, la transmisión ha mantenido una tendencia de crecimiento de casos y muertes. La pandemia ha significado retos para la división político-administrativa (diferencias entre alcaldes y el gobierno nacional, retos de ordenamiento territorial en zonas metropolitanas), para la comunidad académica (representada, pero no limitada, por el debate entre economía y cuarentena), entre muchos otros.

La respuesta más efectiva ante la COVID-19 parece haber estado relacionada con el uso de tecnologías digitales para vigilancia epidemiológica. Países como Singapur, Corea del Sur, Taiwan y China, todos cuentan con ejemplos de innovaciones tecnológicas para el rastreo de contactos y el uso de grandes datos para toma de decisiones en salud pública relacionada con la pandemia (Lee et al., 2020; Parodi & Liu, 2020; Ting et al., 2020; Wang et al., 2020). En general, hay tres aspectos críticos que han sido resaltados para entender tanto la heterogeneidad de la epidemia por regiones, como la eficacia de las acciones llevadas a cabo por gobiernos (y la respectiva adherencia de sus poblaciones): primero, la heterogeneidad del contagio, o la importancia de los superdispersores del virus. Segundo, las congregaciones

que facilitan la dispersión (eventos atípicos en contexto de cuarentena), como celebraciones religiosas, funerales, carnavales, entre otros. Tercero, la conectividad regional e internacional (Hébert-Dufresne et al., 2020; MacIntyre, 2020; Shim, Tariq, Choi, Lee & Chowell, 2020).

Objetivo

Este texto pretende hacer una reflexión sobre algunos temas críticos (heterogeneidad del contagio, eventos de contacto atípicos, y movilidad) en el marco de la respuesta de Colombia a la pandemia y el proceso de toma de decisión en salud pública. Para ello, se presentan resultados de análisis de datos utilizados en el marco de las diferentes intervenciones implementadas en Colombia, y se discuten su alcance y pertinencia. Esto incluye:

- Caracterizar el análisis de grandes datos en términos de super-dispersión, eventos atípicos de contacto, y movilidad intra e inter focos epidémicos, con el fin de ilustrar los insumos que tienen las políticas de salud pública relacionadas con la COVID-19.
- Analizar la toma de decisión basada en los insumos discutidos anteriormente, con el fin de discutir la efectividad de las políticas basadas en datos en contexto de la COVID-19.
- Discutir y reflexionar sobre el papel de las políticas en salud pública (vigilancia epidemiológica basada en grandes datos) en el contexto colombiano.

Aspectos importantes para tener en cuenta al pensar en la COVID-19 desde los grandes datos

Dos tipos de datos son de particular interés: los datos de movilidad en unidades espaciales definidas que permiten monitorear los cambios en los flujos de movilidad y los datos de los contactos de infecciones activas para poder tomar medidas de aislamiento de contactos.

Ambos tipos de datos pueden obtenerse a partir del rastreo de teléfonos celulares y proporcionan información invaluable para la toma de decisiones en política pública.

Los datos agregados y anonimizados de movilidad proporcionados por los operadores privados permiten obtener información casi en tiempo real de los patrones de movilidad humana (Buckee et al., 2020). De esta manera, es posible, primero, evaluar el nivel de acatamiento de las medidas asociadas a la cuarentena y el distanciamiento social a través del monitoreo cuantitativo de cambios (reducciones o aumentos) de movilidad al interior de unidades espaciales o entre unidades espaciales. Segundo, identificar mecanismos y consecuencias de las medidas adoptadas. Tercero, predecir la tendencia, en periodos cortos de tiempo, de cómo progresa la enfermedad en relación con los cambios de movilidad comparados con situaciones similares ocurridas previamente en otros lugares del mundo, una aproximación propensa a la automatización por medio del entrenamiento de algoritmos o *machine learning*. Cuarto, delimitar de manera cuantitativa unidades espaciales de conectividad (o aislamiento) de acuerdo con los niveles de conectividad reflejados en los flujos de movilidad. La integración de este tipo de datos con información epidemiológica (p. ej. número de casos, utilización hospitalaria) y demográfica (p. ej. ingreso per cápita; pirámide poblacional) permite la elaboración de políticas dirigidas a distintos sectores de la población.

Los datos anonimizados de contactos de una persona contagiada o de una persona con síntomas y su red de contactos identificados mediante datos de cercanía de celulares en una determinada unidad de tiempo permiten identificar y hacer el rastreo y aislamiento de casos sospechosos (Kishore et al., 2020; Oliver et al., 2020). Estos datos son particularmente útiles en la medida en que permiten el aislamiento de casos y el rastreo de la cadena de contactos. Este tipo de información (información de redes) puede, además, ser utilizada, potencialmente, para determinar parámetros de importancia en la progresión de la infección que pueden utilizarse en modelos mecánicos de la infección (Chinazzi et al., 2020).

Un problema con respecto a los modelos que tratan de simular las condiciones en que progresa la infección y predecir su tamaño es que muchos de los parámetros sobre los cuales depende el evento son inciertos o mal comprendidos (Holmdahl & Buckee, 2020). Uno de ellos es, precisamente, la movilidad humana, vista como la probabilidad de introducción de un evento infeccioso en una población y el carácter estocástico del evento infeccioso, en el sentido en que un individuo puede contagiar un número muy grande de personas mientras que otros, por su parte, pueden no contagiar a nadie (Hébert-Dufresne, Althouse, Scarpino & Allard, 2020). Este efecto refleja, por lo tanto, la heterogeneidad de la distribución de las infecciones secundarias. Los datos de rastreo de la cadena de contactos permiten obtener datos acerca de esta distribución, incorporarlos eventualmente en los modelos y así disminuir la incertidumbre en la estimación del tamaño de la infección (Hébert-Dufresne et al., 2020).

Métodos

El análisis de datos presentado en este artículo es una muestra de los informes y productos que se han desarrollado para diferentes entes territoriales y de salud pública en Colombia. Hay dos fuentes de datos principales: datos de movilidad, agregados y dispuestos en forma de Infraestructura como Servicio (IAAS, por su sigla en inglés) en la plataforma GeoInsights de Facebook. Estos datos están compuestos por diferentes vectores que denotan movimiento entre dos puntos georreferenciados en intervalos de ocho horas, con información correspondiente a la línea base de personas, el valor observado para ese intervalo de tiempo, y algunos cálculos estadísticos basados en estos valores y la distribución (cambio porcentual, valor-z, significancia, diferencia, entre otros). Adicionalmente está disponible un conjunto de datos que caracteriza el movimiento para cada municipio con base en dos variables: cambio porcentual en número de cuadrantes visitados por usuario, y proporción de usuarios que visitaron un solo cuadrante. Con los datos agregados de movilidad entre puntos geográficos se construyó una red pesada en la cual los nodos son municipios y las aristas el número de personas que se movilizan entre municipios.

A esta representación en red de la movilidad se le hizo un análisis de estadística descriptiva; asimismo, se realizó un ejercicio de detección de comunidades. Finalmente, para algunas comunidades se implementó un algoritmo de predicción a corto plazo (*nowcasting*) basado en datos para evaluar la evolución de la epidemia en los siguientes nueve días. Dado que la epidemia en Colombia llevaba un retraso con respecto a otros lugares del mundo, para evaluar las dinámicas con más información, se implementó un algoritmo de búsqueda de dinámicas similares (*Dynamic Time Warping*, DTW, por su sigla en inglés) para todo el mundo, con la cual se entrenó un algoritmo de inteligencia artificial que para un conjunto de datos de movimiento interno y externo, y una distribución espacial de casos, proyectara en el tiempo los casos para nueve días.

Por otro lado, gracias a la colaboración con la empresa Servinformación, se encuentra disponible un conjunto de datos de movilidad de teléfonos celulares que provienen de un servicio de agregación de aplicaciones para minería de datos sobre usuarios con fines de mercadeo digital disponibles a nivel mundial. Estos datos son anónimos y están disponibles comercialmente; asimismo, han sido utilizados en el pasado con fines de investigación para caracterizar la movilidad de las personas. El conjunto de datos cuenta con un ID anonimizado para cada usuario y coordenadas que indican desde dónde interactuó con diferentes aplicaciones. Con base en estos datos, se construyó una matriz de distancia entre usuarios que compartieron el mismo tiempo y espacio, que puede ser representada como una red de contactos, donde A está conectado con B si se encontraron en el mismo lugar y en tiempos similares —con una diferencia entre las dos observaciones no mayor a diez minutos, definida arbitrariamente—. Para esta red, construida para intervalos de días y semanas, se implementaron métricas de red básicas, como centralidad de grado, de valor propio y *pagerank*, así como modularidad y otros algoritmos para detección de comunidades.

Si bien ambos datos pueden contener sesgos importantes por acceso a tecnología en dispositivos móviles, también han sido utilizados exitosamente para caracterizar la movilidad desde un interés académico y comercial, y en menor medida, para la formulación de políticas

públicas. Los datos de morbilidad y mortalidad por COVID-19 están disponibles a nivel municipal para Colombia para el público general y se pueden descargar directamente y en su versión más actualizada desde la página del Instituto Nacional de Salud (INS) de Colombia.

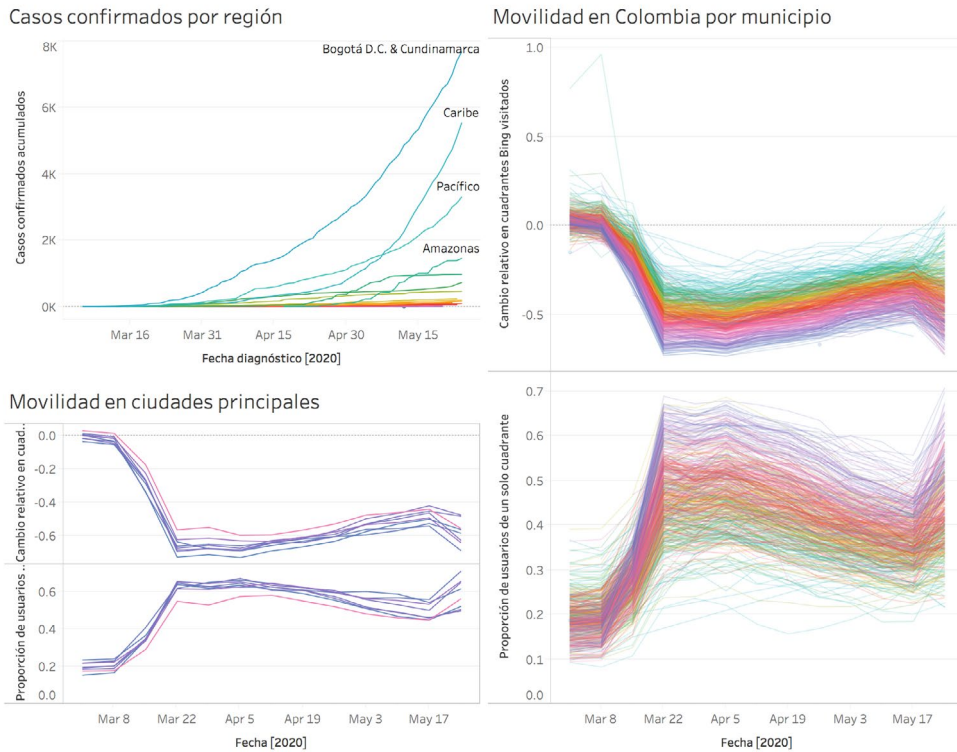
Resultados del análisis de grandes datos, insumos para políticas públicas

Los reportes producidos para tomadores de decisión en políticas públicas contienen información similar a la presentada en esta sección. Esta información hace parte de una amplia gama de soportes —muchos de los cuales no están contenidos en este documento pues desborda su objetivo— con los que se toman decisiones a escalas sin precedentes en la historia reciente. Cerrar fronteras y vuelos internacionales, decretar cuarentenas nacionales, cerrar sectores de una ciudad o de la economía, relajar dichas restricciones, y desplegar equipos y recursos como epidemiología de campo, todas estas son decisiones que se toman con base en, entre otras, esta información.

Movilidad

El gobierno nacional decretó cuarentena obligatoria el 24 de marzo, cinco días después de que varios alcaldes de algunas de las ciudades más importantes del país iniciaran un simulacro. La movilidad del país se redujo significativamente, como lo muestran las diferentes representaciones de la figura 1, como consecuencia de las medidas preventivas implementadas por el gobierno (primero el simulacro a nivel subnacional y, luego, el decreto por el cual se reglamenta la restricción de movilidad). Los datos muestran, además, que la cuarentena, lentamente, ha perdido su adherencia en la población. En general, vemos que las ciudades principales (en colores púrpura y rosa) han reducido mucho más la movilidad que otros territorios con ciudades más pequeñas o de vocación rural (azul, verde, amarillo, naranja y rojo).

Figura 1. Movilidad y casos confirmados de Covid-19 en Colombia por región y ciudades principales



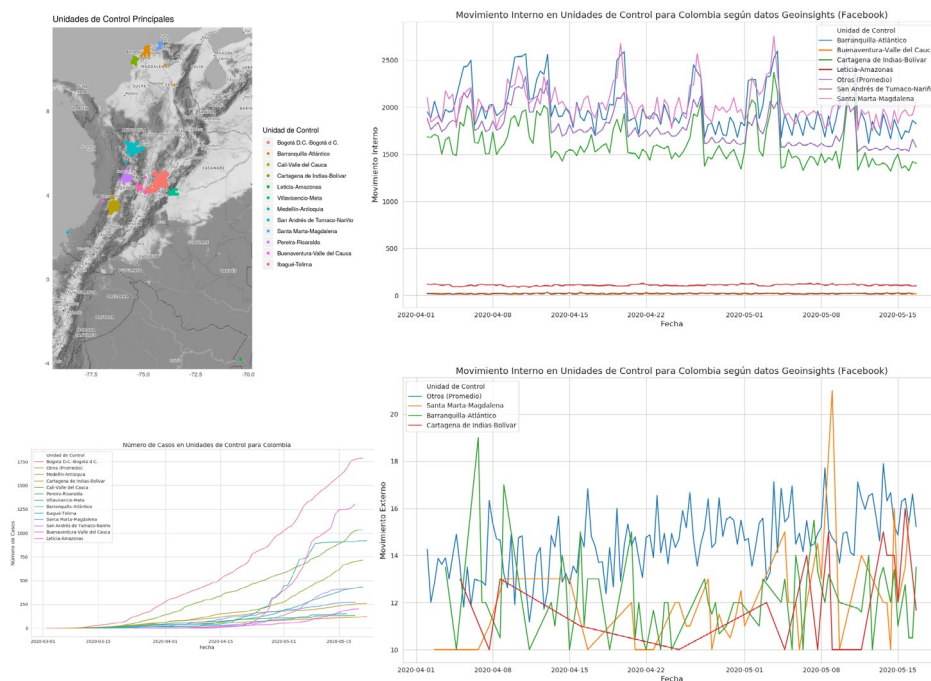
Arriba, izquierda: Casos confirmados de COVID-19 para Colombia, por región. Abajo, izquierda: Movilidad para algunas ciudades de Colombia (en orden decreciente: Barrancabermeja, Santa Marta, Medellín, Bucaramanga, Cali, Cartagena, Barranquilla y Bogotá). La parte superior muestra el cambio relativo en el número de cuadrantes visitados. La parte inferior corresponde a la fracción de usuarios que visitaron solamente un cuadrante. Derecha: Movilidad para Colombia por municipio. Cada línea representa un municipio diferente. Se puede ver que la movilidad se redujo en más del 50 % con el inicio de la cuarentena y, desde entonces, ha subido progresivamente en todo el país. La parte superior muestra el cambio relativo en el número de cuadrantes visitados. La parte inferior corresponde a la fracción de usuarios que visitaron solamente un cuadrante.

Fuente: Elaboración de los autores con base en los datos de movilidad de GeoInsights de Facebook. Los datos sobre los casos confirmados con COVID-19 son tomados del Instituto Nacional de Salud, que los publica diariamente.

La misma información, pero agregada a nivel de unidades funcionales (generadas con el algoritmo de detección de comunidades en redes de movimiento) (Ross, 2012 ; Nelson & Rae, 2016), permite caracterizar el riesgo territorialmente. En la figura 2 se ve cómo la movilidad

interna y externa caracteriza el riesgo en las regiones en términos de importar casos, contagio interno, o la combinación de ambas.

Figura 2. Movilidad y casos confirmados de COVID-19 en Colombia por unidades funcionales



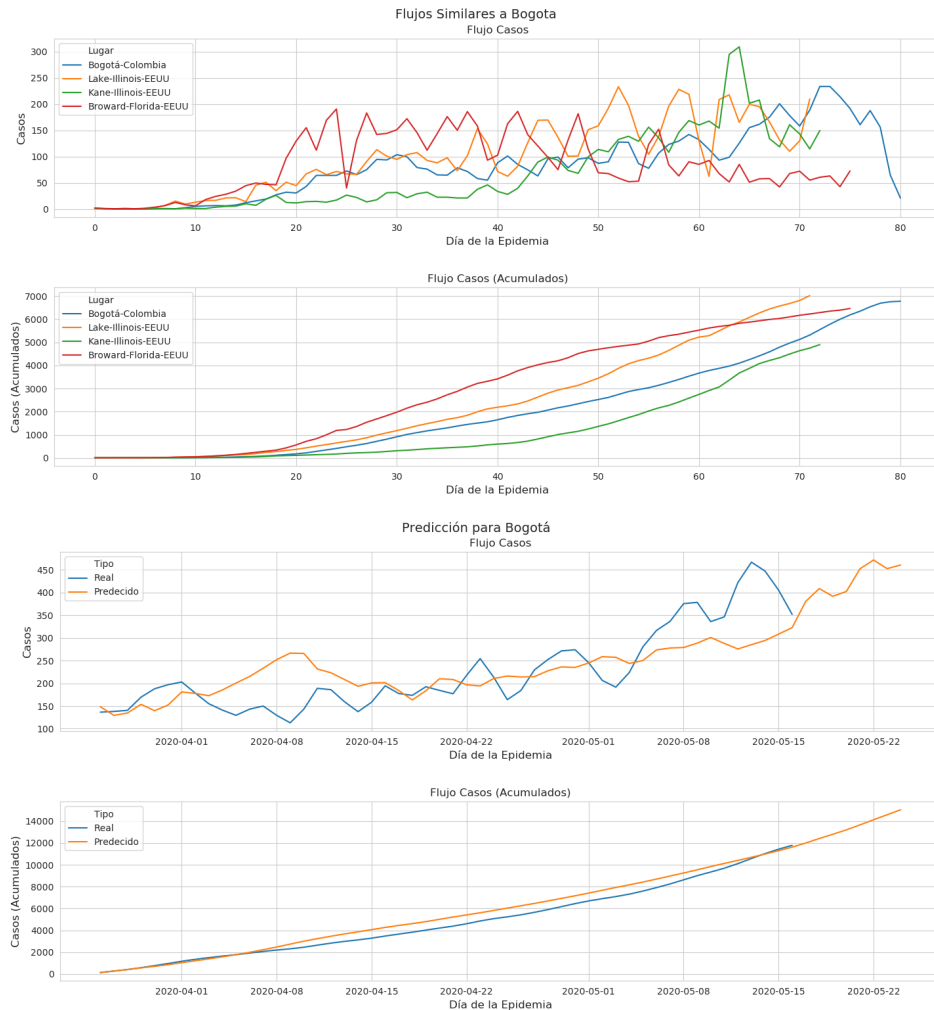
Arriba, izquierda: Unidades funcionales detectadas por el algoritmo de detección de comunidades con base en los datos de movilidad de GeoInsights de Facebook. Abajo, izquierda: Casos confirmados con COVID-19 por unidad funcional. Arriba, derecha: movimiento interno para cada una de las unidades funcionales. Abajo, derecha: Movimiento externo entre unidades funcionales.

Fuente: elaboración de los autores con base los datos de GeoInsights de Facebook y el Instituto Nacional de Salud.

Con esta información se pueden hacer estimaciones para los casos en un futuro cercano, entrenando el algoritmo con las dinámicas de regiones con proyecciones similares en otros lugares del mundo. La figura 3 muestra las series vecinas a la progresión de casos y movilidad de Bogotá, y la predicción que se hace para la capital en el con-

texto global de la pandemia, pero hecho específicamente para sus características.

Figura 3. Predicción de casos de COVID-19 para Bogotá basada en datos

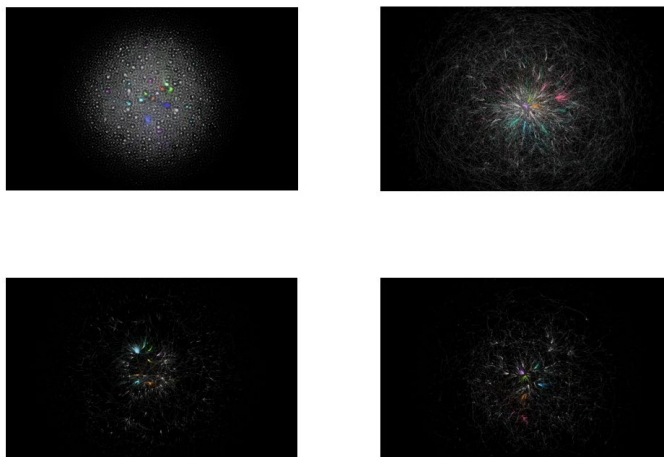


Con base en el algoritmo DTW se identifican regiones de otros países con tendencias similares a las presentadas por Bogotá (se toma como ejemplo dado que es la región con más casos de Colombia y, por lo tanto, la que produce resultados con mejor ajuste).

Fuente: elaboración de los autores con base en los datos de Servinformación.

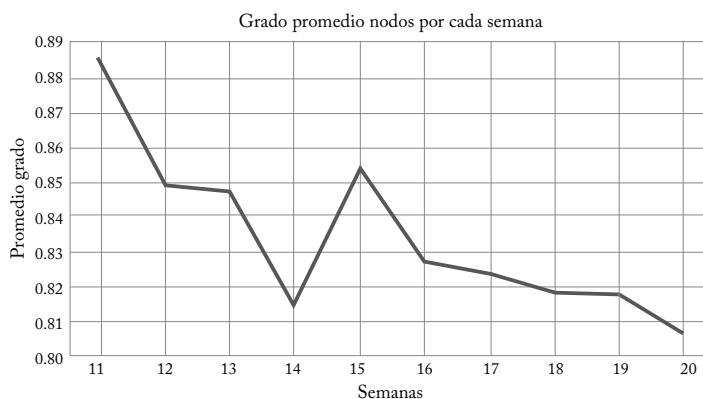
Redes

En Colombia, el trabajo de campo epidemiológico y el rastreo digital de contactos ha revelado aspectos importantes de la epidemia. En Bogotá, uno de los brotes más fuertes hasta la fecha ha sido el de la localidad Kennedy, pero los primeros casos de la ciudad —y del país, dicho sea de paso— tuvieron lugar en las localidades de Usaquén y Chapinero. ¿Cómo y por qué el virus migró a Kennedy (también Suba y Engativá han sido afectadas por la segunda oleada) y se dispersó por toda la ciudad? El evento de dispersión a otras zonas probablemente tuvo lugar alrededor de la semana nueve, por lo cual decidimos explorar las dinámicas de contactos alrededor de esas fechas. La figura 4 y figura 5 muestran el cambio en la red de contactos en el tiempo y, en particular, para la semana en cuestión (también se incluye la gráfica de la semana diez). En esta gráfica podemos ver cómo un evento durante esa semana cambió drásticamente la red de contactos observada para todo el periodo general. Al explorar la naturaleza de este fenómeno, encontramos que dicho incremento en las tasas de contacto (eventos asociados a la transmisión del virus en múltiples contextos) tuvo lugar en proximidad a la plaza de Lourdes, en la calle 63 con carrera 9, para ser más precisos. Esta es una zona de bastante tráfico, pero donde además hay varios establecimientos nocturnos. Sobre todo, este método muestra que los datos de proveedores para mercadeo digital pueden utilizarse para detectar contextos sociales asociados a la transmisión del virus casi en tiempo real, y para responder preguntas retrospectivas con el fin de entender procesos que facilitan o previenen la dispersión del virus.

Figura 4. Red de contactos para Bogotá en el tiempo

Arriba, izquierda: Esta visualización permite ver todos los contactos observados para todos los dispositivos de los cuales se tiene registro en una red cuyos contactos están distribuidos en toda la red (no hay un solo individuo que tenga un número mucho más alto de contactos que el resto de la red, a diferencia, por ejemplo, de la red creada por vínculos de páginas web). Se pueden observar algunos *clusters*, resaltados en colores, que representan pequeñas comunidades en las cuales los individuos tienen más contactos entre sí que con personas fuera de la comunidad. Arriba, derecha: red de contactos para la semana nueve, con los mismos métodos usados para construir la red completa. Se observa un evento atípico, marcado por una comunidad de personas que tuvieron contacto con un dispositivo. Este evento tuvo lugar antes de que fuera instaurada la cuarentena. Abajo, izquierda: red para la semana diez. Abajo, derecha: Red de contactos para Bogotá en la semana diecisiete, en la cual se presentó un brote en Corabastos.

Fuente: elaboración de los autores con base en los datos de Servinformación.

Figura 5. Serie de tiempo de los grados promedio de las redes por semana

Esta serie de tiempo permite mostrar, en promedio, cuántos contactos tienen las personas cada semana. Las semanas se codifican de acuerdo al estándar ISO-8601. En la gráfica se puede observar una tendencia decreciente entre la semana once y la semana veinte.

Fuente: elaboración de los autores con base en datos de Servinformación.

Reflexiones sobre datos y políticas públicas

Sin duda, uno de los temas más visibles como medida de salud pública en Colombia ante la epidemia de COVID-19 ha sido el rastreo digital de contactos, que ha seguido los métodos acá descritos. Hasta ahora, desde nuestro conocimiento, tres entidades territoriales han implementado estas medidas: Bogotá, Valle del Cauca y Antioquia, con resultados bastante heterogéneos. ¿Por qué en un lugar (Antioquia) sí ha funcionado el empoderamiento tecnológico de la salud pública, mientras que en los otros no parece haber dado resultados tan impactantes? En diversas comunicaciones, tanto la gobernación como la alcaldía de Medellín han dado parte de victoria sobre el uso de tecnología, mientras que Cali y Bogotá han visto la epidemia crecer, probablemente a una tasa más baja que si no contaran con medidas de confinamiento.

Es importante hacer dos aclaraciones. La primera es que Bogotá ha implementado un modelo de confinamiento diferencial basado, entre otras cosas, en datos de rastreo digital. Mientras tanto, Antioquia ha fortalecido su sistema de seguimiento epidemiológico de campo con la tecnología como insumo, con gran interés en la epidemiología de campo.

Finalmente, es importante señalar la importancia de que las decisiones de política pública sean tomadas a partir del análisis de datos y con criterios técnicos que reúnan la opinión de expertos. Las decisiones en varios países que han despreciado las opiniones calificadas han resultado costosas tanto en vidas humanas como social y económicamente. También, es relevante resaltar que la diversidad de implementaciones en políticas públicas para un mismo análisis de datos muestra un panorama alentador para Colombia, especialmente en un contexto político donde algunos líderes mundiales han apoyado medidas que más bien han promovido la dispersión del virus. En Colombia, el análisis de datos de última generación ha influido directamente en políticas públicas de respuesta a la pandemia y no como un hecho aislado.

Este ejemplo ilustra que los datos en sí mismos no son una política pública sino un insumo. Más allá de comparar la evolución del brote en Medellín y Bogotá —algo que consideramos a todas luces inapropiado por múltiples razones— este ejemplo ilustra que el análisis de datos es solo un eslabón más en una larga cadena de procesos que componen una política pública. El éxito y su pertinencia no solamente está dada por las características de los datos y la calidad del análisis, sino también por el contexto cultural y político en el cual están inmersos. En otras palabras, los grandes datos en políticas públicas no pueden ser entendidos como un atajo para superar dificultades en un entorno específico, sino, por el contrario, como producto de dicho contexto y que nutre las prácticas de políticas públicas previas.

Por otro lado, los datos de movilidad que componen tanto las unidades funcionales para el control de COVID-19 en Colombia como su respectivo perfil de movilidad y casos presentan desafíos de otra índole. Si bien es cierto que han sido utilizados como herramienta para evaluar el grado de cumplimiento de medidas de distanciamiento social, en términos de políticas públicas, si se ha podido aprovechar más este conjunto de datos, es poco. En cualquier caso, esta información ha sido un insumo clave para alimentar modelos de simulación con los que posteriormente se toman decisiones, así como han alimentado el análisis que derivó en la creación de las zonas de cuidado especial en Bogotá —aunque el insumo principal en términos de grandes datos ha sido la red de contactos—. Esto no significa que la información no sea útil, sino que uno de los retos que existe es entender la manera como se pueden incorporar divisiones político-administrativas de manera más fluida en el ordenamiento territorial y del gasto, así como el potencial que tienen para identificar condiciones de vulnerabilidad.¹ Sin duda, atender este reto será de vital importancia de cara a la siguiente etapa de la epidemia: el relajamiento de las medidas de distanciamiento social para aliviar la crisis económica derivada de las medidas de salud pública.

¹ De hecho, se ha reportado en todo el mundo que los sectores más vulnerables de una ciudad son los que menos posibilidad tienen de resguardarse y de cumplir la cuarentena.

Referencias

- Ainslie, K. E. C., Walters, C. E., Fu, H., et al. (2020). Evidence of initial success for China exiting COVID-19 social distancing policy after achieving containment. *Wellcome Open Research*, 5(81). <https://doi.org/10.12688/wellcomeopenres.15843.1>
- Buckee, C. O., Balsari, S., Chan, J., et al. (2020). Aggregated mobility data could help fight COVID-19. *Science*, 368(6487), 145-146. <https://doi.org/10.1126/science.abb8021>
- Chin, T., Kahn, R., Li, R., Chen, J. T., Krieger, N., Buckee, C. O., Balsari, S., & Kiang, M. V. (2020). US county-level characteristics to inform equitable COVID-19 response. *medRxiv*. <https://doi.org/10.1101/2020.04.08.20058248>
- Chinazzi, M., Davis, J. T., Ajelli, M., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, 368(6489), 395-400. <https://doi.org/10.1126/science.aba9757>
- Hébert-Dufresne, L., Althouse, B. M., Scarpino, S. V., & Allard, A. (2020). Beyond R_0 : Heterogeneity in secondary infections and probabilistic epidemic forecasting. *medRxiv*. <https://doi.org/10.1101/2020.02.10.20021725>
- Holmdahl, I. & Buckee, C. (2020). Wrong but useful—what covid-19 epidemiologic models can and cannot tell us. *New England Journal of Medicine*. <https://doi.org/10.1056/NEJMp2016822>
- Kishore, N., Kiang, M. V., Engø-Monsen, K., Vembar, N., Balsari, S., & Buckee, C. (2020). Mobile phone data analysis guidelines: applications to monitoring physical distancing and modeling COVID-19. <https://doi.org/10.31219/osf.io/5arjy>
- Kraemer, M. U., Yang, C.-H., Gutierrez, B., et al. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490), 493-497. <https://doi.org/10.1126/science.abb4218>
- Lai, S., Ruktanonchai, N. W., Zhou, L., et al. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*. <https://doi.org/10.1038/s41586-020-2293-x>
- Lee, V. J., Chiew, C. J., & Khong, W. X. (2020). Interrupting transmission of COVID-19: lessons from containment efforts in Singapore. *Journal of Travel Medicine*, 27(3): taaa039. <https://doi.org/10.1093/jtm/taaa039>

- MacIntyre, C. R. (2020). Global spread of COVID-19 and pandemic potential. *Global Biosecurity*, 1(3). <https://doi.org/10.31646/gbio.55>
- Mitchell, T. (2002). *Rule of experts: Egypt, techno-politics, modernity*- University of California Press. <https://www.jstor.org/stable/10.1525/j.ctt1ppnxp>
- Nelson, G. D., & Rae, A. (2016). An Economic Geography of the United States: From Commutes to Megaregions. *PLoS one*, 11(11): e0166083. <https://doi.org/10.1371/journal.pone.0166083>
- Oliver, N., Lepri, B., Sterly, H., et al. (2020). Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Science Advances* 6(23): eabc0764. <https://doi.org/10.1126/sciadv.abc0764>
- Parodi, S. M., & Liu, V. X. (2020). From containment to mitigation of COVID-19 in the US. *JAMA*, 323(15), 1441-1442. <https://doi.org/10.1001/jama.2020.3882>
- Ross, C. L. (Ed.). (2012). *Megaregions: Planning for global competitiveness*. Washington: Island Press.
- Shim, E., Tariq, A., Choi, W., Lee, Y., & Chowell, G. (2020). Transmission potential and severity of COVID-19 in South Korea. *International Journal of Infectious Diseases*, 93, 339-344. <https://doi.org/10.1016/j.ijid.2020.03.031>
- Ting, D. S. W., Carin, L., Dzau, V., & Wong, T. Y. (2020). Digital technology and COVID-19. *Nature medicine*, 26(4), 459-461. <https://doi.org/10.1038/s41591-020-0824-5>
- Wang, C. J., Ng, C. Y., & Brook, R. H. (2020). Response to COVID-19 in Taiwan: Big Data Analytics, New Technology, and Proactive Testing. *JAMA*, 323(14), 1341-1342. <https://doi.org/10.1001/jama.2020.3151>
- Wang, Z., & Tang, K. (2020). Combating COVID-19: health equity matters. *Nature Medicine*, 26(4), 458. <https://doi.org/10.1038/s41591-020-0823-6>
- Zhou, P., Yang, X.-L., Wang, X.-G., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 579(7798), 270-273. <https://doi.org/10.1038/s41586-020-2012-7>
- Zhu, N., Zhang, D., Wang, W., et al. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* 382:727-733. <https://www.doi.org/10.1056/NEJMoa2001017>