

# Estadísticas y Análisis del Covid-19 en Ecuador Utilizando Microsoft Power BI

Statistics and Analysis of Covid-19 in Ecuador Using Microsoft Power BI

Vivas K.<sup>1</sup>; Vivas D.<sup>2</sup>

<sup>1</sup> Universidad de las Fuerzas Armadas ESPE, Departamento de Ciencias Exactas, Quito-Ecuador  
e-mail: akvivas@espe.edu.ec

<sup>2</sup> Universidad San Francisco de Quito, Facultad de Ingeniería, Quito-Ecuador  
e-mail: dvivasp@asig.edu.ec

Información del artículo

Recibido: 13/01/2021

Aceptado: 17/03/2021

DOI <https://doi.org/10.29166/ingenio.v4i1.3068>

## RESUMEN

Dada la emergencia que el Ecuador y el mundo están atravesando a causa del covid-19, se vuelve indispensable la descripción rápida y el análisis efectivo de las diferentes variables que explican la evolución de una enfermedad contagiosa. El propósito de este trabajo es mostrar la versatilidad que Microsoft Power BI como herramienta de inteligencia de negocios posee para describir y analizar la información emitida por las entidades oficiales de salud que permitan generar un rápido conocimiento y análisis de los datos de modo que se puedan tomar decisiones correctas respecto al control de la pandemia. Para esto, se recolectó la información oficial que provee el Ministerio de Salud Pública del Ecuador (MSP) y otros portales de noticias. Esta información fue procesada y analizada mediante la plataforma Power BI y el software estadístico R. El análisis de los datos recientes sobre covid-19 muestran que la enfermedad aún se expande en algunas localidades del país. La propuesta sería dar más y mejor acceso a datos abiertos e innovar agregando inteligencia de negocios para controlar esta pandemia.

**Palabras clave:** *covid-19, Power BI, estadística, datos, pandemia.*

## ABSTRACT

Given the emergency that Ecuador and the world are going through due to covid-19, the rapid description and effective analysis of the different variables that explain the evolution of a contagious disease becomes essential. The purpose of this work is to show the versatility that Microsoft Power BI as a business intelligence tool has to describe and analyze the information issued by official health entities that allow generating rapid knowledge and analysis of the data so that it can be taken correct decisions regarding pandemic control. For this, the official information provided by the Ministry of Public Health of Ecuador (MSP) and other news portals was collected. This information was processed and analyzed using the Power BI platform and R software. Analysis of recent data on covid-19 shows that the disease is still spreading in some localities of the country. The proposal would be to give more and better access to open data and innovate by adding business intelligence to control this pandemic.

**Keywords:** *covid-19, Power BI, statistics, data, pandemic.*

## 1. Introducción

El 29 de febrero del 2020 se confirma el primer caso de covid-19 en Ecuador, la enfermedad se expandió rápidamente, de modo que para el 1.º de mayo se acumularon más de 26.000 casos y al día de realización de este trabajo (21 de noviembre) ya son más de 186.000 [1]. Además, de acuerdo a los análisis a nivel regional, Ecuador presenta el índice de letalidad más alto en Sudamérica [2]. Las principales ciudades como Guayaquil y Quito, tienen un crecimiento alarmante en el número de nuevos casos. La región Costa y Sierra son las más afectadas mientras que la región Oriente e Insular presentan menos casos.

Los países de la región han adoptado medidas estrictas desde el principio para contener la primera ola de covid-19, incluidos bloqueos nacionales, toques de queda generales y focalizados, restricciones vehiculares, sin embargo, estas medidas no alcanzaron para contener la enfermedad. Las medidas se vieron socavadas por la fragilidad existente de los sistemas de salud, que se caracterizan por una inversión insuficiente en recursos sanitarios, disparidades regionales, sistemas de información modestos, corrupción (en el caso de Ecuador) y una comunicación y coordinación deficientes. Fuimos testigos de que nuestros hospitales y centros de salud se desbordaron en los primeros 3 meses de la pandemia, las UCI (unidad de cuidados intensivos), alcanzaron una ocupación del 100% en algunas localidades, tanto así que muchas personas tuvieron que viajar a otras ciudades para encontrar atención médica [3].

Uno de los factores que permiten el control de una epidemia es la recolección, procesamiento y visualización de los datos que representan las variables de la enfermedad. El Comité de Operaciones de Emergencia Nacional (COE) junto con el Ministerio de Salud Pública (MSP) son las entidades oficiales encargadas de manejar la información generada por el proceso y las distintas variables de la pandemia. Se vuelve necesaria la descripción, procesamiento y el análisis de la información obtenida, de modo que se pueda conocer de forma continua, rápida y efectiva la visualización de las diferentes variables con el fin de proveer resultados que sirvan para tomar decisiones y acciones pertinentes por parte de los las personas y entes gubernamentales.

Dado que el número de personas infectadas está aumentando muy rápidamente a nivel nacional y en todo el mundo, es necesario obtener nuevas soluciones a los problemas emergentes de gestión de la atención médica para hacer frente a la epidemia y evitar el colapso de los sistemas de salud. En este sentido, la Inteligencia de Negocios (BI) puede analizar datos en tiempo real sobre cómo se está propagando la epidemia y dónde están los focos principales. Estos datos son fundamentales para la respuesta rápida de la atención médica [4] the first patient was tested positive for SARS-CoV-2 at Codogno hospital in the Lombardy region. From that date, the Regional Emergency Medical Services (EMS). Hay evidencia de que la BI, disponible para los líderes y gerentes del sistema de salud, serviría la toma de decisiones que resulte en un mejor desempeño organizacional [5].

Este estudio tiene como objetivo presentar un análisis de los principales indicadores y variables para la pandemia del covid-19 en Ecuador tomando como fuente las cifras oficiales brindadas por el Ministerio de Salud Pública del Ecuador [1] y el portal periodístico Primicias Ec. [6]. Se utilizó como herramientas: la plataforma de Microsoft Power BI y el software R que proporcionan principalmente información visual efectiva e interactiva a partir de los datos, son de fácil utilización y principalmente usada para inteligencia de negocios (*business intelligence*). Para esto, Power BI contiene un conjunto de herramientas y servicios de análisis empresarial que acceden a fuentes de datos para procesarlos con el fin de realizar análisis y visualización de los mismos [7].

La vigilancia de enfermedades constituye la base de la respuesta a las epidemias. La pandemia del covid-19 proporciona un ejemplo actualizado de por qué se deben tomar las consideraciones que emite la frase «persona, lugar y tiempo» y que son cruciales para su control, esto se debe a que la epidemia requiere conocer las tendencias de las variables de la enfermedad en diferentes subgrupos y ubicaciones [8].

## 2. Materiales y métodos

El Ministerio de Salud Pública (MSP) provee de boletines informativos que nos permiten actualizar la información. De esta entidad se tomó el

reporte actualizado al 21 de noviembre 2020. La información presenta el número de casos confirmados de covid-19 en el país por cada día, casos confirmados en las provincias por día, casos totales al 21 de noviembre en cantones, el número de fallecidos confirmados y probables en el país, grupos etarios y casos por sexo. Esta información fue recolectada en documentos Excel de modo que sea fácilmente manejable por Power BI. Los datos de casos acumulados para las provincias de interés fueron tomados del portal Primicias, desde el 13 de marzo hasta el 15 de noviembre.

## 2.1 Microsoft Power BI

Microsoft Power BI es una plataforma de análisis e inteligencia empresarial que consta de aplicaciones y servicios diseñados para proporcionar información visual coherente e interactiva sobre los datos, más específicamente, es un conjunto de herramientas y servicios de análisis empresarial que funcionan en conjunto para acceder a fuentes de datos, dar forma a las diferentes fuentes, realizar análisis y visualización de los mismos y compartir conocimientos a través de estos procesos. Cabe mencionar que para cada proyecto se considera un conjunto específico de herramientas, sin embargo, la dinámica y versatilidad de estas siempre están disponibles al utilizar las características únicas de múltiples herramientas como parte de soluciones integradas que abarcan diversas fuentes de datos y tipos de visualización [7].

Los componentes de Power BI se dividen en dos grupos [9]: Complementos de Excel y Servicios en la Nube que están disponibles mediante Office 365.

### Los complementos de Excel

Power Query, que permite conectar a una amplia variedad de tipos de fuentes de datos para luego limpiarlos, filtrarlos o transformar los datos que serán tratados en Excel.

Power Pivot, que es el motor de la base de datos y usa los modelos de datos de Excel que permiten tomar grandes volúmenes de datos (mucho más que en una plantilla de Excel sola).

Power View, que es la reciente característica de Excel que permite crear tableros (*dashboards*) atractivos e interactivos que incluyen tablas,

gráficos y mapas de fácil manejo para el creador y el usuario.

Power Map, que es el complemento que permite trabajar con datos geográficos para crear mapas 2-D o 3-D.

### Los servicios en la nube

Mobile BI, que es una aplicación Windows 8.1 que nos permite observar partes de un trabajo realizado en Power BI como tablas, gráficos, tableros, entre otros.

Power BI Data Catalogue, permite a los administradores el acceso a una gran variedad de datos públicos o corporativos.

Power BI Sites, que permite una organización más profesional de los reportes o tableros y manejarlos mientras sus colegas también los revisan.

Se puede observar que la arquitectura y los componentes principales (Fig. 1) se relacionan de modo que el desarrollador este siempre en contacto con las fuentes de información (Power BI puede leer una gran cantidad de tipos de datos) que pueden venir de fuentes diversas, cuenta con variadas aplicaciones extra que pueden incorporarse fácilmente al sistema, existe además la conexión en línea para otros usuarios que deseen acceder a la información o sus resultados por medio de consultas.

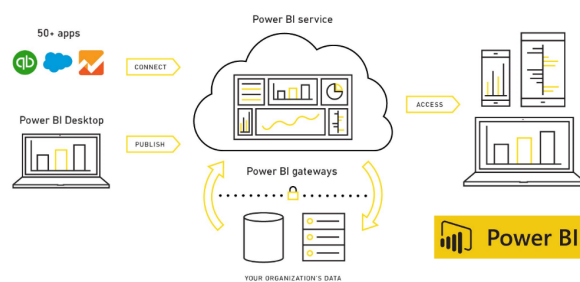


Figura 1. Arquitectura y componentes principales de Power BI

Fuente: [Microsoft \[10\]](#).

Adicionalmente, se debe tomar en cuenta los modelos de datos que Power BI genera automáticamente o que se pueden generar manualmente. Un modelo de datos es solo un conjunto de tablas vinculadas por relaciones. Un modelo de tabla única ya es un modelo de datos, aunque no muy intere-

sante. Tan pronto como tenga varias tablas, la presencia de relaciones hace que el modelo sea mucho más poderoso e interesante de analizar [11].

## 2.2 Modelos matemáticos

El modelamiento matemático ha jugado un papel importante en la planificación de acciones y el control de la dinámica de las enfermedades. Los epidemiólogos, en respuesta a una emergencia sanitaria, obtienen y analizan los datos observados. Usando datos, observaciones, ciencia y teoría proceden a planificar o implementar políticas que mejoren su impacto. Naturalmente, comprender las causas y los modos de transmisión de cada enfermedad es fundamental para pronosticar o mitigar su impacto dentro y entre las poblaciones en riesgo [12].

Este trabajo utilizó dos modelos de regresión no lineales: Logístico y de Gompertz, cuyas curvas se caracterizan por ser sigmoideas (forma de s) y que poseen por lo tanto un punto de inflexión donde la curva cambia desde un «crecimiento rápido» (matemáticamente ocurre donde la curva es convexa) en un «crecimiento lento» (matemáticamente ocurre donde la curva es cóncava). Además, cuentan con un parámetro que es el límite superior de la trayectoria de la pandemia en términos de casos de infección o muertes, dato importante sobre cuántas infecciones y cuántas muertes se pueden esperar en total. Esta información puede ser crucial para las autoridades responsables de abordar y mitigar la pandemia en las diferentes localidades [13].

El modelo logístico fue propuesto por el biólogo matemático holandés Pierre Francois Verhulst en 1840 como un modelo para el crecimiento poblacional. Este modelo cumple con varias condiciones que se dan en el crecimiento de la cantidad de sujetos infectados por un virus específico. Principalmente, el modelo considera que: el crecimiento en la cantidad de infectados es proporcional al producto de la cantidad de la población actual con la cantidad de crecimiento futuro. Además, dicha cantidad llegaría a su *capacidad de soporte* en algún momento, y que, si la población tiende al valor de la capacidad de soporte, la velocidad de crecimiento disminuye de modo que la población de infectados se estabiliza [14].

El modelo se expresa

$$\frac{dw}{dt} = \frac{kw(\alpha-w)}{\alpha}; (k>0) \quad (1)$$

Para este modelo, es la cantidad  $\left(\frac{dw}{dt}\right)_w$  la que declina linealmente conforme aumenta  $w$ , es decir, la tasa de crecimiento es relativa al tamaño presente. La solución a (1) está dada por

$$w = \frac{\alpha}{1 + \beta e^{-kt}} \quad (2)$$

Donde

$\alpha$  es la capacidad de soporte del modelo, cantidad total de casos o muertes que el modelo predice.

$\beta$  es la relación .

$k$  es la tasa de crecimiento relativa del modelo, y  $w$  es la cantidad de casos acumulados.

Esta función tiene su punto de inflexión en  $t_i = \frac{\ln \beta}{k}$ , lo cual significa que antes de este tiempo, la curva es cóncava hacia arriba ( $\frac{d^2w}{dt^2} > 0$ ) y luego del mismo, la curva es cóncava hacia abajo ( $\frac{d^2w}{dt^2} < 0$ ).

Por otro lado, si se relaciona el cambio en la población contagiada como proporcional a la población actual y el logaritmo de la relación entre la capacidad de soporte y dicha población actual, se obtiene

$$\frac{dw}{dt} = kw \log \frac{\alpha}{w} \quad (3)$$

La solución a (3) produce el modelo de Gompertz y lleva su nombre en honor a Benjamín Gompertz (1779-1865). Es una función sigmoidea que describe el crecimiento como más lento al comienzo y al final de un período de tiempo dado

$$w = \alpha e^{-\beta e^{-kt}} \quad (4)$$

Donde

$\alpha$  es la capacidad de soporte del modelo, cantidad total de casos o muertes que el modelo predice.

$\beta$  es el valor que determina el punto del crecimiento exponencial.

$k$  es la tasa de crecimiento relativa del modelo, y  $w$  es la cantidad de casos acumulados.

Esta curva también es sigmoidea, pero no es simétrica alrededor de su punto de inflexión cuyo tiempo es  $t_i = \frac{\ln \beta}{k}$  [15].

Para ambos modelos podemos calcular el coeficiente de determinación  $R^2$  como una medida estadística de qué tan cerca están los datos de la línea de regresión ajustada. La definición de  $R^2$  es bastante sencilla: es el porcentaje de la variación en la variable de respuesta que es explicado por un modelo. Es decir:  $R^2 = \text{Variación explicada} / \text{variación total}$  [15].

### 3. Resultados y discusión

Una vez obtenidos los datos y con la menor cantidad de inconsistencias posibles, realizamos los pasos necesarios para lograr el análisis con R y la visualización por medio de la herramienta Power BI. En esta etapa identificamos las variables y los resultados que deseamos mostrar junto con los objetos visuales que mejor se acomoden para una mejor lectura y comprensión por parte de los usuarios.

### 3.1 Resultados generales

Los resultados generales al 21 de noviembre (Fig. 2) se muestran mediante un tablero resumen o *dashboard* con la siguiente información: casos totales 184876, total hombres 97.984 que representan el 52,9% y total mujeres, 87.076 que representan un 47,1 % del total de casos. El total de fallecidos confirmados 8787 y probables 4352, casos recuperados 160.639. Los casos totales han sido clasificados por grupo etario en el que se puede observar que el grupo más afectado es el de las personas entre 20 y 49 años con el 60% seguido por el de 50 a 64 años que representa casi el 20% de la población afectada. Adicional se puede observar un mapa de localización de provincias con sus respectivos círculos que se relacionan con la cantidad de contagios incluida la región insular.

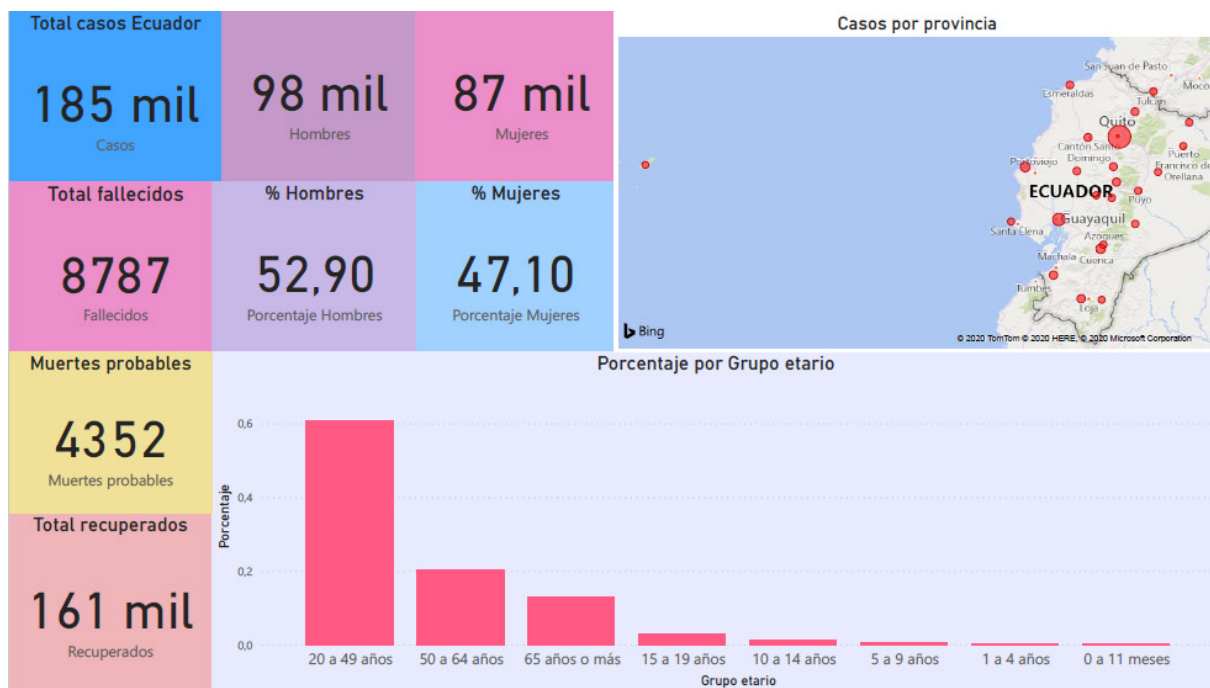


Figura 2. Principales variables presentadas mediante un tablero en Power BI (dashboard)

### 3.2 Casos acumulados por provincia por cada 1000 habitantes

Se presentan los casos acumulados para las provincias más afectadas por cada 1000 habitantes (Fig. 3). Los datos van desde el 13 de marzo hasta

el 15 de noviembre, con el fin de observar la dinámica de cambio comparativa entre estas provincias. Es evidente que la provincia de Pichincha (más específicamente Quito) es la más afectada, seguida por Guayas y Manabí.



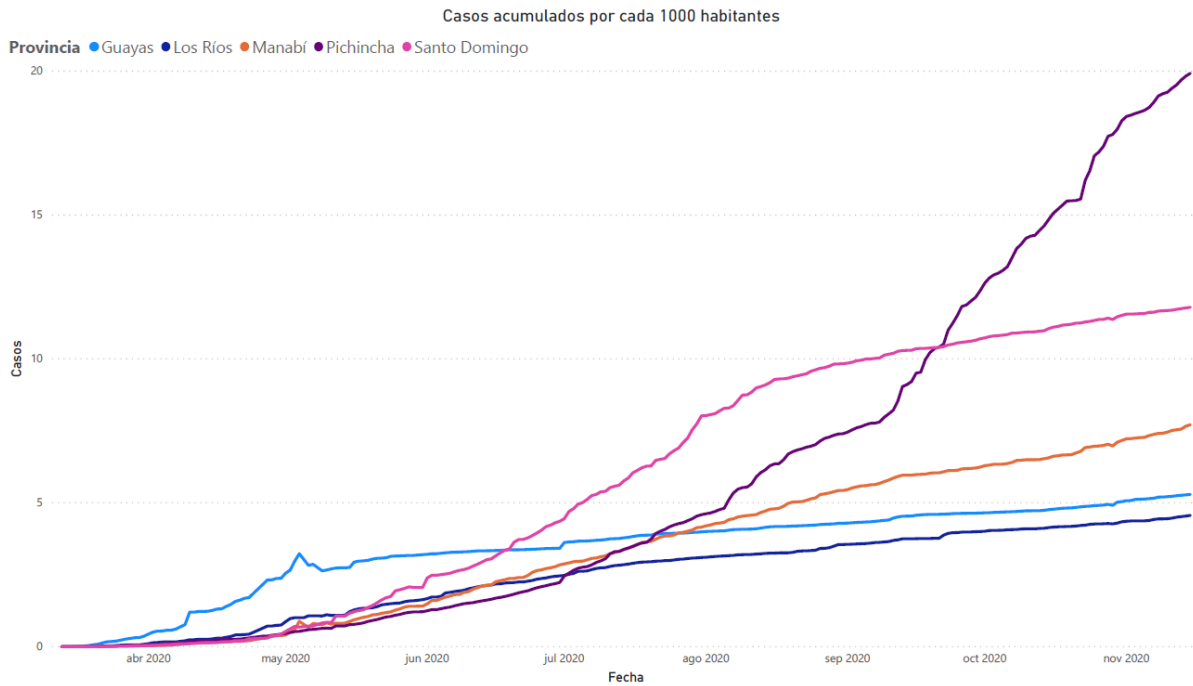


Figura 3. Casos acumulados del 13 de marzo hasta el 15 de noviembre para algunas de las provincias más afectadas

### 3.3 Modelos de ajuste y predicción para el número de casos por provincia

Se aplicaron los modelos logístico y Gompertz por medio del software R mediante el paquete easynls [16]. Para los casos acumulados en las provincias Pichincha, Guayas, Manabí y Santo Domingo y se obtuvieron los parámetros  $\alpha$ ,  $\beta$ ,  $k$  y  $R^2$  y en cada caso (Tabla 1). Los resultados muestran un alto nivel de ajuste de acuerdo al coeficiente  $R^2$  para Pichincha, Manabí, Los Ríos y Santo Domingo. En el caso de Guayas, el número de casos real; 23.227, supera al valor de 20.357 y 20.834 para los modelos logístico y de Gompertz respectivamente, con

lo cual los modelos no podrían predecir valores a futuro para esta provincia. Además, los datos presentan un pico alrededor del día 50 (30 de abril) lo cual representa un error en la recolección debido a que es imposible que los datos tomen valores más bajos posteriormente. En la siguiente sección se realiza una corrección para esta provincia. Podemos observar también los valores de  $k$  de lo cual destacamos la provincia de Pichincha cuyo valor es el más alto para ambos modelos. Esto nos indica según el modelo Gompertz, que la ciudad aún no ha llegado al valor pico en sus casos diarios.

Tabla 1. Parámetros  $\alpha$ ,  $\beta$ ,  $k$ ,  $R^2$  y  $\sigma$  para los modelos logístico y Gompertz en las provincias de interés.

Parámetros	Guayas	Pichincha	Manabí	Los Ríos	Santo Domingo
Modelo Logístico					
$\alpha$	20357	105661	11597	3923	5255
$\beta$	6,8647	147,905	37,203	16,4675	94,1132
$k$	0,0316	0,02215	0,0272	0,02759	0,03715

$R^2$	0,9360	0,99770	0,9938	0,98390	0,99800
$t_i$	60.9	225.3	132.7	101.5	122.3

Modelo de Gompertz

$\alpha$	20834	416708	13313	4211	5640
$\beta$	2,7128	7,44176	5,69932	4,06144	10,3447
k	0,0226	0,00566	0,01470	0,01689	0,02171
$R^2$	0,9570	0,99780	0,99860	0,99390	0,99760
$t_i$	37.5	354.5	124.3	83.4	107.6

Por medio de los parámetros se construyeron los modelos de cada provincia para compararlos gráficamente en Power BI mediante su herramienta *Gráfico de líneas* (Fig. 4 y 5). Se puede apreciar

que el nivel de ajuste en la mayoría de los casos es alto, permitiendo de esta manera realizar predicciones significativas.

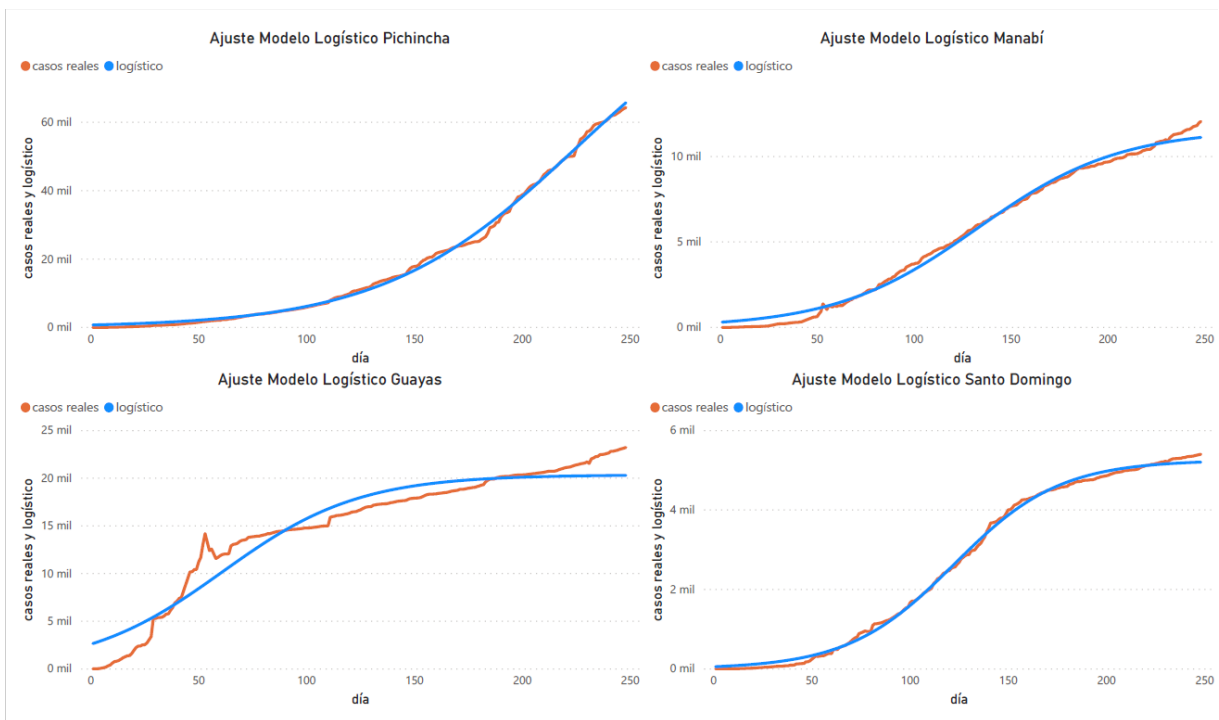


Figura 4. Ajuste del modelo logístico a las provincias de interés

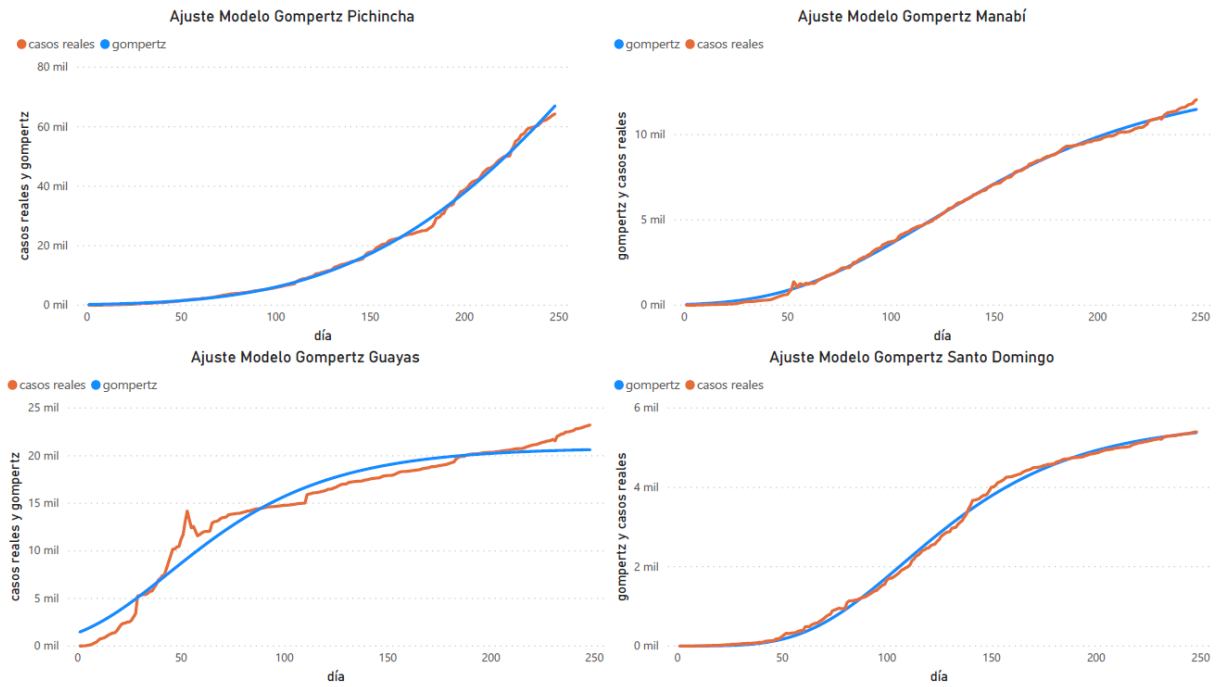


Figura 5. Ajuste del modelo Gompertz a las provincias de interés

La Tabla 2 muestra predicciones para cada una de las provincias de interés para las fechas 1/12/2020 y 31/12/2020.

Tabla 2. Predicciones para las provincias de interés

Provincia	Actual	Modelo Logístico		Modelo Gompertz	
		1/12/2020	31/12/2020	1/12/2020	31/12/2020
Pichincha	64320	74085	86660	78542	101937
Manabí	12052	11282	11456	11839	13245
Santo Domingo	5403	5227	5246	5454	5542

### 3.4 Ajuste de los modelos para Guayas con datos corregidos

Se realizó una corrección a los datos de la provincia del Guayas para obtener un mejor ajuste del modelo. Para esto se realizó una interpolación entre los días 40 y 60 de modo que se evita el pico en esa sección. Además, se ha realizado la regresión con el modelo de Von Bertalanffy cuya ecuación para el número casos viene dada por  $y = \frac{K}{1 + e^{-a(x-x_0)}}$ . Este modelo podría mejorar el ajuste debido a sus cuatro pa-

rámetros, aunque si el crecimiento es isométrico, la potencia se fija en 3 [17]. Los resultados (Fig. 6) muestran que el ajuste mejora considerablemente. Los valores obtenidos para los parámetros fueron:  $a = 22303$ ,  $x_0 = 0.67909$  y  $K = 0.01616$ . Aunque el nivel de ajuste dado por el parámetro  $a$  ha mejorado hasta el valor 0.9901, el valor real de casos acumulados para el último día; 23.227, sigue superando al valor máximo de casos al que se llegaría según el modelo.



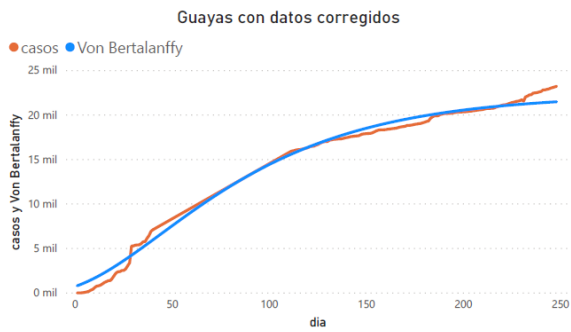


Figura 6. Ajuste del modelo Von Bertalanffy a los datos corregidos Guayas.

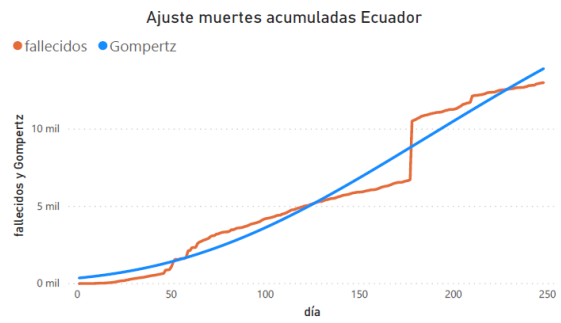


Figura 7. Ajuste del modelo de Gompertz para el número de muertes acumuladas Ecuador

### 3.5 Muertes acumuladas Ecuador

Se realizó el ajuste para muertes acumuladas en el país mediante el modelo de Gompertz (Fig. 7). Los parámetros resultantes fueron:  $\alpha = 25728$ ,  $\beta = 4.296471$  y  $k = 0.0078452$ . El coeficiente del nivel de ajuste fue  $R^2 = 0.9646$ . Podemos observar un salto abrupto de los datos alrededor del día 180. Sin embargo, el coeficiente de ajuste presenta un valor bastante aceptable

### 3.6 Muertes en bruto por provincia

Para el caso de muertes por provincia (Fig. 8), se realizó el análisis al 21 de noviembre tomando los datos en bruto para cada caso y mostrándolos mediante el *Gráfico de columnas* de Power BI. El diagrama fue realizado diferenciando la región a la que pertenece cada provincia mediante colores y se puede observar que Pichincha, Guayas y Manabí lideran la lista. Desde el punto de vista de regiones, la región Costa es la que tiene todas sus provincias, excepto Esmeraldas en el grupo de más afectadas.

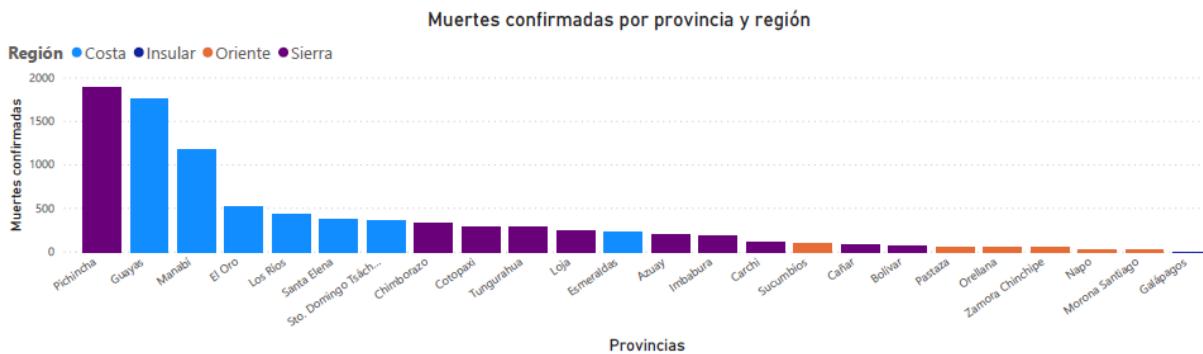


Figura 8. Muertes confirmadas por provincia y región.

### 3.7 Mapa para casos en bruto por provincia

Se muestra ahora el número de casos en bruto por provincia (Fig. 9) al 21 de noviembre. Se han incluido las coordenadas de latitud y longitud con

el fin de utilizar la herramienta *Mapa* de Power BI que nos permite visualizar la información de forma geográfica. El tamaño de los círculos color rojo son proporcionales al número de casos.

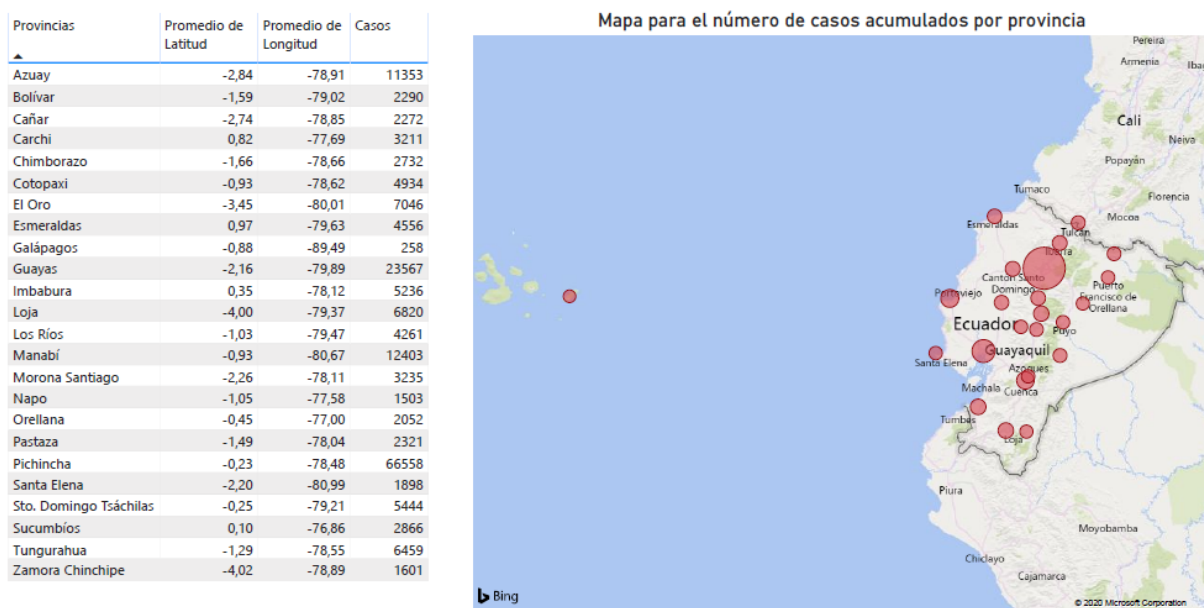


Figura 9. Casos por provincia presentados en forma geográfica que incluye las coordenadas de latitud y longitud.

### 3.8 Tasa de letalidad por provincia

La tasa de letalidad se la obtiene dividiendo el número de muertes para el total de casos confirmados: . La tabla 2 muestra los resultados para cada provincia. Podemos ver que las provincias con mayor tasa de letalidad son Santa Elena, Chimborazo, Manabí y Los Ríos.

La información recolectada por las entidades encargadas no es del todo exacta, existen altibajos abruptos en los datos, días sin información, variables que han cambiado con el tiempo, entre otros. Se debe tomar en cuenta, por ejemplo, la existencia de pacientes asintomáticos, personas con bajos recursos, personas que habitan lejos de los centros de salud y que, a pesar de poseer el virus, no acuden a las pruebas que nos permita contar con información real de lo que está sucediendo. Esta situación limita el control estadístico y se convierte también en una gran limitante para médicos, pacientes y familiares, presas de la incertidumbre a falta de un diagnóstico oportuno. Sin embargo, se ha curado la información de modo que se obtenga el resultado más transparente posible.

Tabla 2. Tasa de letalidad por provincia

Provincia	Casos confirmados	Fallecidos	Tasa de letalidad
Azuay	11353	201	1,77%
Bolívar	2290	67	2,93%
Cañar	2272	86	3,79%
Carchi	3211	109	3,39%
Chimborazo	2732	332	12,15%
Cotopaxi	4934	292	5,92%
El Oro	7046	518	7,35%
Esmeraldas	4556	222	4,87%
Galápagos	258	1	0,39%
Guayas	23567	1755	7,45%
Imbabura	5236	178	3,40%
Loja	6820	238	3,49%
Los Ríos	4261	348	8,17%
Manabí	12403	1172	9,45%
Morona Santiago	3235	24	0,74%
Napo	1503	75	4,99%
Orellana	2052	53	2,58%
Pastaza	2321	60	2,59%
Pichincha	66558	1883	2,83%

Santa Elena	1898	375	19,76%
Santo Domingo	5444	366	6,72%
Sucumbíos	2866	95	3,31%
Tungurahua	6459	284	4,40%
Zamora Chinchipe	1601	53	3,31%

Los datos oficiales publicados por el Ministerio de Salud Pública muestran que existe una tendencia preocupante al alza para el número de casos totales en el país. A esto hay que añadirle las consecuencias propias de un evento tan perjudicial como la afectación económica a nivel nacional. Es urgente la necesaria reactivación económica y social para la población cuya situación en muchos casos se ha vuelto compleja y cuyas consecuencias son, por ejemplo, el incremento de forma alarmante de los índices delincuenciales a todo nivel.

Se conoce que 289.258 contratos laborales han finalizado desde el 16 de marzo de 2020, cuando empezó la emergencia sanitaria de covid-19. La cifra la emitió el Ministerio de Trabajo en una rueda de prensa el 30 de julio de 2020, en la que afirmó, además, que la pequeña y mediana empresa también se ha visto muy afectada, así como la economía informal debido al confinamiento. Los cinco sectores más afectados serían según informes oficiales el comercio, la agricultura, la manufactura, la construcción y el turismo [18].

A pesar de la limitación en el número y distribución de las pruebas diagnósticas para covid-19, las autoridades sanitarias podrían analizar la distribución del virus a niveles más desagregados como cantonales y parroquiales con el fin de hallar mejores estrategias en el control de la enfermedad [19]. En este sentido es que Power BI puede ser una herramienta de gran aporte en el análisis de la situación general. También se vuelve importante el tecnicismo con el que midan las diferentes variables de la enfermedad para recolectar la información adecuada. De acuerdo a [8] tres son los aspectos que se deben tomar en cuenta: 1) Realizar un buen estudio para medir la tasa correcta de contagio en cada localidad, si se evalúan solo personas con síntomas puede ser que se esté subestimando de forma sustancial la verdadera tasa de contagio mientras que, si se miden masivamente puede ser que la tasa se sobreestime. 2) El rendi-

miento de la prueba (sensibilidad, especificidad, valores predictivos) de las pruebas en el campo son todavía en gran parte desconocidos y variarán según el lugar y el tiempo. 3) Al estimar las tasas de mortalidad por infecciones, en lugar de conocer el número real de infectados que han muerto, generalmente tenemos que conformarnos con el número de personas que han dado positivo y fallecido, lo que puede llevar a una sobreestimación o subestimación de casos.

La tasa de letalidad es alarmante para Ecuador, al 27 de Noviembre, se encuentra en 7,11%, siendo la tasa más alta de los países de la región (Perú 3,74%, Colombia 2,82%, Venezuela 0,87%, Chile 2,78%, Argentina 2,70%, Brasil 2,78%, Bolivia 6,19%, Paraguay 2,15% y Uruguay 1,47%) [20]. Las razones serían múltiples: baja calidad de la información, falta de recursos sanitarios, densidad poblacional, irresponsabilidad en el cuidado personal y más. El gran tamaño de la población puede hacer que sea más difícil aumentar la capacidad del sector de la salud. En términos de densidad, Colombia y Ecuador son claramente valores atípicos, lo que puede afectar la capacidad de mitigar la propagación de covid-19 y mantener el distanciamiento social en áreas públicas como, por ejemplo, el transporte público [3].

Los modelos logísticos y de Gompertz resultaron ser sencillos, pero eficaces para obtener predicciones confiables sin tomar parámetros externos, sino solamente los datos obtenidos de casos diarios y acumulados como en [21], [22], [13]. Los valores para  $R^2$  también resultaron ser altos como en [22] y [13].

Se puede realizar un ajuste de los datos cuando estos presentan grandes inconsistencias como en [22]. En nuestro caso, se realizó un ajuste a los datos del Guayas por medio de una interpolación lo cual mejoró notablemente al modelo en su ajuste, aunque no en sus predicciones. En esta provincia se pudo observar un «rebrote» de casos, ya que se presenta un aumento repentino al final de los mismos lo que no hizo posible su predicción con los modelos dados.

#### 4. Conclusiones

Las plataformas de inteligencia de negocios resultan ser una herramienta con muchas ventajas a la

hora de presentar resultados eficientes y rápidos. Por eso se comprende la importancia de que las entidades públicas cuenten con plataformas de estas características, más aún en una crisis global en la que se genera información diaria que debe ser analizada de forma permanente y eficiente. Es de suma importancia contar con datos abiertos y de calidad y en este sentido, Ecuador necesita mejorar sus sistemas de adquisición de datos de modo que la información sea completa, sistematizada y probada para los fines investigativos pertinentes. Esto requiere del esfuerzo ciudadano y de políticas que permitan ir construyendo estructuras que generen esta información, más aún si tenemos los índices más alarmantes para las variables de la pandemia a nivel regional. Finalmente se debe tomar en cuenta que todas las consecuencias nefastas que esta pandemia está generando pueden ser mitigadas con el solo hecho de adquirir, procesar y analizar datos de forma correcta.

### Referencias

- [1] MSP, “Ministerio de Salud Pública,” *Actualización de casos de coronavirus*, 2020. <https://www.salud.gob.ec/actualizacion-de-casos-de-coronavirus-en-ecuador/> (accessed Aug. 06, 2020).
- [2] G. Quiroz, “La letalidad por el covid-19 de Ecuador es la más alta de Sudamérica | El Comercio,” 2020. <https://www.elcomercio.com/actualidad/peru-chile-brasil-covid-19.html> (accessed Nov. 26, 2020).
- [3] M. Benítez, C. Velasco, A. Sequeira, J. Henríquez, F. Menezes, and F. Paolucci, “Responses to COVID-19 in five Latin American countries,” *Heal. Policy Technol.*, vol. 9, no. January, pp. 525–559, 2020, [Online]. Available: <https://doi.org/10.1016/j.hlpt.2020.08.014>.
- [4] G. M. Sechi *et al.*, “Business intelligence applied to emergency medical services in the lombardy region during sars-COV-2 epidemic,” *Acta Biomed.*, vol. 91, no. 2, pp. 39–44, 2020, doi: 10.23750/abm.v91i2.9557.
- [5] L. Loewen and A. Roudsari, “Evidence for Business Intelligence in Health Care: A Literature Review,” *Stud. Health Technol. Inform.*, vol. 235, no. January, pp. 579–583, 2017, doi: 10.3233/978-1-61499-753-5-579.
- [6] Primicias, “Primicias. El periodismo comprometido. Noticias, política, economía, sociedad,” 2020. <https://www.primicias.ec/> (accessed Jun. 23, 2020).
- [7] B. Powell, *Creating Business Intelligence Solutions of Analytical Models, Reports and Dashboards*. Birmingham: Packt Publishing Ltd., 2017.
- [8] N. Pearce, J. Vandenbroucke, T. VandenWeele, and S. Greenland, “Accurate Statistics on COVID-19 Are Essential for Policy Guidance and Decisions,” *Am. J. Public Health*, vol. 110, no. 7, pp. 964–966, 2020, doi: 10.2105/AJPH.2020.305708.
- [9] C. Webb, *Power Query for Power BI and Excel*. New York: Apress, 2014.
- [10] Microsoft, “Data Visualization | Microsoft Power BI,” 2020. <https://powerbi.microsoft.com/en-us/> (accessed Nov. 27, 2020).
- [11] A. Ferrari and M. Russo, *Analyzing Data with Power BI and Power Pivot for Excel*, 1st ed. USA: Microsoft Press, 2017.
- [12] F. Brauer, C. Castillo-Chavez, and Z. Feng, *Mathematical Models in Epidemiology*. 2019.
- [13] J. Mazurek, Z. Neničková, J. Mazurek, and Z. Neničková, “Predicting the number of total COVID-19 cases and deaths in the USA by the Gompertz curve.” Accessed: Jun. 23, 2020. [Online]. Available: <https://www.researchgate.net/publication/340809381>.
- [14] J. Stewart, *Single Variable Calculus*, 8th ed. Boston, 2012.
- [15] N. R. Draper and H. Smith, *Applied Regression Analysis*, vol. 87, no. 10. Wiley, 1998.
- [16] E. Arnhold, “easynls: Easy Nonlinear Model,” *R Packag. version 5.0*, pp. 1–9, 2017. Available:

- ble: <https://cran.r-project.org/package=easynls>.
- [17] R. Villaseñor, J. Jurado, and D. Salas, "Solución de la ecuación diferencial de crecimiento en peso, de Von Bertalanffy (1938), por dos métodos distintos," *An. DEL Inst. CIENCIAS DEL MAR Y Limnol.*, vol. 19, 1993, Accessed: Mar. 14, 2021.
- [18] Reuters, "Más de 289.000 personas han perdido sus empleos durante pandemia," 2020. <https://www.primicias.ec/noticias/economia/despidos-registro-pandemia-empleo-ecuador/> (accessed Nov. 27, 2020).
- [19] E. Farach and A. Romero, "Escenarios estadísticos del COVID-19 en Honduras: análisis desde la inteligencia de negocios y la transparencia de datos," *Innovare Rev. Cienc. y Tecnol.*, vol. 9, no. 2, pp. 89–96, 2020, doi: 10.5377/innovare.v9i2.10193.
- [20] Johns Hopkins, "COVID-19 Map - Johns Hopkins Coronavirus Resource Center," 2020. <https://coronavirus.jhu.edu/map.html> (accessed Nov. 27, 2020).
- [21] P. Sánchez-Villegas and A. Daponte Codina, "Predictive models of the COVID-19 epidemic in Spain with Gompertz curves," *Gac. Sanit.*, May 2020, doi: 10.1016/j.gaceta.2020.05.005.
- [22] V. Arias and M. Alberto, "Estimación de población contagiada por Covid-19 usando regresión Logística generalizada y heurísticas de optimización," Apr. 2020. Accessed: Jun. 23, 2020. [Online]. Available: <http://www.kerwa.ucr.ac.cr/handle/10669/80859>.