

INTELIGENCIA DE NEGOCIOS CON POWER BI: CASOS DE COVID-19 EN MÉXICO

Flor Margarita Calderón Cuevas

Escuela Superior De Ciencias y Tecnologías De La Información

Correo Electrónico: florcalderon@uagro.mx

RESUMEN

Este artículo tiene como propósito presentar un trabajo de inteligencia de negocios utilizando el software Power BI de Microsoft; se hizo la creación de un panel interactivo con los datos obtenidos de la página oficial del Gobierno de México, la cual proporciona un catálogo de datos sobre los casos diarios de personas confirmadas, sospechosas, negativas y las defunciones a causa del Covid-19 en todos los estados de la República mexicana. El panel interactivo contiene gráficas de líneas, un mapa de la República mexicana, tarjetas en donde se muestra los casos confirmados, sospechosos, negativos y las defunciones en total desde el 31 de diciembre del 2019 hasta la fecha del 19 de mayo del 2021; así como también los casos totales cada 24 horas y los top de estados y municipios afectados por el Covid-19.

Palabras clave: *Power BI, Covid-19, México, datos, estados y casos.*

ABSTRACT

This article is intended to present a business intelligence paper using Microsoft's Power BI software; an interactive panel was created with the data obtained from the official website of the Government of Mexico, which provides a catalog of data on the daily cases of confirmed, suspected, negative people and deaths due to Covid-19 in all the states of the Mexican Republic. The interactive panel contains line graphs, a map of the Mexican Republic, cards showing confirmed, suspected, negative cases and deaths in total from December 31, 2019, to the date of May 19, 2021; as well as the total cases every 24 hours and the top states and municipalities affected by covid-19.

Keywords: *Power BI, Covid-19, Mexico, data, states and cases.*

1. INTRODUCCION

En la actualidad el mundo entero está padeciendo de una situación muy grave: casi 2 millones de personas se han afectado por una angustiosa enfermedad, la COVID-19, provocada por un virus poco conocido, el SARS-CoV-2, que ha provocado alrededor de 126,604 de fallecidos, incluidos jóvenes, aunque la afectación es mayor en personas de la tercera edad y aquellas que tienen factores de riesgo que los hacen más vulnerables como la hipertensión arterial, la diabetes mellitus, entre otras (Velázquez Pérez, 2020). Lo que empezó con una serie de

casos un diciembre del 2019 de pacientes hospitalizados a causa de un nuevo coronavirus (SARS-CoV-2), en la provincia de Hubei, China se convirtió el 11 de marzo, una pandemia la cual fue declarada en una rueda de prensa mundial por Tedros Adhanom Ghebreyesus, director general de la Organización Mundial de la Salud (Ferrer, 2020).

Después de los aumentos de los casos por la enfermedad en algunos países un 27 de febrero del 2020 se presentó el primer caso de Covid-19 en México. El 30 de abril, 64 días después de este primer diagnóstico, el número de pacientes aumentó exponencialmente, alcanzando un total de 19.224 casos confirmados y 1.859 (9,67%) fallecidos (Suárez, Suarez Quezada, Oros Ruiz, & Ronquillo de Jesús, 2020).

Con la estrategia de hacer un análisis de los casos confirmados, sospechosos, negativos y las defunciones en toda la república mexicana, se hizo uso de la inteligencia de negocios la cual nos permite hacer análisis de negocios, implementar la minería de datos y ayudar a las organizaciones a tomar decisiones basadas en los datos. Existen muchas herramientas que nos permiten hacer soluciones de inteligencia de negocios entre las que se destaca Power BI.

Power BI es una herramienta que permite la unificación de datos de numerosos orígenes para crear paneles interactivos y envolventes, así como informes que ofrecen conocimientos útiles y mejoran los resultados empresariales (BI, 2021). Para poder realizar el panel interactivo el cual lleva por nombre CASOS DE COVID-19 EN MÉXICO POR ESTADOS, se hizo la descarga de los datos a través de la página oficial del Gobierno de México, a continuación, se describe el proceso para la creación del panel.

1.1 OBJETIVO GENERAL

Creación de un panel interactivo con Power BI donde se visualicen gráficas, tablas, mapas y tarjetas con los datos obtenidos de la página oficial del Gobierno de México sobre el Covid-19 en México por estados y municipios; para el cual posteriormente hacer un análisis exhaustivo.

2. METODOLOGÍA

Para llevar a cabo el panel se utilizó una serie de fases las cuales fueron: fuentes y limpieza de datos, selección y transformación de los datos con Power BI, construcción del panel interactivo y pruebas y diseño final.

2.1 Fuentes y limpieza de datos

Los datos de casos diarios por Estado y Municipios debido a la infección COVID-19 en México se obtuvieron del sitio web oficial del Gobierno de México a través de la Dirección General de Epidemiología (México, 2021). El período analizado fue entre el 31 de diciembre del 2019 y el 19 mayo del 2021. El conjunto de datos fue descargado en formato .CSV, fueron ocho conjunto de datos los cuales eran casos confirmados, sospechosos, negativos y las defunciones de estados y municipios, los cuales contenían los siguientes campos:

- Cve_ent: Clave de entidad.
- Población.
- Nombre del estado en el caso de los datos de estado y nombre del municipio en caso de los datos de municipios.

- Las fechas venían por columnas, contenían muchas columnas ya que las fechas fueron colocadas como columnas.

Después de hacer la revisión de todos los conjuntos de datos, se prosiguió a seleccionar los campos necesarios para el panel, se noto un error en la última fecha que era la del 20 de mayo del 2021, así que para que no hubiera complicaciones mas adelante en el trabajo se elimino de todos los conjuntos de datos; así como también se elimino la fila “Nacional” que estaba en todos los conjuntos de los estados, ya que contenía el resumen total de los estados por día y eso podría habernos ocasionado problemas adelante. El conjunto de datos de Estados quedo como lo muestra la figura 2.1.

cve_ent	poblacion	nombre	31/12/2019	01/01/2020	02/01/2020	03/01/2020	04/01/2020	05/01/2020	06/01/2020	07/01/2020	08/01/2020	09/01/2020
1	1434635	AGUASCALIENTES	0	0	0	0	0	0	0	0	0	0
2	3634868	BAJA CALIFORNIA	0	0	0	0	0	0	0	0	0	0
3	804708	BAJA CALIFORNIA SUR	0	0	0	0	0	0	0	0	0	0
4	1000617	CAMPECHE	0	0	0	0	0	0	0	0	0	0
7	5730367	CHIAPAS	0	0	0	0	0	0	0	0	0	0
8	3801487	CHIHUAHUA	0	0	0	0	0	0	0	0	0	0
9	9018645	DISTRITO FEDERAL	0	0	0	0	0	0	0	0	0	0
5	3218720	COAHUILA DE ZARAGOZA	0	0	0	0	0	0	0	0	0	0
6	785153	COLIMA	0	0	0	0	0	0	0	0	0	0
10	1868996	DURANGO	0	0	0	0	0	0	0	0	0	0
11	6228175	GUANAJUATO	0	0	0	0	0	0	0	0	0	0
12	3657048	GUERRERO	0	0	0	0	0	0	0	0	0	0
13	3086414	HIDALGO	0	0	0	0	0	0	0	0	0	0
14	8409693	JALISCO	0	0	0	0	0	0	0	0	0	0
15	17427790	MEXICO	0	0	0	0	0	0	0	0	1	0
16	4825401	MICHOACAN DE OACAN	0	0	0	0	0	0	0	0	0	0
17	2044058	MORELOS	0	0	0	0	0	0	0	0	0	0
18	1288571	NAYARIT	0	0	0	0	0	0	0	0	0	0
19	5610153	NUEVO LEON	0	0	0	0	0	0	0	0	0	0
20	4143593	OAXACA	0	0	0	0	0	0	0	0	0	0
21	6604451	PUEBLA	0	0	0	0	0	0	0	0	1	0
22	2279637	QUERETARO	0	0	0	0	0	0	0	0	0	0
23	1723259	QUINTANA ROO	0	0	0	0	0	0	0	0	0	0
24	2866142	SAN LUIS POTOSI	0	0	0	0	0	0	0	0	0	0
25	3156674	SINALOA	0	0	0	0	0	0	0	0	0	0
26	3074745	SONORA	0	0	0	0	0	0	0	0	0	0
27	2572287	TABASCO	0	0	0	0	0	0	0	0	0	0
28	3650602	TAMAULIPAS	0	0	0	0	0	0	0	0	0	0
29	1380011	TLAXCALA	0	1	0	0	0	0	0	1	1	0
30	8539862	VERACRUZ	0	0	0	0	0	0	0	0	0	0
31	2259098	YUCATAN	0	0	0	0	0	0	0	0	0	0
32	1666426	ZACATECAS	0	0	0	0	0	0	0	0	0	0

Figura 2.1 Conjunto de datos por Estados corregidos en formato .CSV

2.2 Limpieza y transformación de los datos con Power BI

Después de haber llevado a cabo la investigación y análisis de los campos que se encontraban en la base de datos (ya descritos anteriormente) se prosiguió a seleccionar los campos útiles para la investigación; para ello las base de datos en formato .CSV fueron cargados al software Power BI; ya en Power BI se seleccionó el editor de consultas (ícono Transformar datos), ya dentro del editor se les cambio el nombre a cada tabla o consulta por: EstadoConfirmados, EstadoDefunciones, EstadoNegativos, EstadoSospechosos, MunicipioConfirmados, MunicipioDefunciones, MunicipioNegativos, MunicipioSospechosos. Después a cada tabla se les aplico lo siguiente:

- Usar la primera fila como encabezado.
- Después como los datos estaban organizados que cada día tenía una columna como se menciono anteriormente, se hizo uso de anular la dinamización de las columnas y de esta forma las columnas se convirtieron en dos las cuales correspondían a atributo y valor, la de atributo correspondía a la fecha y la otra a los valores de los casos ya sea confirmados, sospechosos, etc.
- Se cambio el nombre de Atributo por Fecha, se verifico que los datos válidos estaban al 100%.
- En la columna fecha y desde la pestaña de Transformar en el apartado de Fecha se puso a opción Analizar fecha, se reconoció la fecha y de esa manera quedo organizada.
- Se eliminaron los campos que no se iban a utilizar como la población y la clave; se les hizo exactamente lo mismo a todas las tablas y se aplicaron los cambios en la figura 2.2 se puede observar como quedaron las tablas en Power BI.

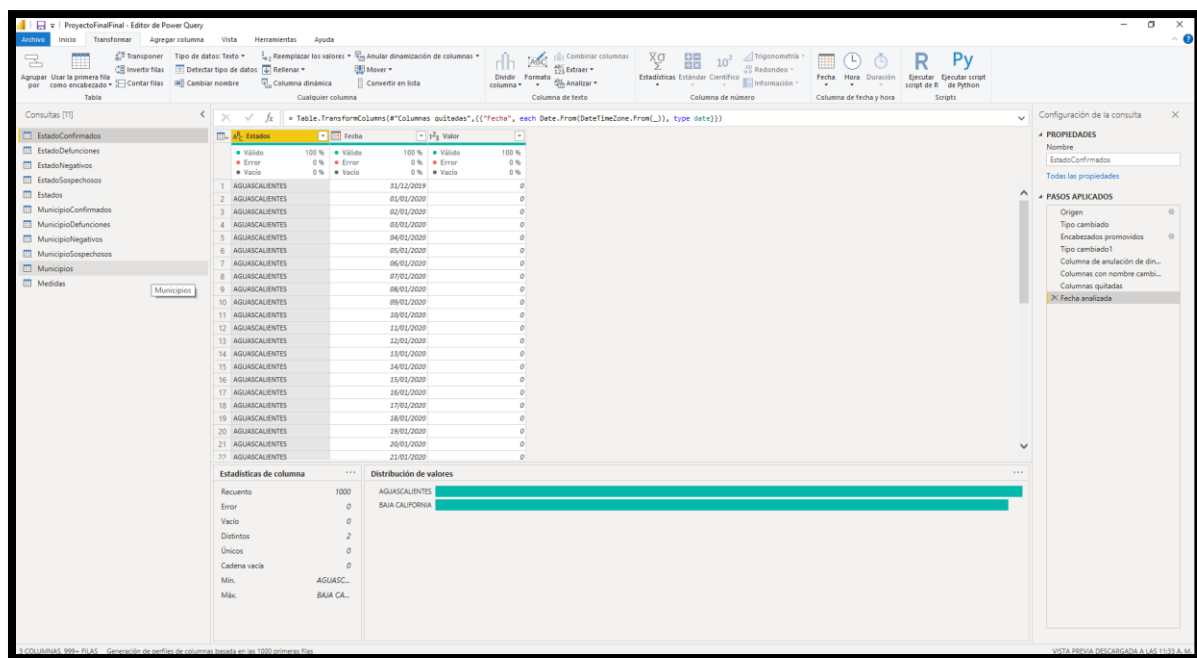


Figura 2.2 Pantalla final en Power BI de la limpieza y transformación de los datos.

Después de hacer la limpieza se prosiguió a crear el modelo de conexiones, para la conexión se tuvo que crear tres tablas una que contuviera las fechas, otra los estados y la ultima los municipios, y así con ello establecer las relaciones entre las tablas. Para la creación de la tabla fechas en la pestaña datos, se dio clic en Nueva tabla, se le dio el nombre de Tabla de fechas, se usó la función Calendar() con el siguiente comando:

Tabla de fechas = CALENDAR(MIN(EstadoConfirmados[Fecha]),MAX(EstadoConfirmados[Fecha]))

La cual nos devuelve la mínima fecha de la tabla EstadoConfirmados y la máxima de la misma columna.

Para las tablas de Estados y Municipios se entro al editor de Power BI y se duplicaron las tablas de EstadoConfirmados y MunicipiosConfirmados, se les quito las columnas Fecha y valores, se les aplico la opción quitar duplicados, se les cambio el nombre por Estados y Municipios, se cerró y se aplicó los cambios.

En la opción de modelo de la principal pantalla de Power BI se puede observar como se ha creado el modelo de conexión entre las tablas (ver figura 2.3).

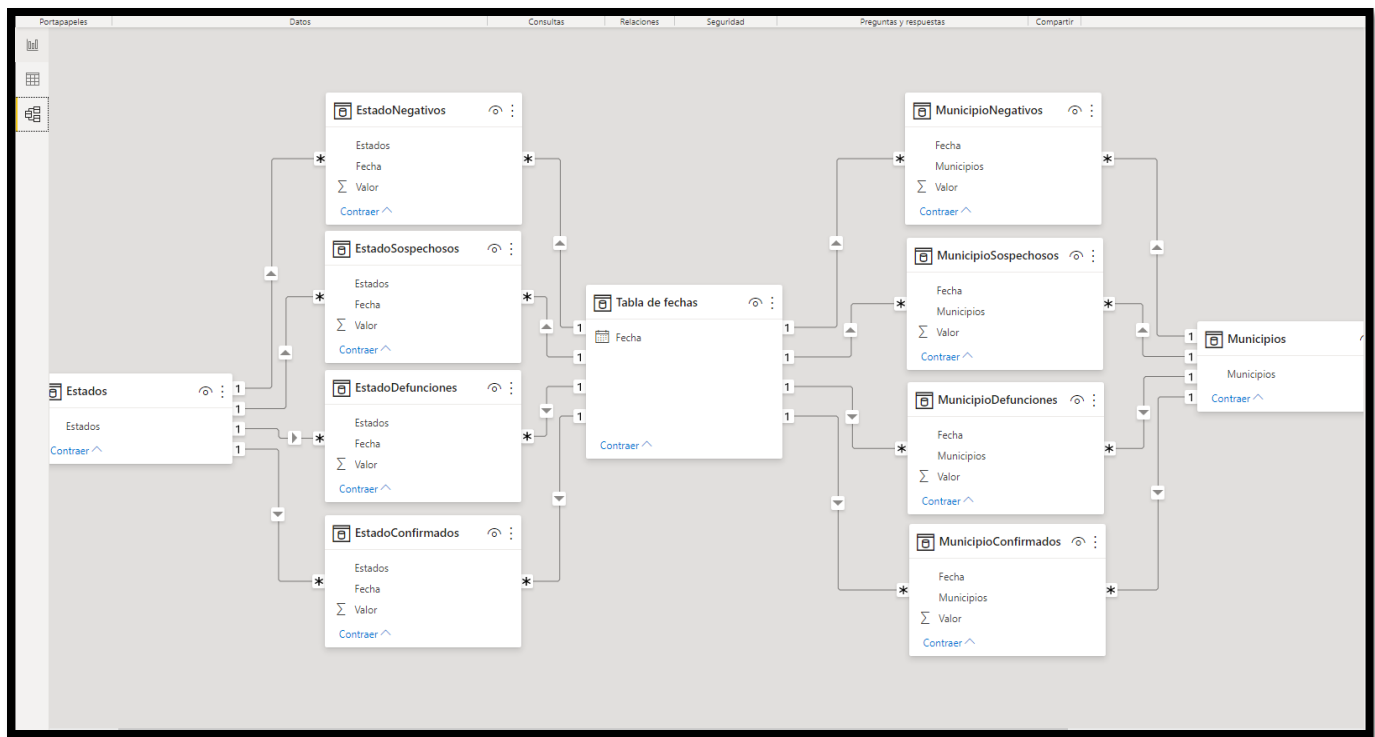


Figura 2.3 Pantalla en Power BI del modelo de conexión.

2.3 Creación de las medidas para el panel interactivo

En esta parte se empezó a añadir las medidas al informe. Para agregar medidas al informe se creo primero una tabla llamada Tabla Medidas en la cual se irán creando todas medidas que se van a necesitar para el panel; las primeras medidas devuelven el acumulado de los casos del último día y son las siguientes:

Total confirmados estados = `CALCULATE(SUM(EstadoConfirmados[Valor]),FILTER("Tabla de fechas",'Tabla de fechas'[Fecha]= MAX("Tabla de fechas"[Fecha])))`

Total defunciones estados = `CALCULATE(SUM(EstadoDefunciones[Valor]),FILTER("Tabla de fechas",'Tabla de fechas'[Fecha]= MAX("Tabla de fechas"[Fecha])))`

Total negativos estados = `CALCULATE(SUM(EstadoNegativos[Valor]),FILTER("Tabla de fechas",'Tabla de fechas'[Fecha]= MAX("Tabla de fechas"[Fecha])))`

Total sospechosos estados = `CALCULATE(SUM(EstadoSospechosos[Valor]),FILTER("Tabla de fechas",'Tabla de fechas'[Fecha]= MAX("Tabla de fechas"[Fecha])))`

Para crearlas dentro de la tabla Medidas se tuvo que seleccionar la tabla y en el ícono de nueva medida dar clic, se escribieron las medidas presentadas anteriormente pero por separado; se utilizó la función `CALCULATE()`, la cual

es una expresión que indica una suma de valores ya sea en el caso de los confirmados, sospechosos, negativos, etc. Y también se usó la función **FILTER()** la cual nos ayudó a indicar que la columna fecha de la tabla de fechas debía ser igual al máximo que aparece en la columna fecha de la tabla fechas. De esta forma se obtuvo la suma de los valores confirmados, defunciones, negativos y sospechosos de la fecha más reciente a la última fecha.

Ya por último se agregaron las siguientes medidas:

Municipios afectados=**CALCULATE(COUNTROWS('Municipios'),FILTER(MunicipioConfirmados,MunicipioConfirmados[Valor] > 0))**

Estados afectados =

CALCULATE(COUNTROWS('Estados'),FILTER(EstadoConfirmados,EstadoConfirmados[Valor] > 0))

Las cuales nos dan el recuento de todos los Estados o Municipios afectados; se hizo uso de la función **CALCULATE()**, como expresión se usó la función **COUNTROWS()** para que contará todas las filas que tenía la tabla Municipios y Estados, se usó la función **FILTER()** para seleccionar las tablas de EstadoConfirmados y Municipios Confirmados y que en el filtro de la expresión nos devuelva el número de los confirmados pero que sea mayor a cero. Después de eso ya se tenían todas las medidas que se iban a utilizar para el cuadro de mando o panel.

2.4 Construcción del panel y diseño final

Para las representaciones y para que el panel luciera mucho más profesional se descargó un tema en la página oficial de Power BI (BI, 2021). Después de instalar el tema se empezó con la creación del panel para ello se hizo lo siguiente:

- Se agregó un título el cual es “CASOS DE COVID-19 EN MÉXICO POR ESTADOS”; se le dio un diseño apropiado al título para que se viera profesional.
- Se crearon dos secciones una donde iban a ir los casos totales y casos totales en 24 Horas.
- En la sección de casos totales se agregaron los valores de cada una de las tablas de estado para que mostrara un recuento total de los casos que se llevan hasta la fecha del 19 de mayo del 2021.
- En la sección de casos totales en 24 horas se colocaron las medidas creadas Total confirmados estados, Total defunciones estados, Total negativos estados y Total sospechosos estados; en las dos secciones se hizo uso de las tarjetas que se encuentran en la sección de visualizaciones de Power BI.
- Se agregó un mapa de la república mexicana, se le dio un color diferente a cada estado y que estuviera de acorde al diseño del panel.
- Se agregaron dos segmentaciones de datos uno para hacer un filtro de los estados y otro para las fechas, se les aplicó un diseño, y se acomodó de tal forma que pudiera ser fácil su manejo.
- Se agregaron dos tablas que muestran los TOP 5 de los Estados y Municipios con más casos confirmados de la última fecha; para ello se hizo uso de las tablas estados, municipios y las medidas de Total confirmados estados y Total de confirmados municipios.
- Por último se agregaron dos gráficos de líneas los cuales muestran en Total de confirmados por fecha y el total de defunciones por fecha.

3. RESULTADOS

Los resultados obtenidos por el panel fueron los siguientes:

- Desde la fecha del 31 de diciembre del 2019 al 19 de mayo del 2021 se cuentan un total de 32 Estados afectados de los 32 en total que componen la República mexicana y 2236 Municipios de los 2,446 que existen en la República.
- Los casos totales para cada situación es la siguiente:
 - Confirmados 2,390,126
 - Defunciones 221,071
 - Negativos 4,077,061
 - Sospechosos 437,834
- En el caso de los totales las ultimas 24 horas del 19 de mayo del 2021 tenemos:
 - Confirmados 1077
 - Defunciones 35
 - Negativos 5520
 - Sospechosos 3442
- En el TOP 5 de los Estados más afectados de la fecha del 19 de mayo del 2021 se encuentran:
 - El Distrito Federal con 371 casos confirmados
 - Quintana Roo con 60
 - Nuevo León con 59
 - Yucatán con 56
 - México con 55
- En el TOP 5 de Municipios más afectados de la fecha del 19 de mayo del 2021 se encuentran:
 - Álvaro de Obregón
 - Benito Juárez
 - Mérida
 - Iztapalapa
 - Azcapotzalco
- Las gráficas de líneas nos arrojan la siguiente información:
 - Para el caso de la gráfica con el total de confirmados por fecha podemos ver como del 31 de diciembre al 07 de enero del 2020 se encuentra un registro de solo un caso confirmado.
 - Ya para el 22 de marzo los casos iban en aumento en tan solo pocos meses.
 - Ya a partir de esa fecha anterior los casos se dispararon, se puede observar como los casos aumentan y disminuyen es decir se encuentran rebrotes de contagios.
 - Para el caso de la gráfica de total de defunciones por fecha podemos encontrar que el 17 de marzo se encontraron las primeras defunciones.
 - Para el 20 de enero del 2021 se puede observar que fue la fecha donde se hubo mucho más registros de defunciones por estados.



Figura 2.4 Cuadro de mando completo

4. CONCLUSIONES

El trabajo anterior presentado permitió hacer un análisis exhaustivo de cada estado; el panel es muy interactivo por eso se le agrego filtros para poder interactuar con las fechas y los estados y así obtener una información más precisa para cada estado. El mapa de la República mexicana colocado en el panel permite tener una mejor visualización y que se vea mucho más atractivo. Este trabajo es muy importante en la actualidad ya que hasta la fecha aun se siguen presentando casos confirmados de Covid-19, se espera para posibles trabajos futuros poder actualizar las fechas y tener un mejor análisis completo.

5. REFERENCIAS

- BI, M. P. (1 de junio de 2021). *Microsoft Power BI*. Obtenido de Microsoft Power BI: <https://powerbi.microsoft.com/es-es/>
- BI, M. P. (20 de mayo de 2021). *Microsoft Power BI*. Obtenido de <https://community.powerbi.com/t5/Themes-Gallery/bd-p/ThemesGallery>
- Ferrer, R. (2020). Pandemia por COVID-19: el mayor reto de la historia del intensivismo. *Medicina Intensiva*, 323-324.
- México, G. d. (19 de Mayo de 2021). *Covid-19 México Información General*. Obtenido de <https://datos.covid-19.conacyt.mx/#DOView>
- Suárez, V., Suarez Quezada, M., Oros Ruiz, S., & Ronquillo de Jesús, E. (2020). Epidemiología de COVID-19 en México: del 27 de febrero al 30 de abril de 2020. *Revista Clínica*, 463-471.
- Velázquez Pérez, L. C. (2020). La COVID-19: reto para la ciencia mundial. *Anales de la Academia de Ciencias de Cuba*, 763.