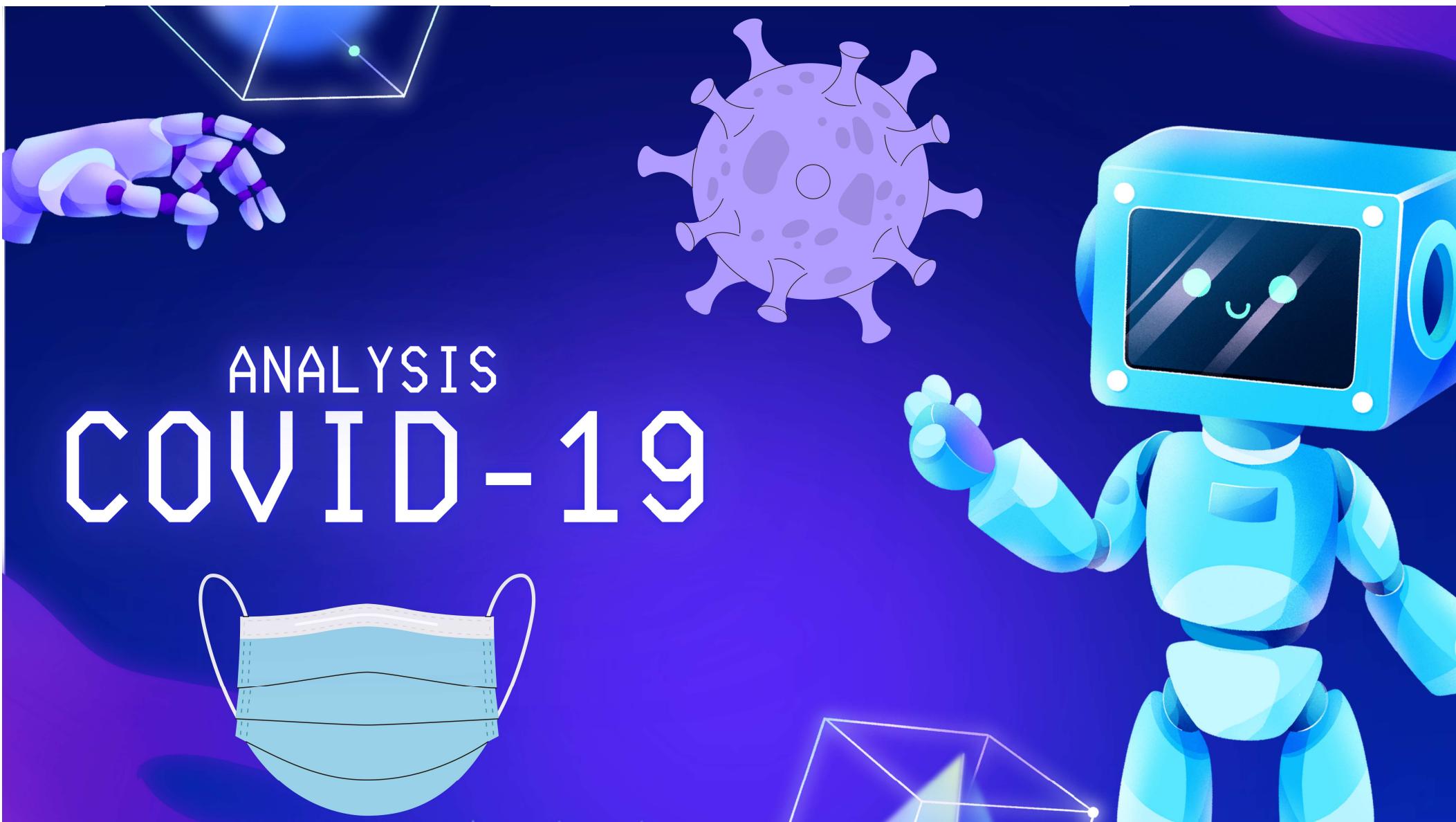


ANALYSIS COVID-19



INTRODUCTION

En este proyecto, nos adentramos en el análisis detallado de la pandemia del COVID-19, una crisis global que ha impactado a comunidades de todo el mundo. Enfocándonos en la situación específica de México, Perú y Colombia, nos proponemos explorar y visualizar datos relacionados con la propagación de este virus.

La importancia de comprender las dimensiones de esta crisis es evidente, especialmente para las entidades gubernamentales encargadas de la gestión de la salud en estos países.

El propósito fundamental es proporcionar información detallada y significativa que permita la toma de decisiones informadas y eficaces



ETAPAS DEL PROYECTO



RECOLECTAR DATA

Se recolectó data de las páginas oficiales de la ONU, tanto la información del COVID-19 como el compacto demográfico de todos los países.



EDA

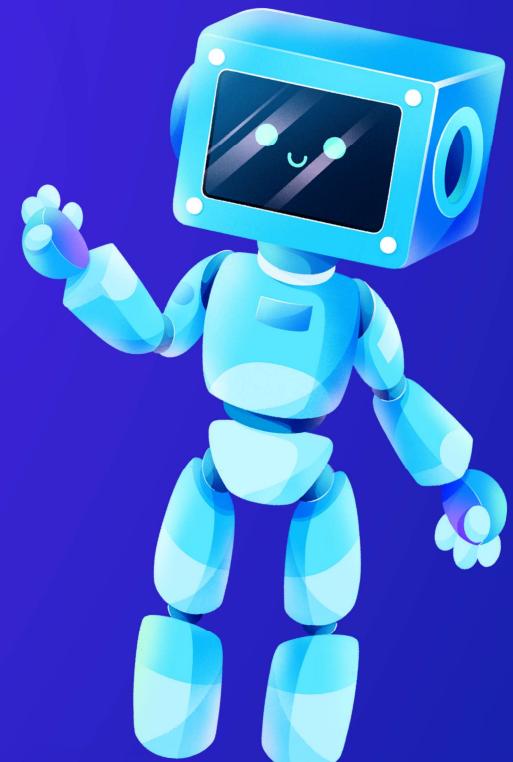
Procedimos a realizar un análisis profundo de la raw data recolectada, enfrentamos muchos desafíos durante el proceso. Conseguimos una data limpia y transformada al final.



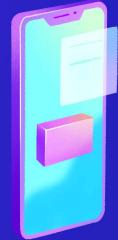
PREGUNTAS CLAVE

Se plantearon 7 preguntas clave, las cuales nos ayudaron a generar un entendimiento profundo de la data y sacar conclusiones efectivas.

RECOLECTAR DATA



RECOLECCIÓN DE LA DATA



Para el desarrollo del proyecto hemos empleado dos conjuntos de datos fundamentales, los cuales nos sirvieron para poder llevar a cabo todo nuestro análisis.

El primer conjunto de datos abarca información detallada sobre la propagación del virus COVID-19, casos y decesos diarios, como sus acumulados correspondientes. Así como países, fechas y códigos de países.

<https://covid19.who.int/WHO-COVID-19-global-data.csv>



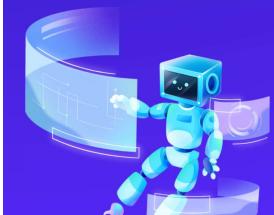
RECOLECCIÓN DE LA DATA



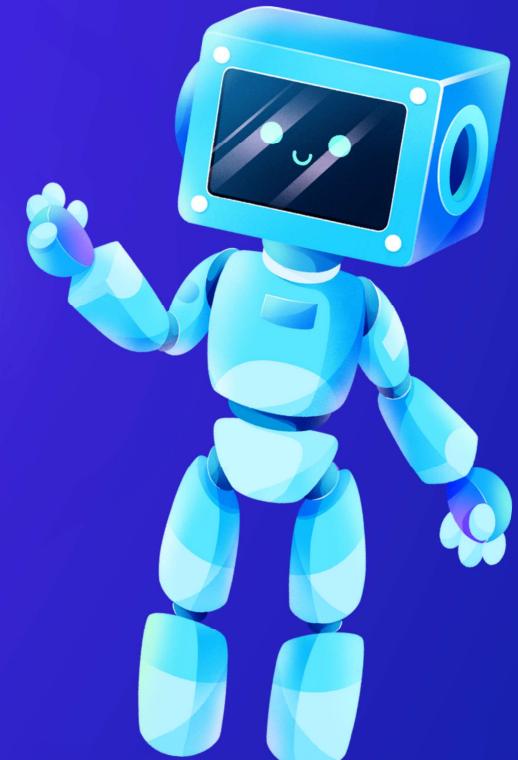
El segundo conjunto de datos se centra en el compacto demográfico mundial, donde nos brinda mucha información sobre población, densidad, mortalidad, fertilidad, entre otros muchos más indicadores.

https://raw.githubusercontent.com/davidcarrillo10288/Covid-19-analysis/master/WPP2022_GEN_F01_DEMOGRAPHIC_INDICATORS_COMPACT_REV1.xlsx

Trabajar en unión con estos dataset nos ayudaron a poder discriminar y entender como se desarrollo el virus a lo largo de toda la pandemia y que indicadores fueron más relevantes.



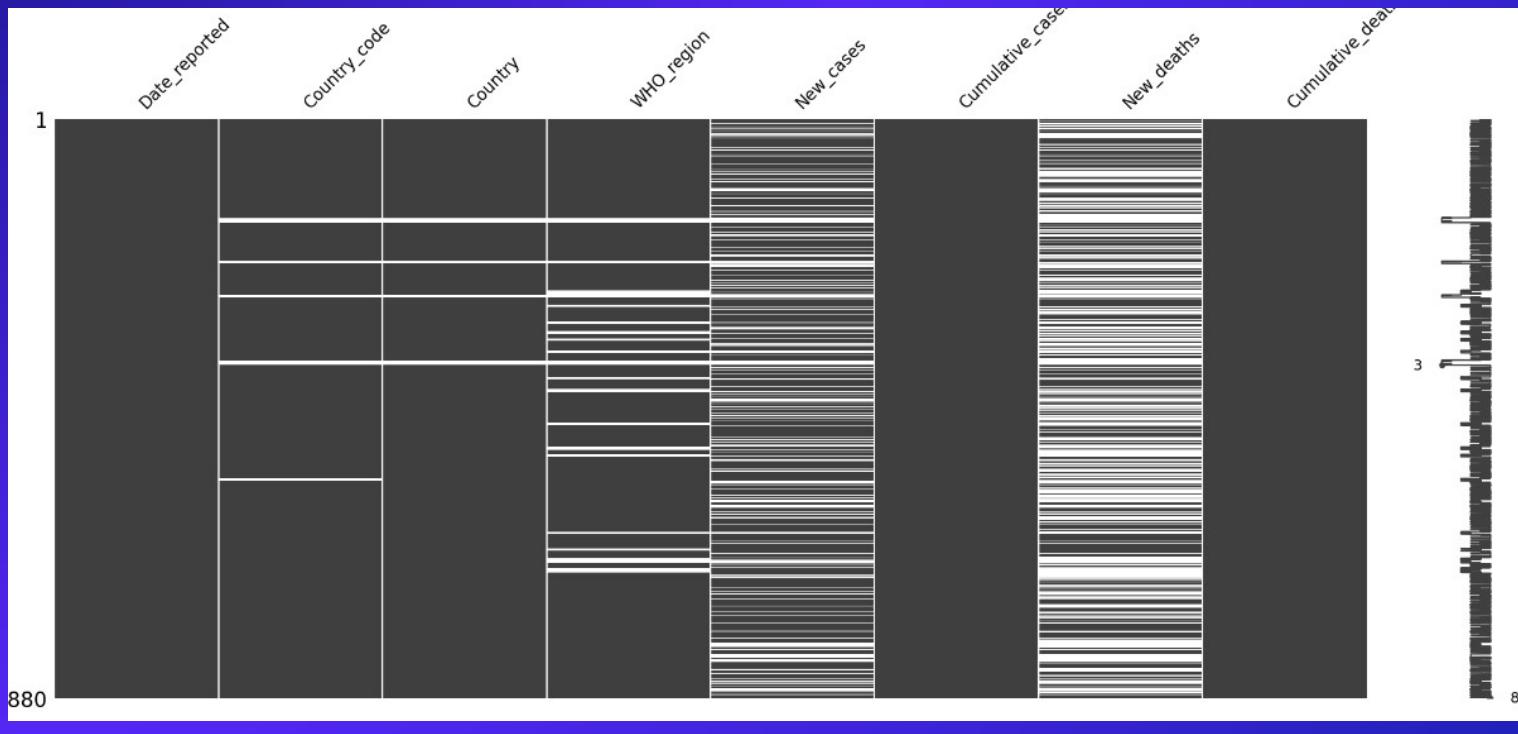
ANALISIS EXPLORATORIO DE DATOS



1. EDA BASE COVID WORLD HEALTH ORGANIZATION

En el análisis de datos sobre la pandemia del COVID-19, superamos desafíos cruciales:

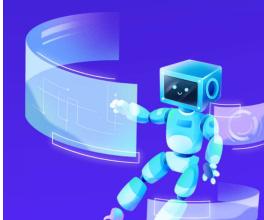
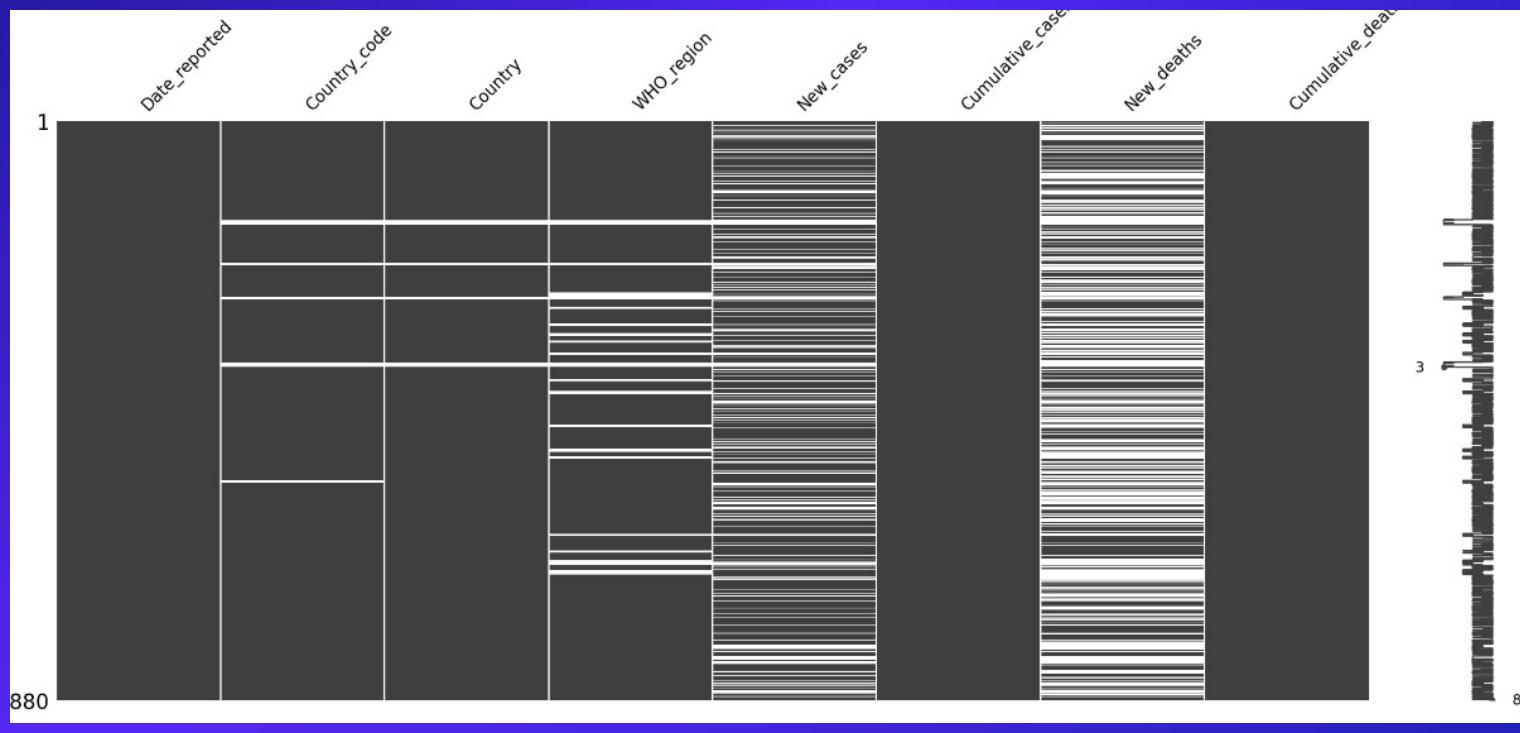
- Tratamos valores nulos con estrategias de imputación y eliminación.



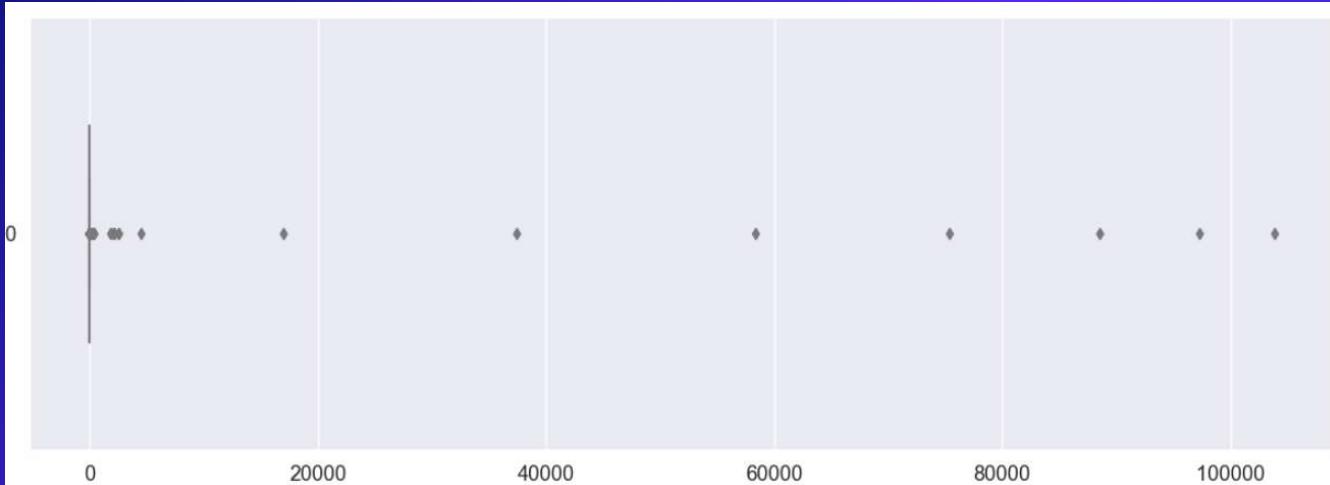
1. EDA BASE COVID WORLD HEALTH ORGANIZATION

En el análisis de datos sobre la pandemia del COVID-19, superamos desafíos cruciales:

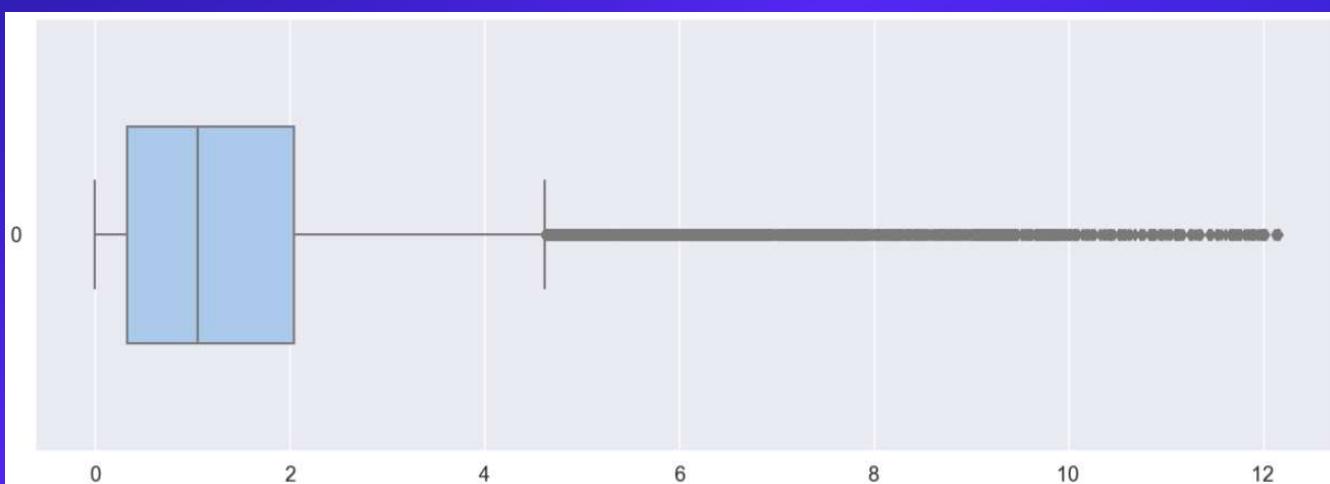
- Tratamos valores nulos con estrategias de imputación y eliminación.



- Introdujimos "lethality_rate" como nueva variable. Feature Engineering.
- Abordamos atípicos con truncamiento por percentiles.



index	lethality_rate
0	count 49393.000000
1	mean 11.773469
2	std 887.488947
3	min 0.000000
4	25% 0.334001
5	50% 1.055630
6	75% 2.047132
7	max 103875.000000



index	lethality_rate
0	count 49393.000000
1	mean 1.559370
2	std 1.935145
3	min 0.000000
4	25% 0.334001
5	50% 1.055630
6	75% 2.047132
7	max 12.156250

- Validamos datos negativos cotejándolos con fuentes oficiales.

<https://data.who.int/dashboards/covid19/cases>

<https://www.worldometers.info/coronavirus/worldwide-graphs/>

	Date_reported	Country_code	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths	lethality_rate
2515	2023-07-09	AU	Australia	10511	11541674	0	22770	0.197285
2516	2023-07-16	AU	Australia	7826	11549500	-76	22694	0.196493
2517	2023-07-23	AU	Australia	6495	11555995	85	22779	0.197118
2541	2024-01-07	AU	Australia	9626	11731031	33	24244	0.206666
2542	2024-01-14	AU	Australia	11114	11742145	86	24330	0.207202
2543	2024-01-21	AU	Australia	10502	11752647	21	24351	0.207196



- El resultado: un DataFrame limpio y confiable para un análisis preciso de la evolución de la pandemia.

	Date_reported	Country_code	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths	lethality_rate
0	2020-01-05	AF	Afghanistan	0	0	0	0	0.000000
1	2020-01-12	AF	Afghanistan	0	0	0	0	0.000000
2	2020-01-19	AF	Afghanistan	0	0	0	0	0.000000
3	2020-01-26	AF	Afghanistan	0	0	0	0	0.000000
4	2020-02-02	AF	Afghanistan	0	0	0	0	0.000000
...
49388	2023-12-24	ZW	Zimbabwe	23	266048	1	5731	2.154123
49389	2023-12-31	ZW	Zimbabwe	23	266071	0	5731	2.153936
49390	2024-01-07	ZW	Zimbabwe	46	266117	3	5734	2.154691
49391	2024-01-14	ZW	Zimbabwe	85	266202	3	5737	2.155130
49392	2024-01-21	ZW	Zimbabwe	55	266257	0	5737	2.154685

49393 rows × 8 columns

2. EDA BASE POPULATION UNITED NATIONS

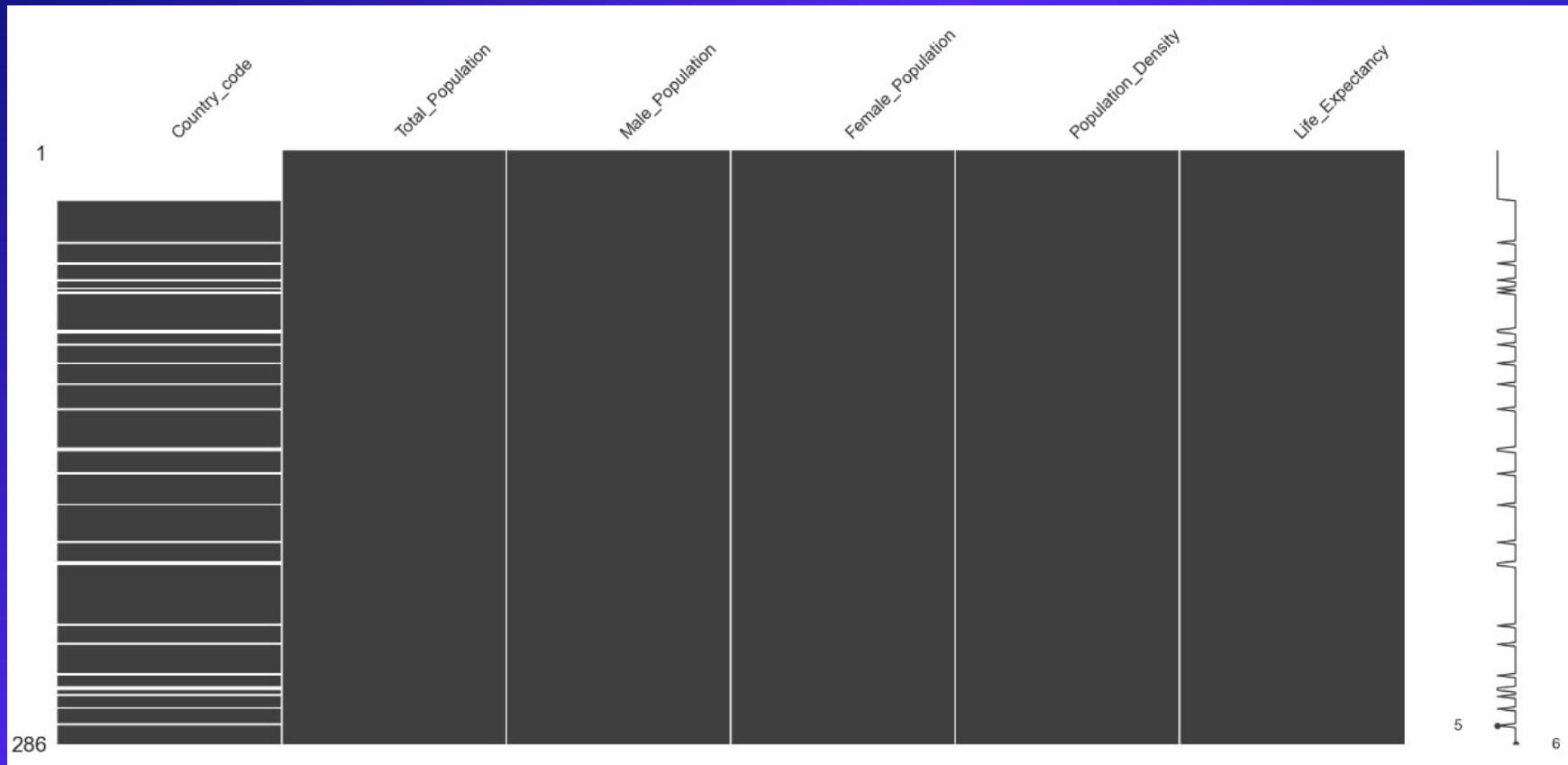
Durante el EDA de la población base según las Naciones Unidas, se abordaron desafíos específicos para garantizar la calidad de los datos:

- Seleccionamos 6 variables cruciales relacionadas con la población, densidad de población y esperanza de vida.

ISO2 Alpha-code	Total Population, as of 1 July (thousands)	Male Population, as of 1 July (thousands)	Female Population, as of 1 July (thousands)	Population Density, as of 1 July (persons per square km)	Life Expectancy at Birth, both sexes (years)	
69	NaN	7764951.032	3906407.855	3858543.178	59.546	72.79
142	NaN	1080677.346	538158.386	542518.959	49.425	60.725
214	NaN	530149.56	273104.233	257045.327	42.485	73.392
286	NaN	2021559.651	1033583.865	987975.786	195.683	70.615
358	NaN	2324282.002	1176368.32	1147913.681	143.117	77.013
...	
20305	WS	211.905	108.147	103.759	74.878	72.157
20377	TK	1.752	0.879	0.874	146	75.076
20449	TO	104.951	51.994	52.956	161.462	70.871
20521	TV	10.956	5.654	5.301	365.2	64.278
20593	WF	11.714	5.648	6.066	83.671	79.617

286 rows × 6 columns

- Identificamos y tratamos valores nulos, así como datos infinitos en el conjunto.
- Optamos por imputar valores nulos con ceros y eliminamos filas con datos infinitos para mantener la integridad del conjunto.



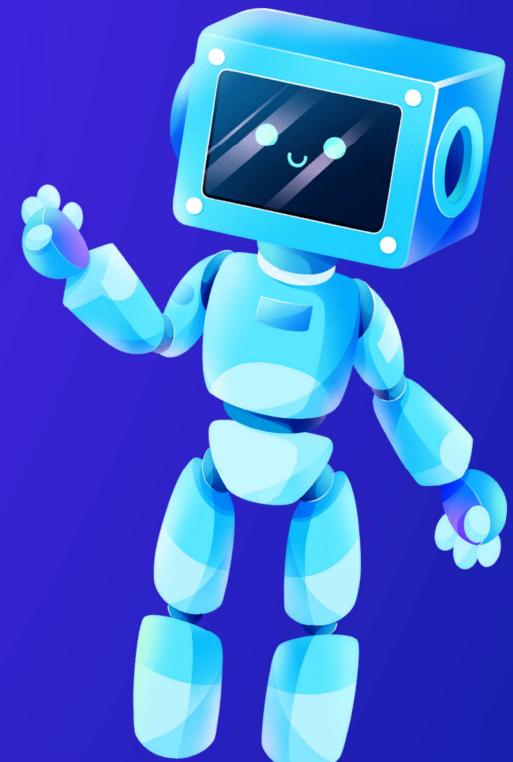
<https://population.un.org/wpp/Graphs/DemographicProfiles/Line/900>

- El resultado es un DataFrame depurado y focalizado, listo para proporcionar insights significativos sobre la dinámica demográfica y la calidad de vida

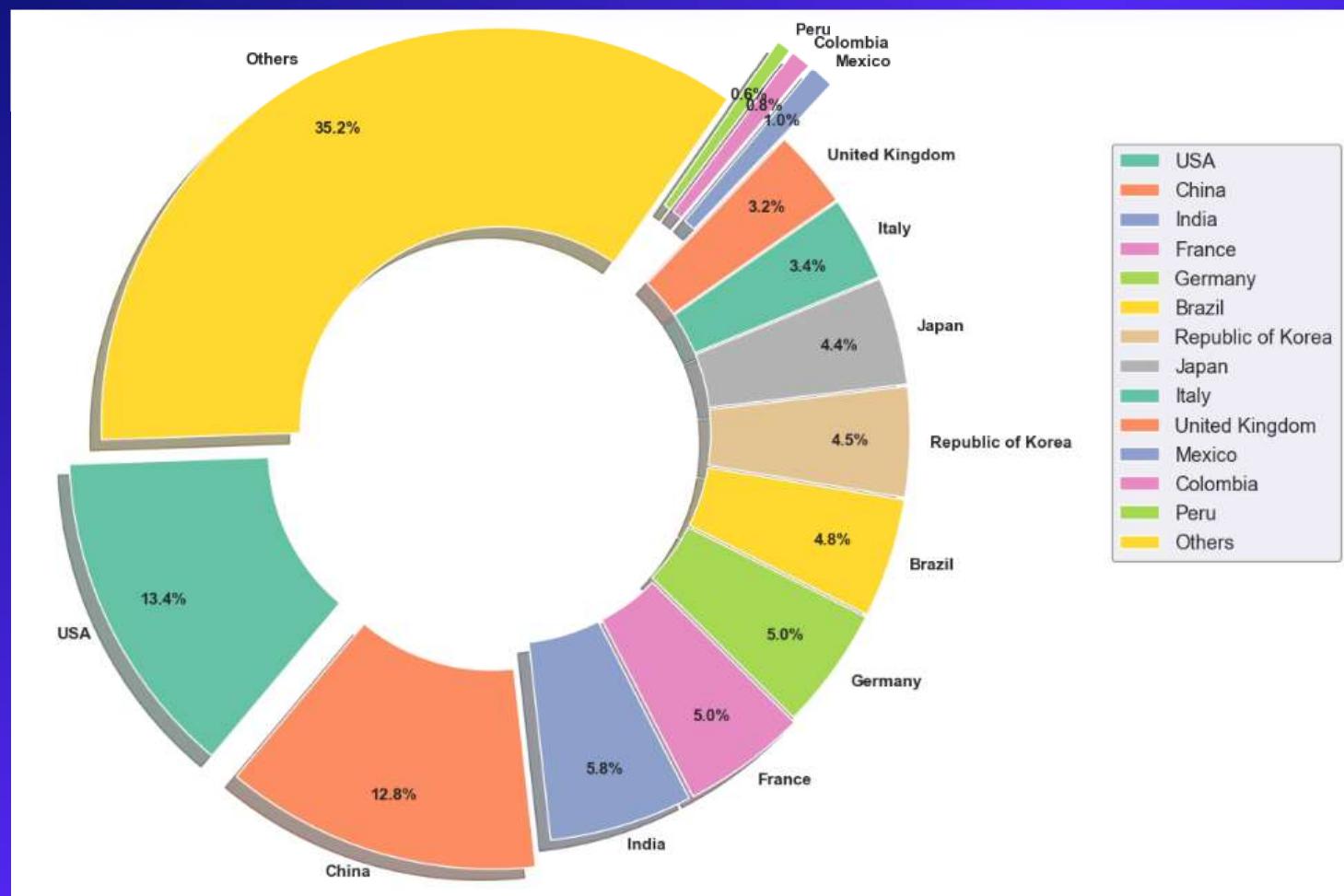
	Country_code	Total_Population	Male_Population	Female_Population	Population_Density	Life_Expectancy
0	BI	11874838.0	5894278.0	5980560.0	457.605	62.351
1	KM	790986.0	397126.0	393860.0	425.033	64.068
2	DJ	1073994.0	533626.0	540367.0	46.333	63.085
3	ER	3498818.0	1724860.0	1773958.0	28.881	67.300
4	ET	114120594.0	57375433.0	56745161.0	114.121	65.838
...
230	WS	211905.0	108147.0	103759.0	74.878	72.157
231	TK	1752.0	879.0	874.0	146.000	75.076
232	TO	104951.0	51994.0	52956.0	161.462	70.871
233	TV	10956.0	5654.0	5301.0	365.200	64.278
234	WF	11714.0	5648.0	6066.0	83.671	79.617

235 rows × 6 columns

PREGUNTAS CLAVE

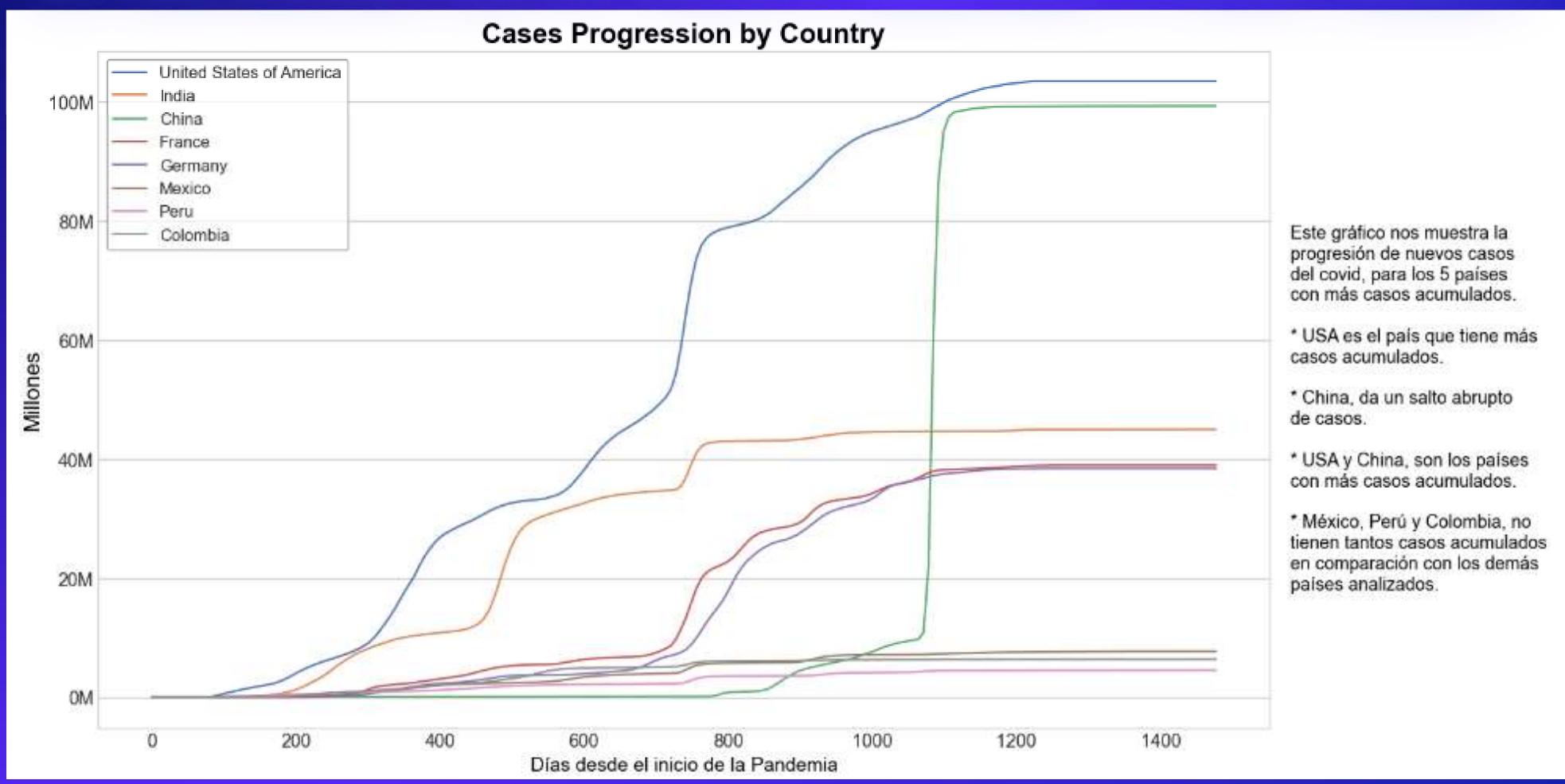


1- ¿CÓMO HA EVOLUCIONADO EL COVID-19 EN MÉXICO-PERÚ-COLOMBIA EN COMPARACIÓN CON EL IMPACTO OBSERVADO A NIVEL GLOBAL?

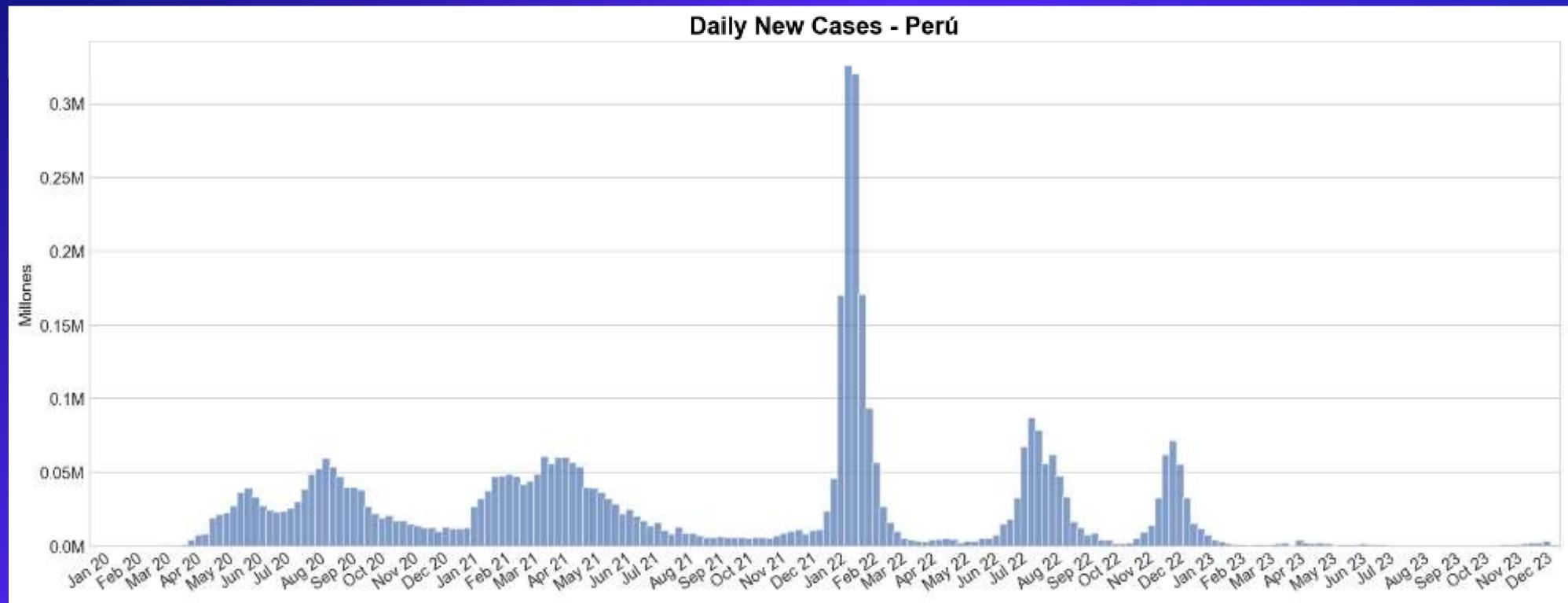


- USA y China son los países con más casos acumulados.
- Resto de países que comprenden el Top 10, tienen un porcentaje de casos acumulados similares que van del 5.8% al 3.2%.
- México-perú-Colombia, tiene porcentajes de casos acumulados más bajos, va en el rango del 1% al 0.5%.

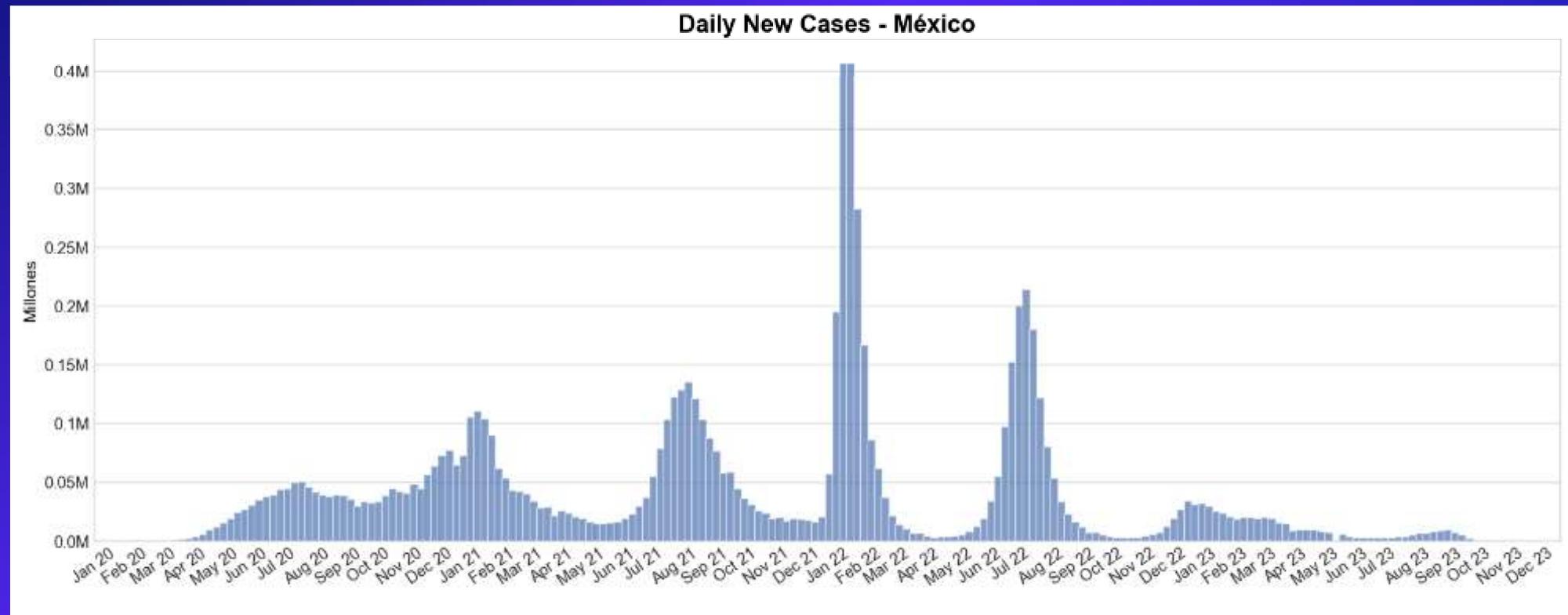
1- ¿CÓMO HA EVOLUCIONADO EL COVID-19 EN MÉXICO-PERÚ-COLOMBIA EN COMPARACIÓN CON EL IMPACTO OBSERVADO A NIVEL GLOBAL?



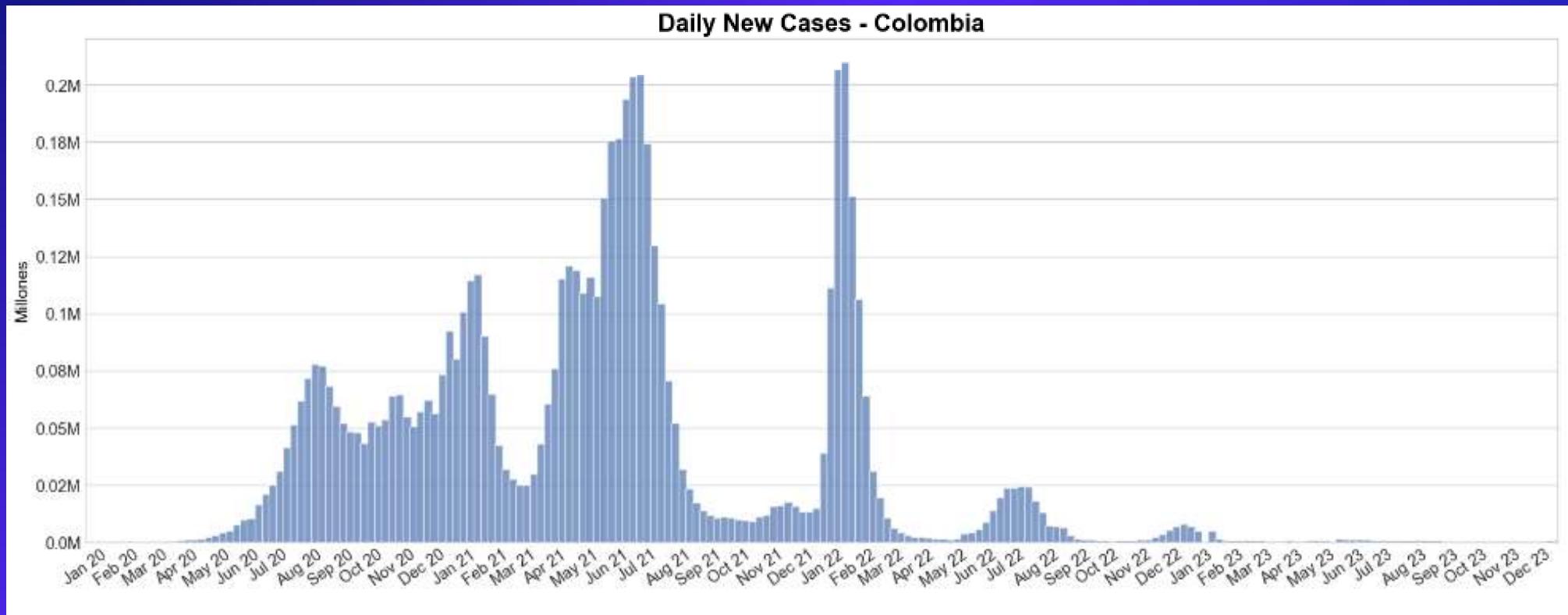
2- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS CASOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



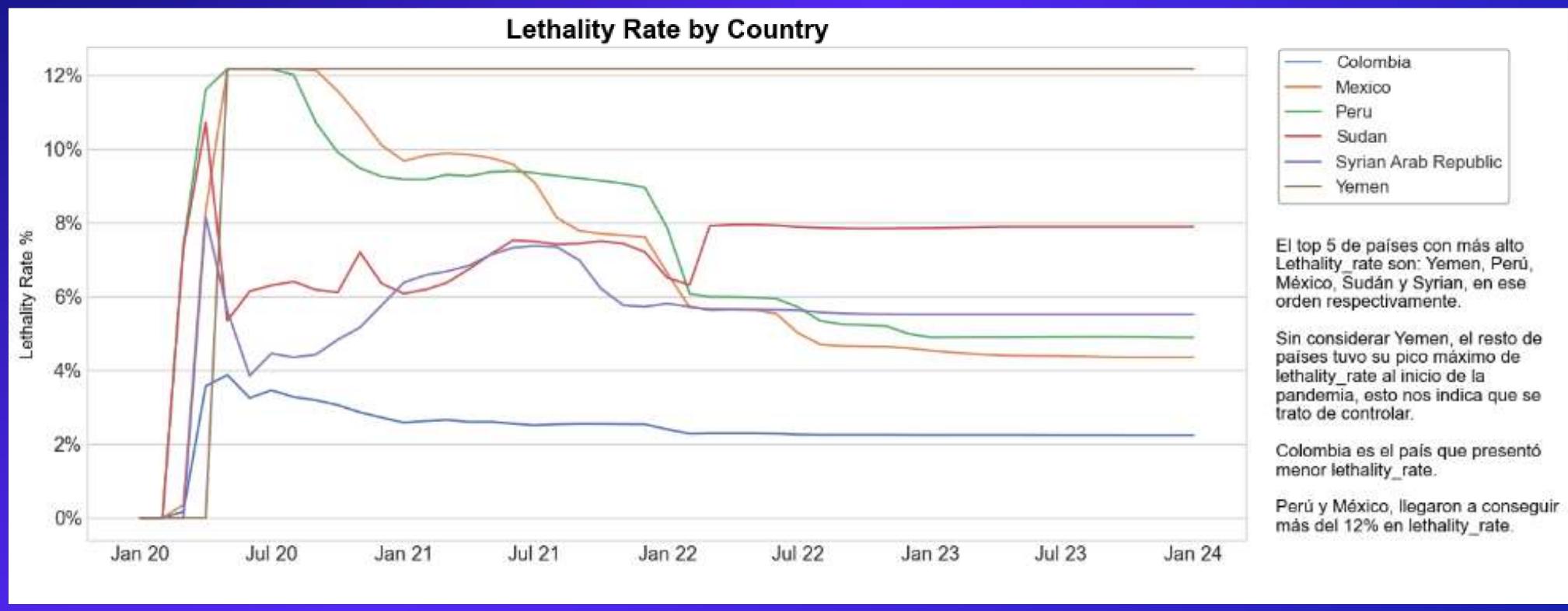
2- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS CASOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



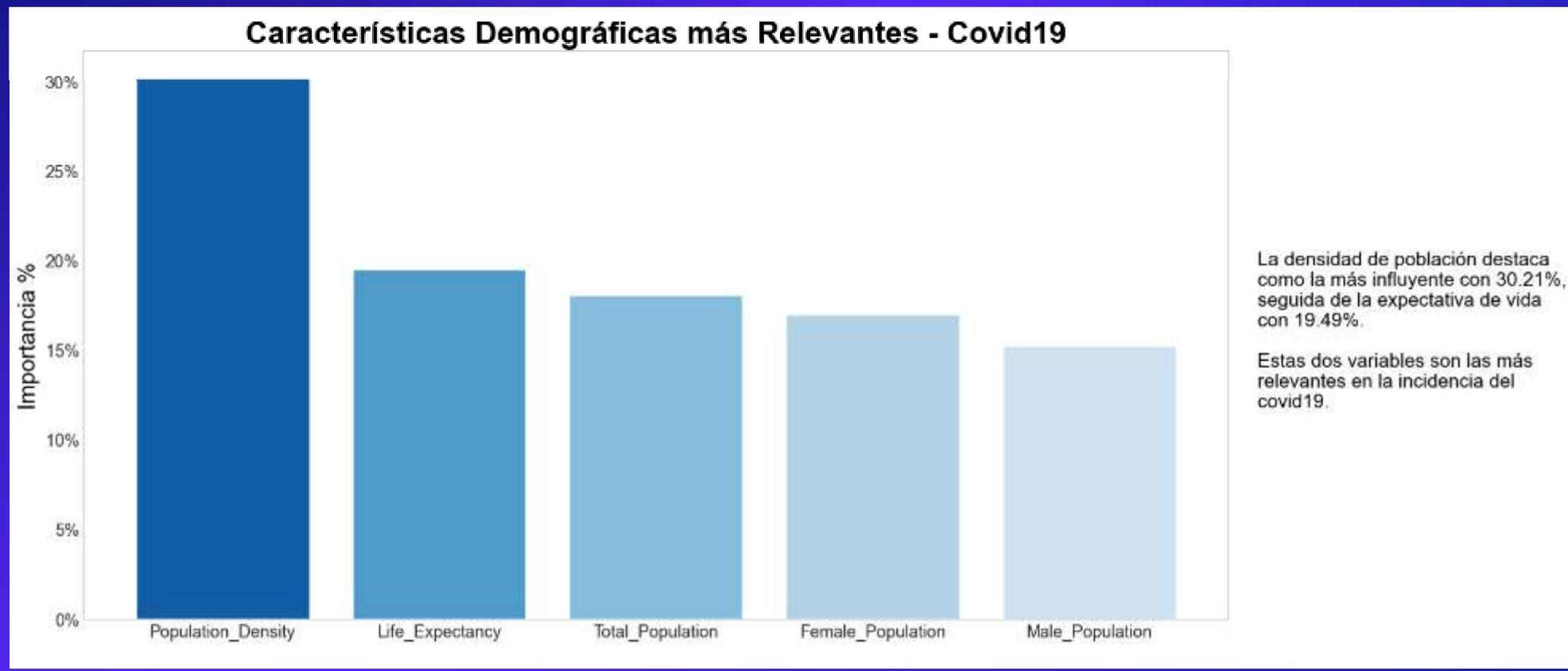
2- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS CASOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



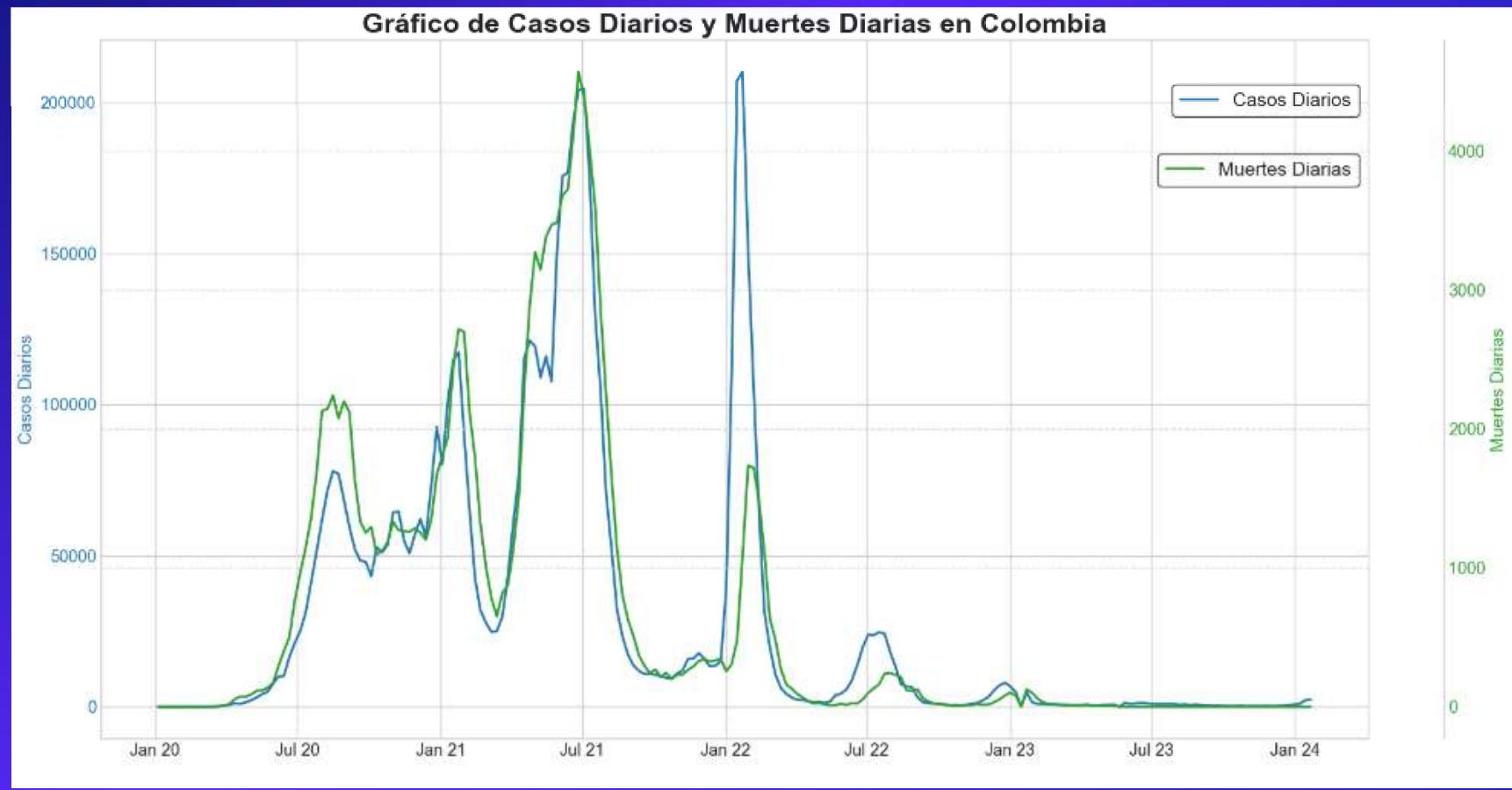
3- ¿CUÁL HA SIDO LA EVOLUCIÓN DEL ÍNDICE DE LETALIDAD DEL COVID-19 EN MÉXICO-PERÚ-COLOMBIA, COMPARADO CON LOS PAÍSES CON LOS ÍNDICES HISTÓRICOS MÁS ELEVADOS?



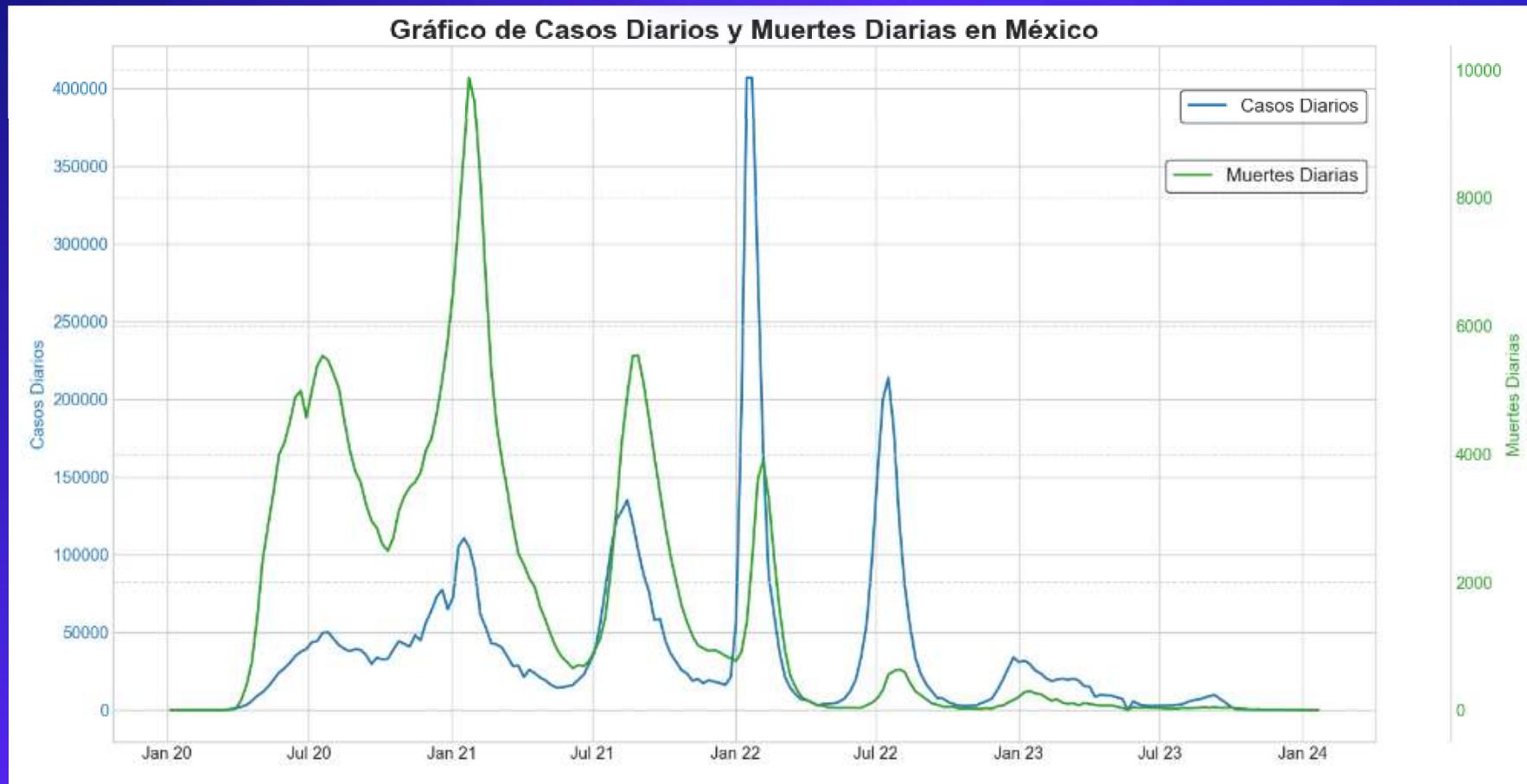
4- DESDE UNA PERSPECTIVA DEMOGRÁFICA, ¿CUÁLES HAN SIDO LAS CARACTERÍSTICAS QUE TUVIERON UN MAYOR IMPACTO EN EL ÍNDICE DE LETALIDAD EN MÉXICO-PERÚ-COLOMBIA?



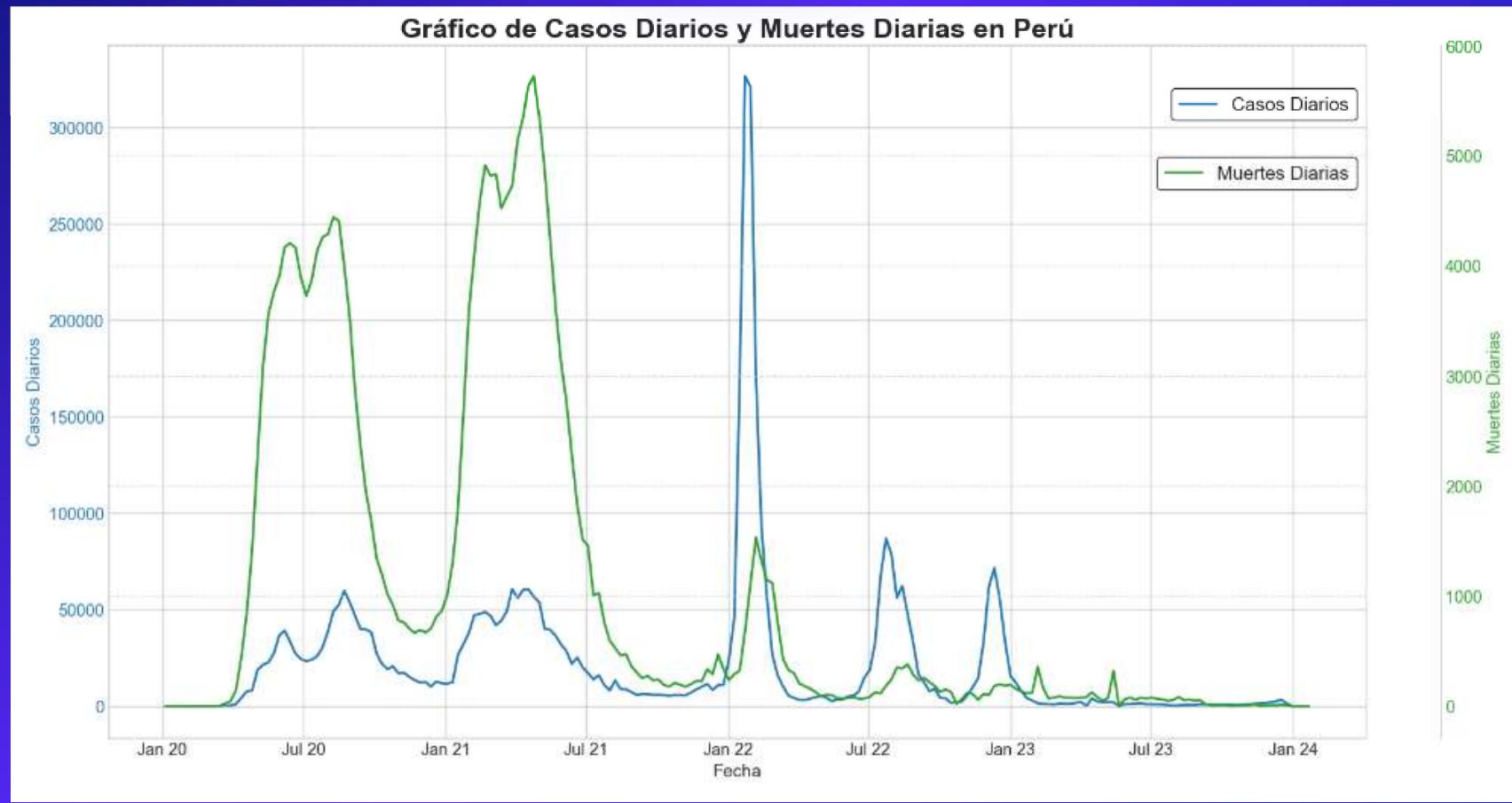
5- ¿RELACIÓN ENTRE EL AUMENTO DE LOS NUEVOS CASOS DIARIOS Y LOS NUEVOS DECESOS DIARIOS PARA LOS PAÍSES MÉXICO-PERÚ-COLOMBIA?



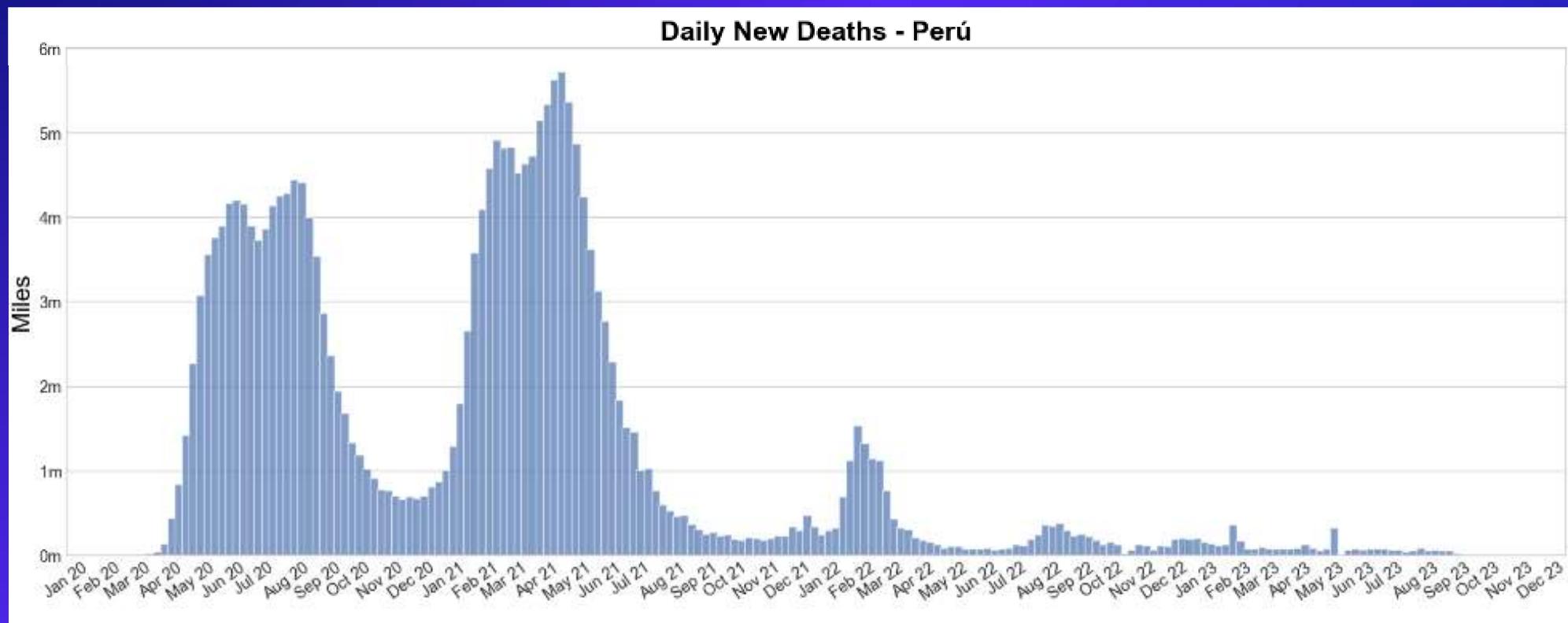
5- ¿RELACIÓN ENTRE EL AUMENTO DE LOS NUEVOS CASOS DIARIOS Y LOS NUEVOS DECESOS DIARIOS PARA LOS PAÍSES MÉXICO-PERÚ-COLOMBIA?



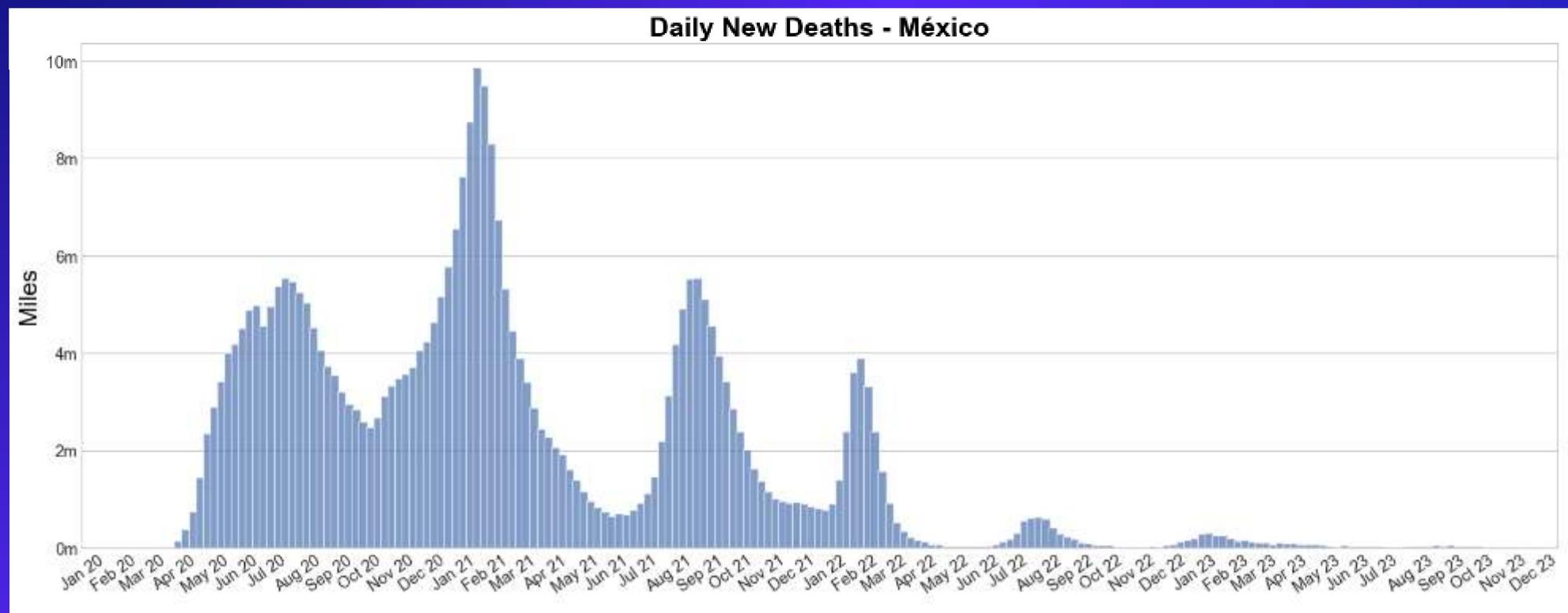
5- ¿RELACIÓN ENTRE EL AUMENTO DE LOS NUEVOS CASOS DIARIOS Y LOS NUEVOS DECESOS DIARIOS PARA LOS PAÍSES MÉXICO-PERÚ-COLOMBIA?



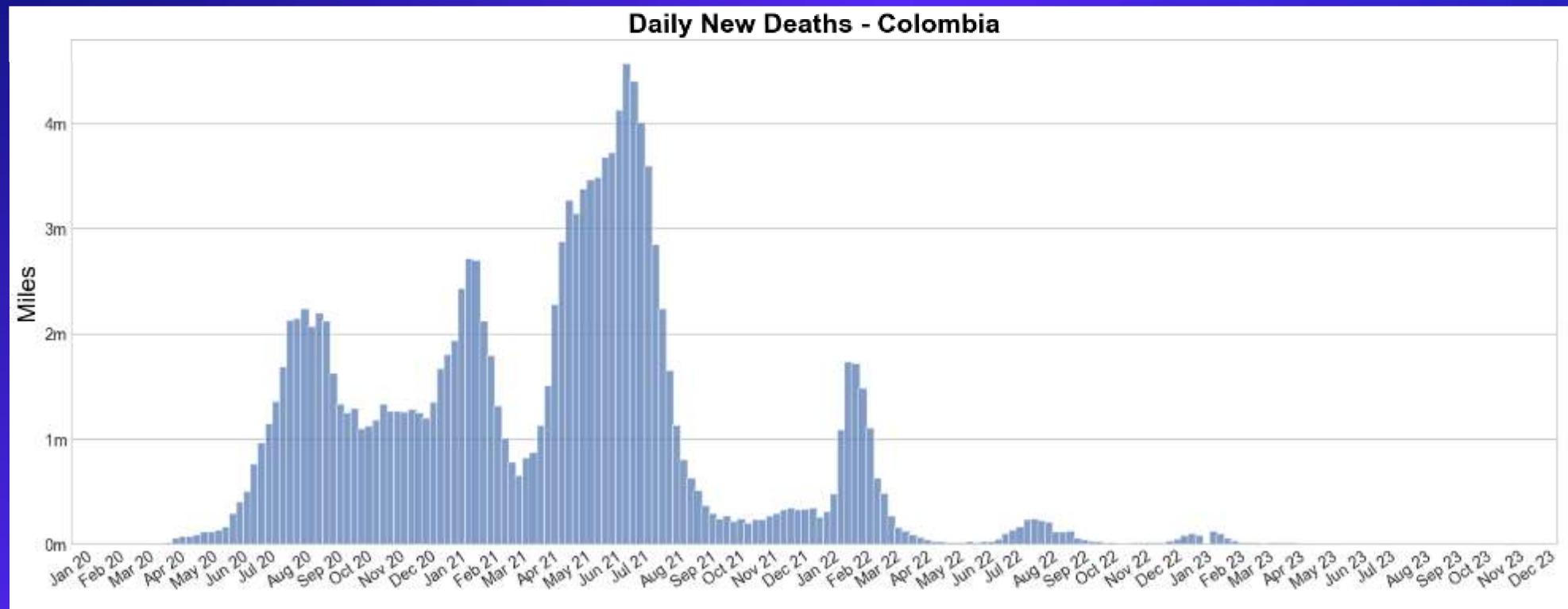
6- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS DECESOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



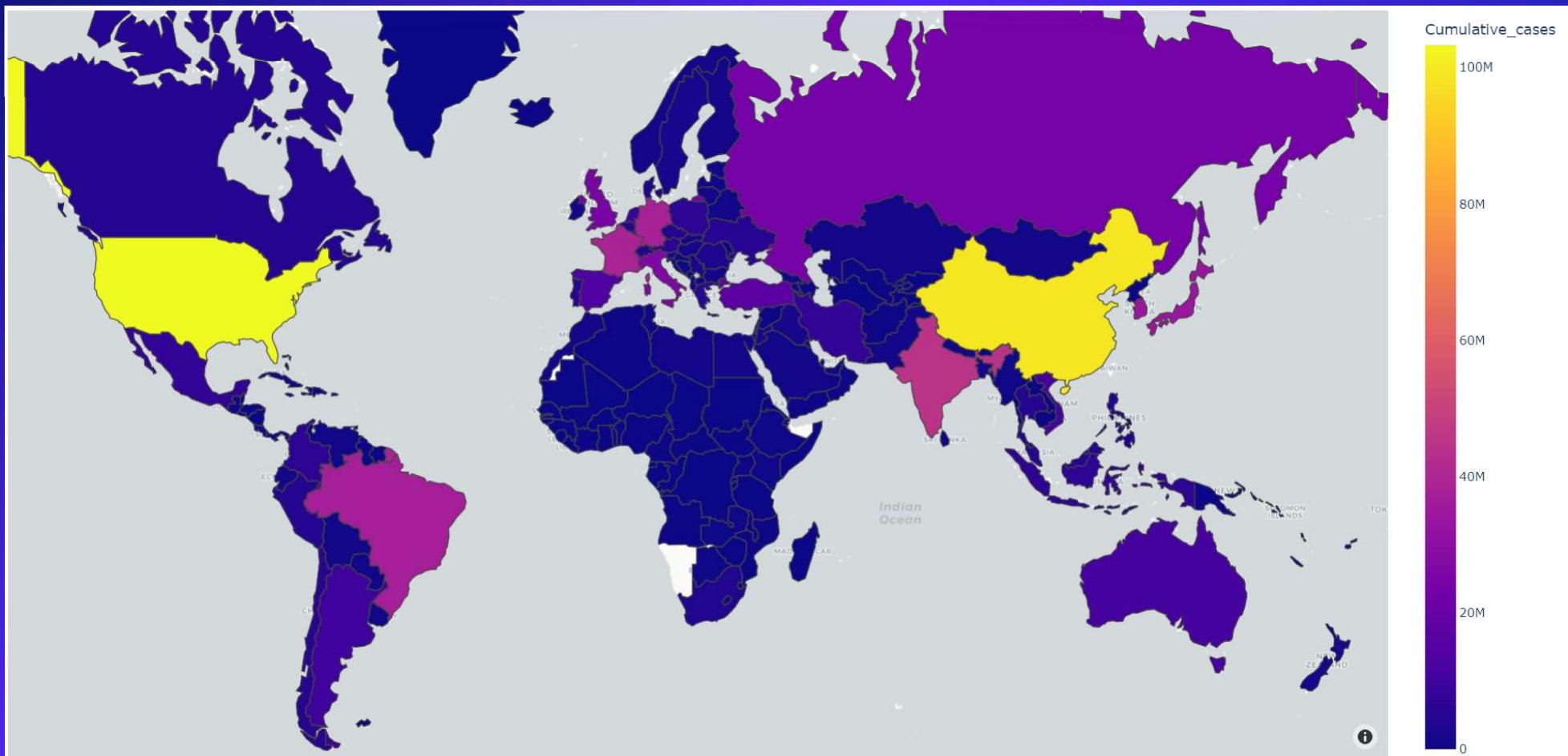
6- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS DECESOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



6- ¿CUÁL HA SIDO LA EVOLUCIÓN DE NUEVOS DECESOS DIARIOS REPORTADOS DE COVID-19 EN MÉXICO-PERÚ-COLOMBIA A LO LARGO DEL TIEMPO?



7- ¿CÓMO EVOLUCIONAN LOS CASOS ACUMULADOS EN TODOS LOS PAÍSES A NIVEL MUNDIAL - GRÁFICO COROPLÉTICO?



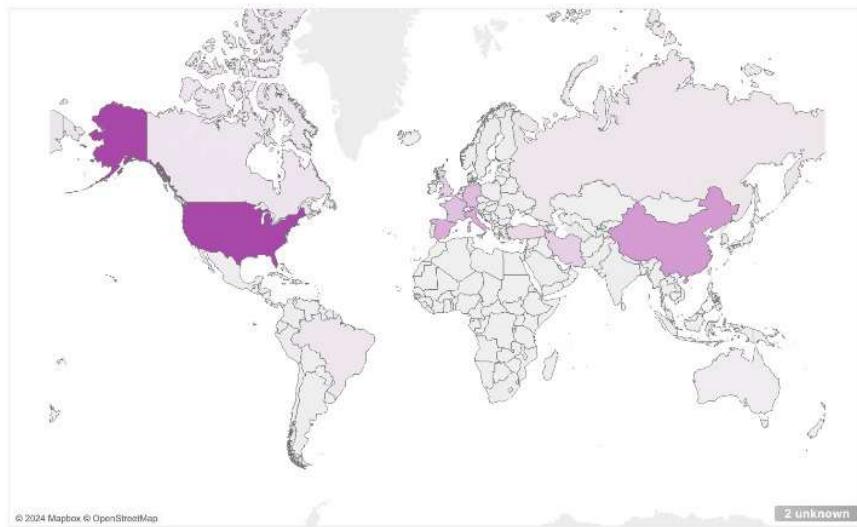
https://davidcarrillo10288.github.io/grafico-interactivo/grafico_interactivo.html

COVID-19: Distribución temporal en el mundo

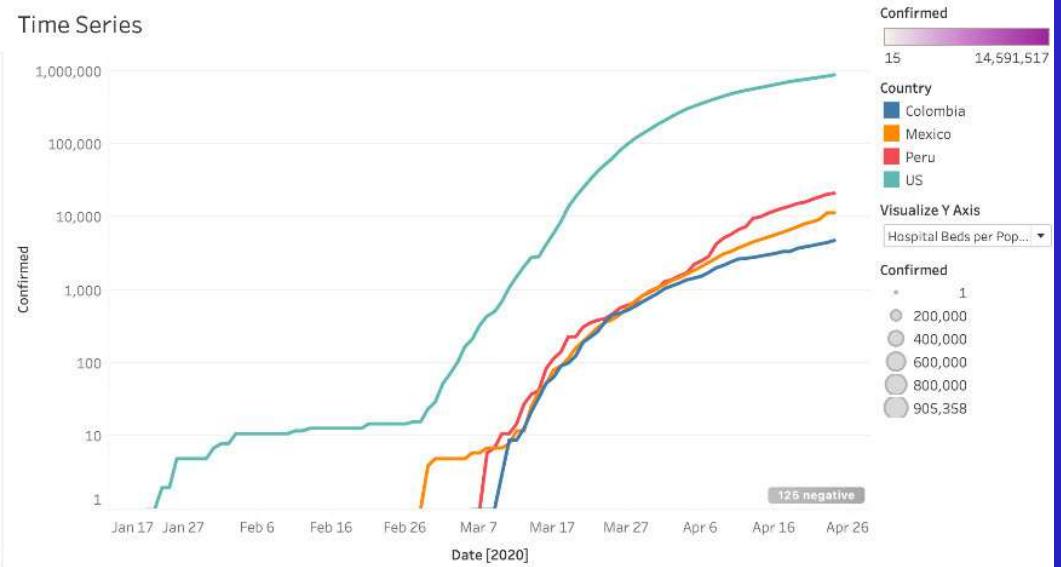


https://davidcarrillo10288.github.io/grafico-interactivo/covid_geo_chart.html

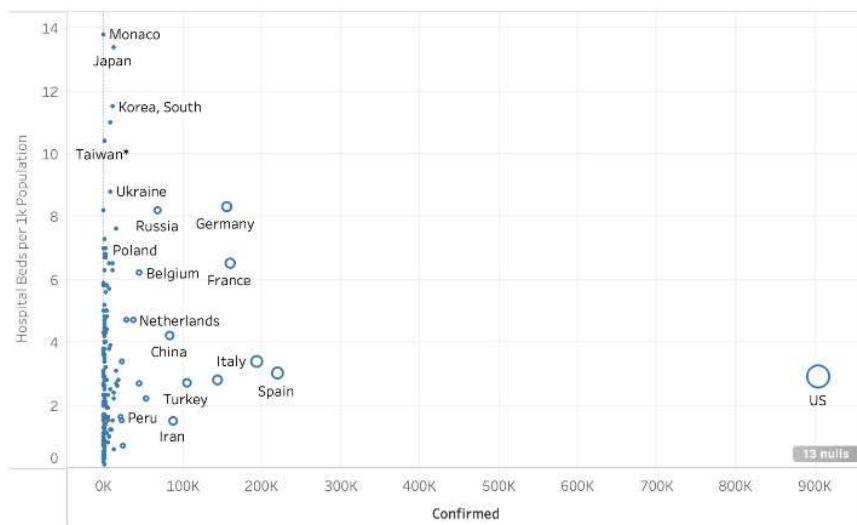
World Map



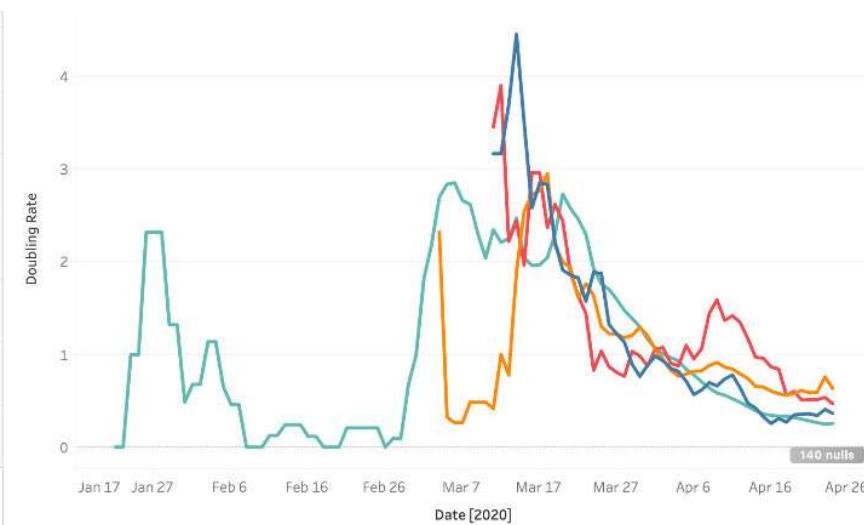
Time Series



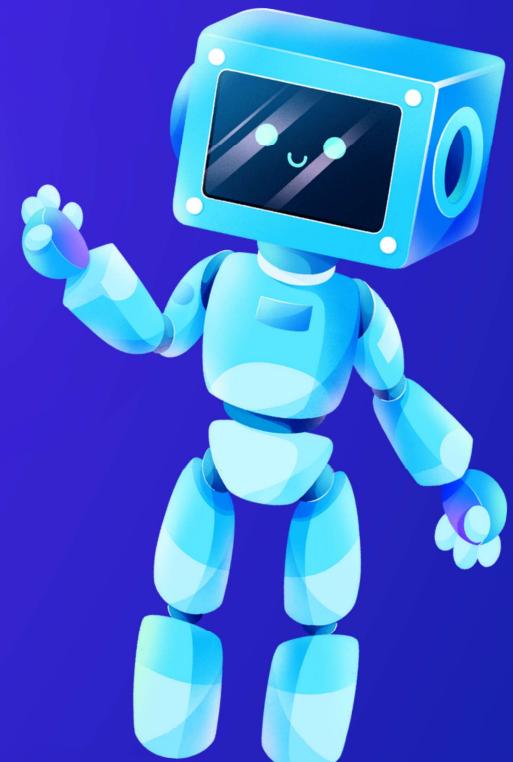
Statistics



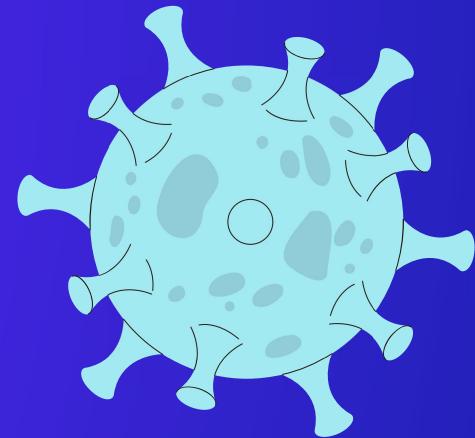
Doubling Rate



CONCLUSIONES FINALES



- USA y CHINA lideraron los casos y decesos de COVID-19.
- México enfrentó picos de casos y decesos, destacando la variante Delta. Problemas iniciales en la adopción de medidas pueden haber contribuido.
- Perú experimentó crisis sanitaria inicial, pero controló la situación con una vacunación eficiente, a pesar de desafíos con la variante Omicrón.



- Colombia sufrió crisis inicial, logrando cierto control con aumento en la vacunación, aunque variantes como Omicrón generaron preocupación intermitente.
- Cada país fue afectado de manera única por las variantes del COVID-19.
- El proceso de vacunación demostró resultados favorables en la disminución de casos y decesos, visible en los gráficos de implementación.

RECOMENDACIONES





Fortalecimiento y Concientización:

- **Reforzar medidas sanitarias y concientizar sobre su importancia.**
- **Invertir en centros de salud y proveer mascarillas gratuitas.**

Agilizar la Vacunación:

- **Establecer alianzas internacionales para el suministro de vacunas.**
- **Fomentar investigación nacional y participación de científicos locales.**

Mejora de Infraestructura de Salud:

- **Construir hospitales temporales, adquirir camas y equipos médicos.**
- **Desarrollar plantas de oxígeno para asegurar suministros.**

Adaptabilidad a Nuevas Variantes:

- **Desarrollar planes de respuesta rápida y establecer socios estratégicos.**
- **Garantizar acceso a nuevas vacunas y recursos hospitalarios.**



Programas Sociales:

- **Crear centros de apoyo psicológico para gestionar traumas post-pandemia.**

NUESTRO EQUIPO



David Carrillo Castillo

Data Scientist Junior

+51957639879

www.linkedin.com/in/davidcarrillocastillo

Lima, Perú



Wilfer Echavarria

Data Scientist Junior

+573005602181

<https://www.linkedin.com/in/wilfer-echavarria-bb19191b/>

Colombia



Mauricio

Data Scientist Junior

+522471081283

<https://www.linkedin.com/in/mauricio-quiroz-flores/>

México

