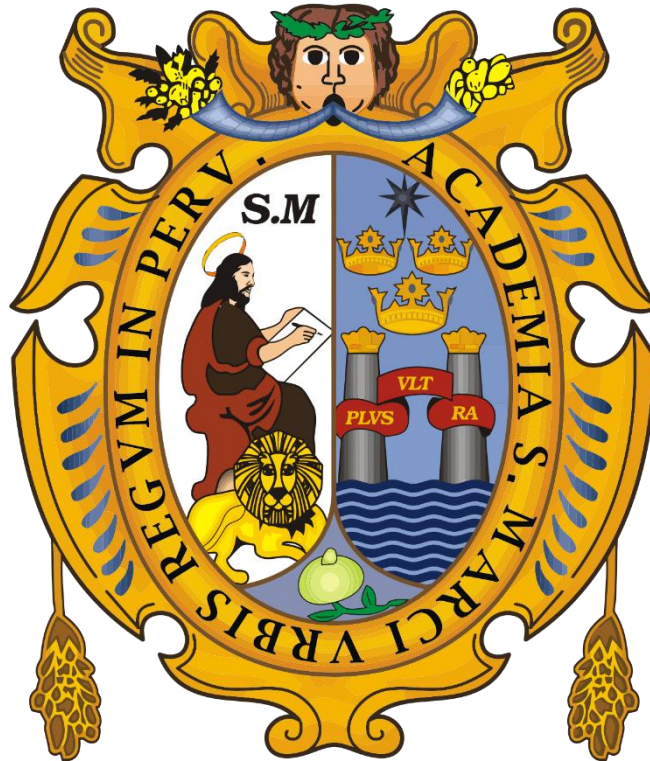


UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS  
FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
E.A.P DE INGENIERÍA DE SOFTWARE



**Proyecto**

Dashboard Global COVID-19

**Curso:** Base de Datos II

**Profesor:** Hugo Villaverde Medrano

**Alumno:** Reyes Cordova, Rodrigo Manuel

Lima, Perú

Agosto - 2022

## 1. Resumen

Este proyecto de Dashboard nace principalmente con el objetivo de tener forma gráfica un monitoreo de los cambios que ha venido y viene presentando el virus del SARS-CoV-2/COVID-19 a través de los diferentes países del mundo en los años 2021-2022.

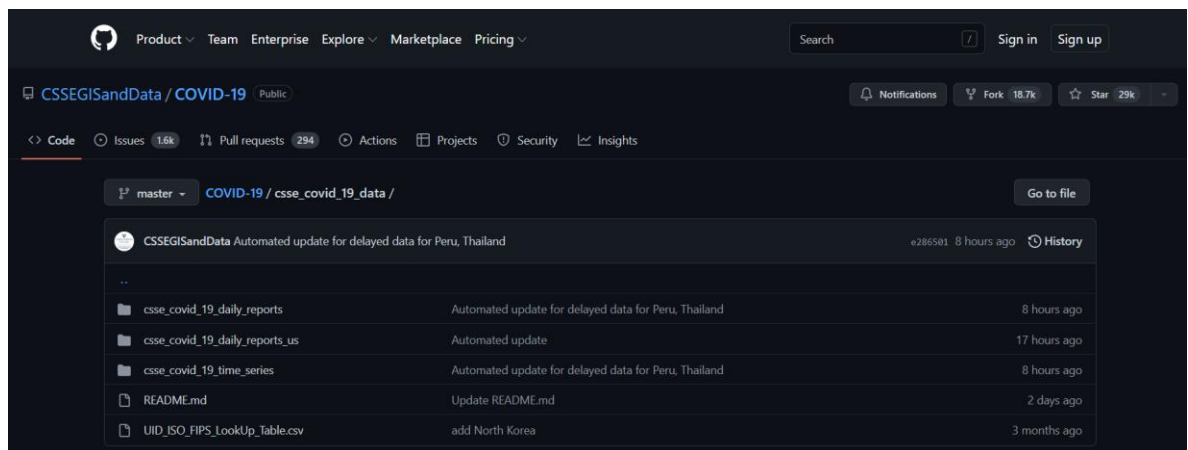
Este Dataset es proporcionado por el Centro de ciencia e ingeniería de sistemas de la Universidad Johns Hopkins (JHU CSSE). Además, con el apoyo del equipo Living Atlas de ESRI y el laboratorio de física aplicada de la Universidad Johns Hopkins (JHU APL).

Toda la data obtenida para este proyecto se descarga a través de un Script de código en lenguaje de programación Python automatizado, todos los días, con su limpieza y homologación respectiva.

Posteriormente estos datos son ingresados en un sistema de gestión de bases de datos relacional (PostgreSQL), de donde el Dashboard se alimenta.

## 2. Fuente de datos

Definimos las fuentes de información de los distintos datos extraídos.



- CSSEGISandData/COVID-19 (Fecha – País):  
Los siguientes enlaces contienen tablas de resumen de series temporales diarias, incluidas las confirmadas, las muertes y las recuperaciones.  
Todos los datos se leen del informe diario de casos. Las tablas de series

temporales están sujetas a actualización si se identifican imprecisiones en los datos históricos de los países.

- Confirmados:  
[https://github.com/CSSEGISandData/COVID-19/raw/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series/time\\_series\\_covid19\\_confirmed\\_global.csv](https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv)
- Fallecidos:  
[https://github.com/CSSEGISandData/COVID-19/raw/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series/time\\_series\\_covid19\\_deaths\\_global.csv](https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv)
- Recuperados:  
[https://github.com/CSSEGISandData/COVID-19/raw/master/csse\\_covid\\_19\\_data/csse\\_covid\\_19\\_time\\_series/time\\_series\\_covid19\\_recovered\\_global.csv](https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv)
- Vacunados:  
[https://github.com/govex/COVID-19/raw/master/data\\_tables/vaccine\\_data/global\\_data/time\\_series\\_covid19\\_vaccine\\_doses\\_admin\\_global.csv](https://github.com/govex/COVID-19/raw/master/data_tables/vaccine_data/global_data/time_series_covid19_vaccine_doses_admin_global.csv)

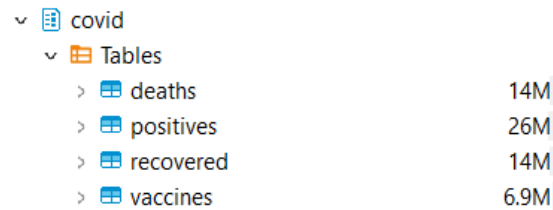
### 3. Script Extracción de Data

Se desarrollo un Script en Python para la descarga diaria de los archivos **csv**.

```
1 import glob
2 import pandas as pd
3 from func.functions import *
4
5 import warnings
6 warnings.filterwarnings("ignore")
7
8 """
9 Links de Descarga
10 """
11 POSITIVE_URL = 'https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv'
12
13 DEATHS_URL = 'https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv'
14
15 RECOVER_URL = 'https://github.com/CSSEGISandData/COVID-19/raw/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv'
16
17 VACCINE_URL = 'https://github.com/govex/COVID-19/raw/master/data_tables/vaccine_data/global_data/time_series_covid19_vaccine_doses_admin_global.csv'
18
19
20 """ Función Principal """
21 def main():
22
23     """ Descarga de archivos .csv's """
24     downlad_file(POSITIVE_URL,'world covid data/positives','.csv')
25     downlad_file(DEATHS_URL,'world covid data/deaths','.csv')
26     downlad_file(RECOVER_URL,'world covid data/recovered','.csv')
27     downlad_file(VACCINE_URL,'world covid data/vaccine','.csv')
28
29     """ Lista de archivos (POSITIVOS - FALLECIDOS - RECUPERADOS - VACUNACIONES) """
30     positive_covid_files = glob.glob("world covid data/positives/*.csv", recursive=True)
31     death_covid_files = glob.glob("world covid data/deaths/*.csv", recursive=True)
32     recovered_covid_files = glob.glob("world covid data/recovered/*.csv", recursive=True)
33     vaccines_list = glob.glob("world covid data/vaccine/*.csv", recursive=True)
34
35     """ Lectura de archivos descargados el día de hoy """
36     positives = pd.read_csv(positive_covid_files[-1])
37     deaths = pd.read_csv(death_covid_files[-1])
38     recovered = pd.read_csv(recovered_covid_files[-1])
39     vaccine_df = pd.read_csv(vaccines_list[-1])
40
41     """ Limpieza de datos y homologación """
42     positives = clean_covid(positives)
43     deaths = clean_covid(deaths)
44     recovered = clean_covid(recovered)
45     vaccine_df = clean_vaccines(vaccine_df)
46
47     """ Importando Data a Tabla ya creada en Base de Datos PostgreSQL """
48     dataframe_import(positives, 'covid.positives')
49     dataframe_import(deaths, 'covid.deaths')
50     dataframe_import(recovered, 'covid.recovered')
51     dataframe_import(vaccine_df, 'covid.vaccines')
52
53     """ Refrescando Vista Materializada """
54     refresh_materialied_view()
55
56 #Ejecución de Script
57 if __name__ == '__main__':
58
59     # NaN Values
60     psycopg2.extensions.register_adapter(float, nan_to_null)
61
62 #Ejecutando Script
63 main()
64
```

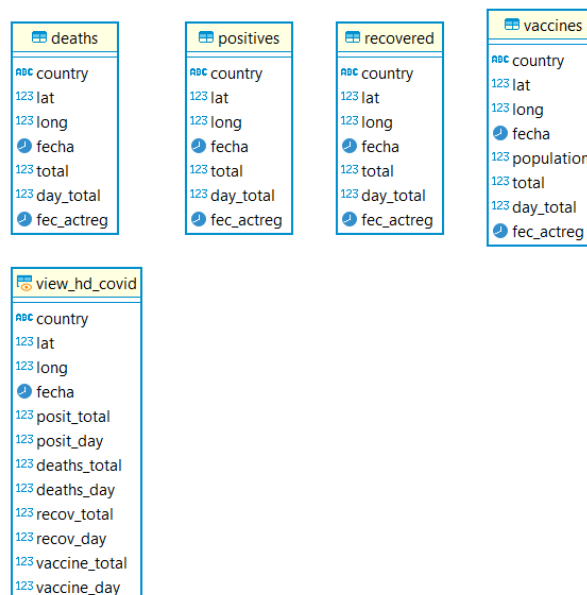
Código del archivo principal: **main.py**

Para finalmente terminar almacenadas en 4 tablas en una base de datos PostgreSQL.



▼ covid	
▼ Tables	
> deaths	14M
> positives	26M
> recovered	14M
> vaccines	6.9M

#### Tablas de destino en Base de Datos en el Esquema **covid**



deaths	positives	recovered	vaccines
<ul style="list-style-type: none"> <li>country</li> <li>lat</li> <li>long</li> <li>fecha</li> <li>total</li> <li>day_total</li> <li>fec_actreg</li> </ul>	<ul style="list-style-type: none"> <li>country</li> <li>lat</li> <li>long</li> <li>fecha</li> <li>total</li> <li>day_total</li> <li>fec_actreg</li> </ul>	<ul style="list-style-type: none"> <li>country</li> <li>lat</li> <li>long</li> <li>fecha</li> <li>total</li> <li>day_total</li> <li>fec_actreg</li> </ul>	<ul style="list-style-type: none"> <li>country</li> <li>lat</li> <li>long</li> <li>fecha</li> <li>population</li> <li>total</li> <li>day_total</li> <li>fec_actreg</li> </ul>

view_hd_covid
<ul style="list-style-type: none"> <li>country</li> <li>lat</li> <li>long</li> <li>fecha</li> <li>posit_total</li> <li>posit_day</li> <li>deaths_total</li> <li>deaths_day</li> <li>recov_total</li> <li>recov_day</li> <li>vaccine_total</li> <li>vaccine_day</li> </ul>

Estructura de Tablas (positives – recovered – deaths – vaccines)

## 4. Indicadores y Dashboard

Para la creación del Dashboard se crearon 12 gráficos para obtener información de forma sencilla con Power BI sobre la data ya antes mencionada.

### a) Tabla Muertos Totales por País

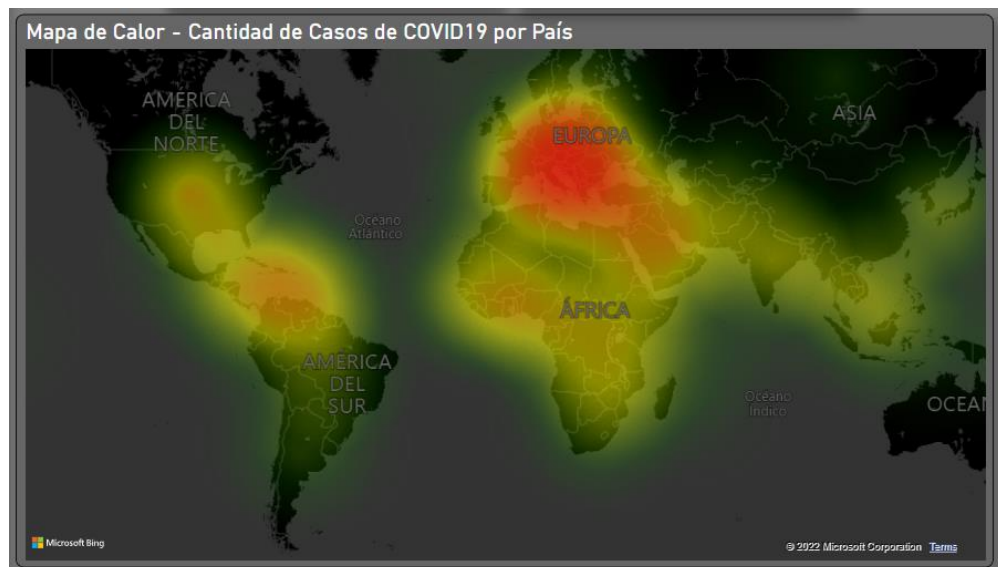
Esta tabla lo que hace es desplegar el total fallecidos por país, mostrando el campo de cantidad de fallecidos con una degradación para identificar el país con mayor cantidad de fallecidos (gradiente rojo)

hasta los que presentan una menor cantidad de fallecidos (gradiente verde).

Muertes Totales por País	
United States	1,039,026.00
Brazil	682,010.00
India	527,206.00
Russia	375,532.00
Mexico	328,871.00
Perú	215,088.00
Italy	174,300.00
Indonesia	157,296.00
Germany	146,214.00
Iran	143,160.00
Colombia	141,287.00
Argentina	129,566.00
Poland	116,798.00
Ukraine	116,511.00
Spain	111,906.00
South Africa	102,066.00

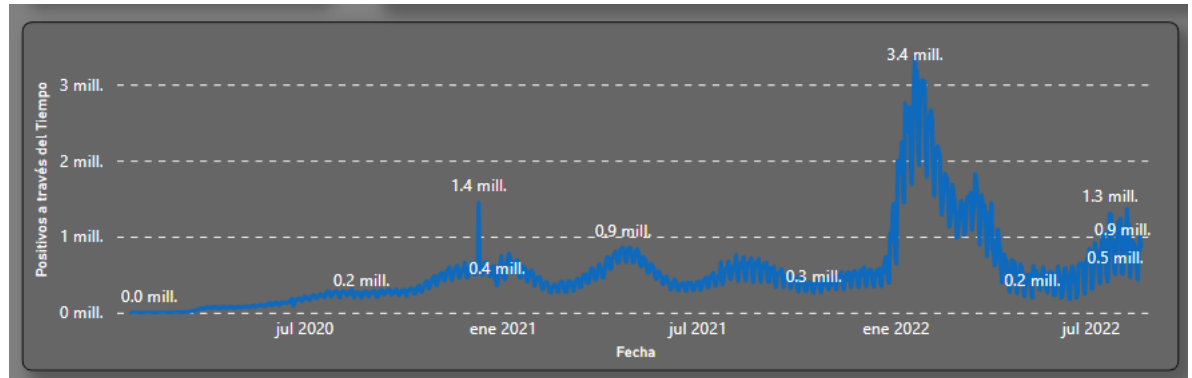
b) Mapa de Calor – Cantidad de Positivos

Este mapa de calor permite saber donde es que se concentra cantidad de casos de COVID positivos alrededor del mundo por su posición geográfica, representado en un mapa.



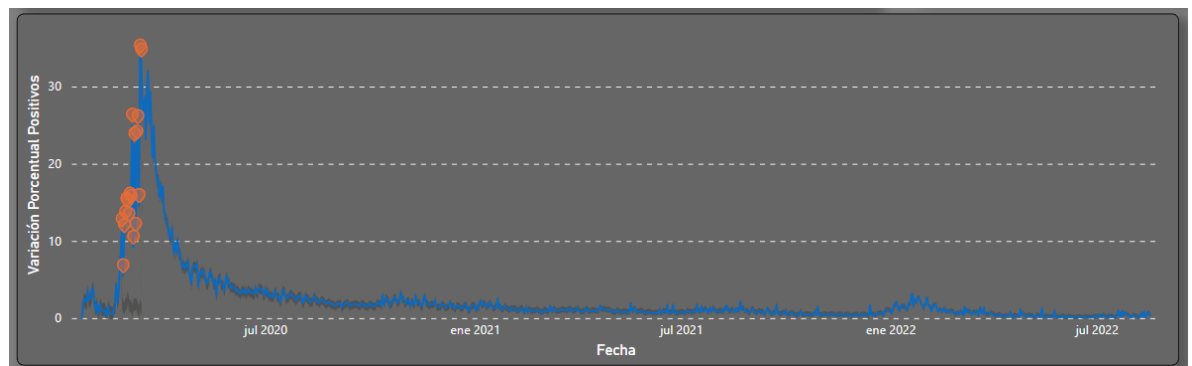
c) Serie de Tiempo – Positivos

Identificar el aumento y descenso de Casos Positivos del virus durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).



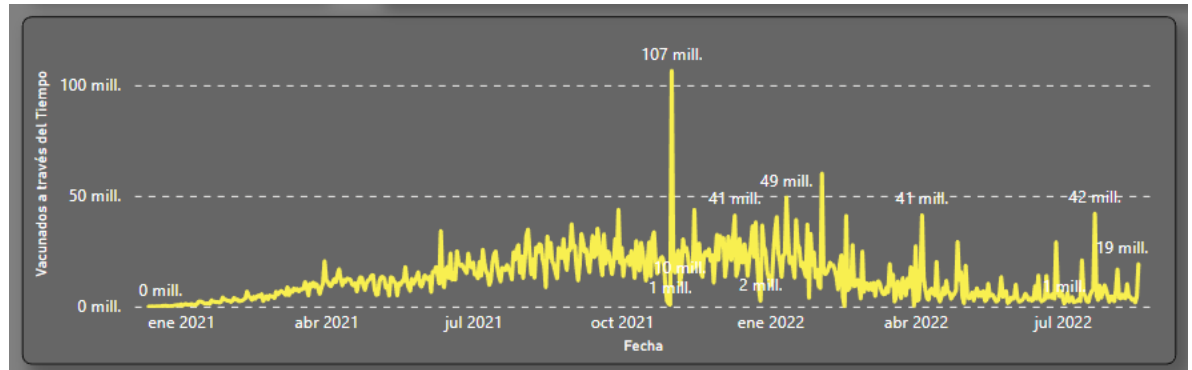
d) Serie de Tiempo – Variación Porcentual de Positivos

Análisis de como el porcentaje de Casos Positivos ha cambiado durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).



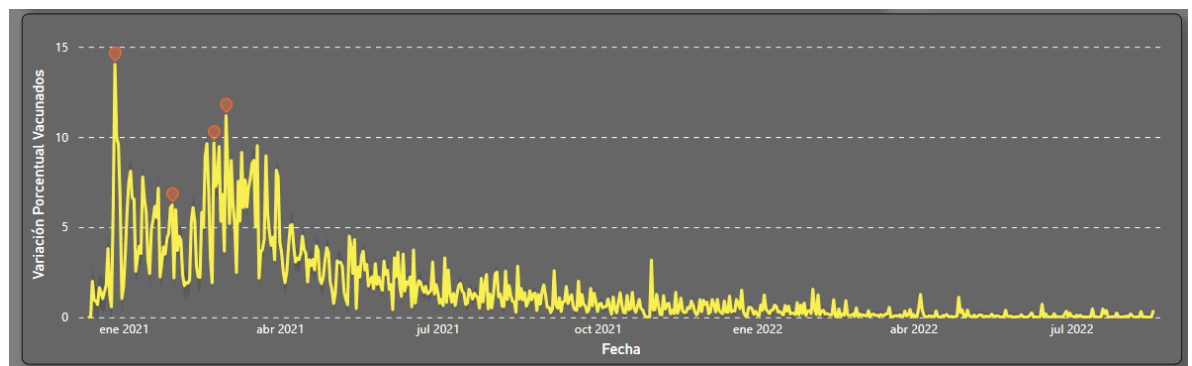
e) Serie de Tiempo – Vacunados

Identificar el aumento y descenso de Vacunas contra el virus durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).



f) Serie de Tiempo – Variación Porcentual de Vacunados

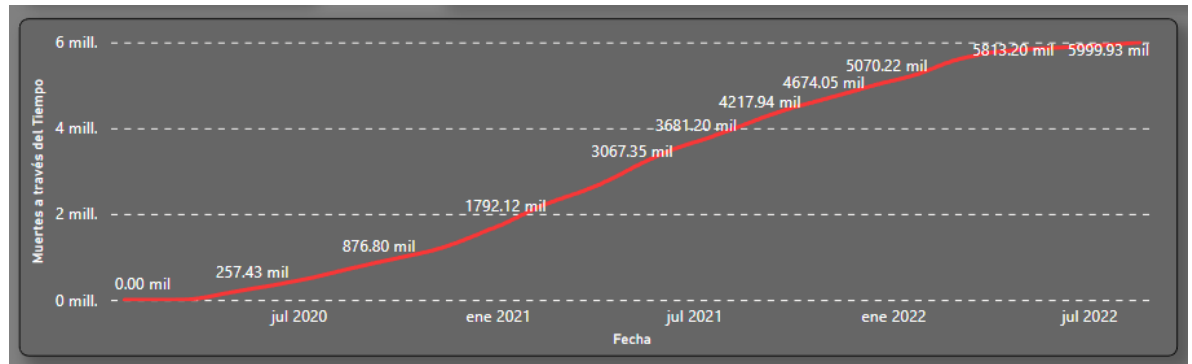
Análisis de como el porcentaje de Vacunaciones ha cambiado durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).



g) Serie de Tiempo – Fallecidos

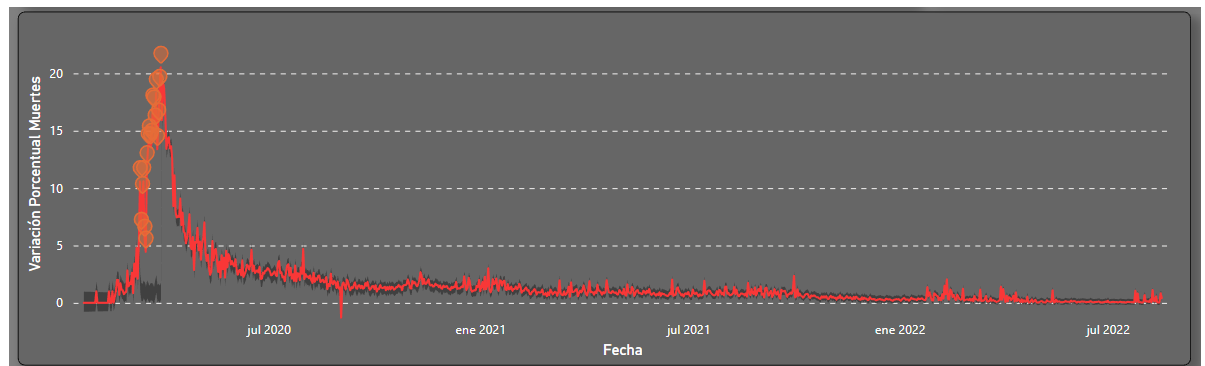
Identificar la cantidad de fallecidos causados por el virus durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).





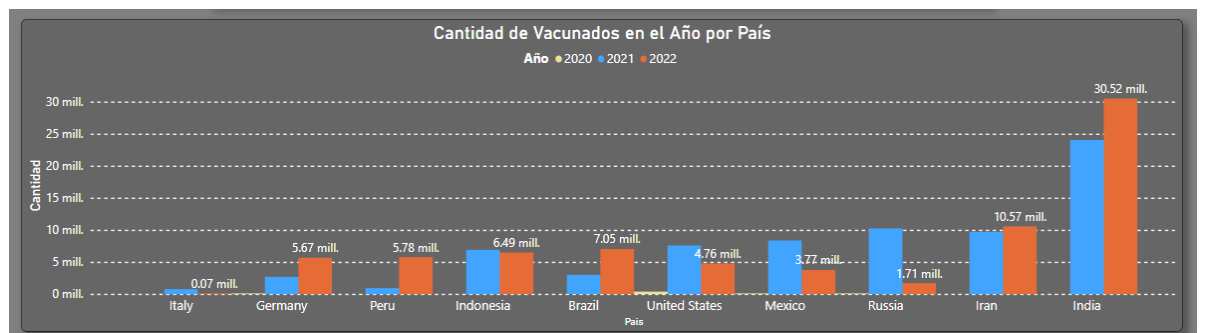
## h) Serie de Tiempo – Variación Porcentual de Fallecidos

Análisis de como el porcentaje de Fallecidos ha cambiado durante el período de tiempo desde el comienzo del Covid-19 a la fecha del día (2021-2022).



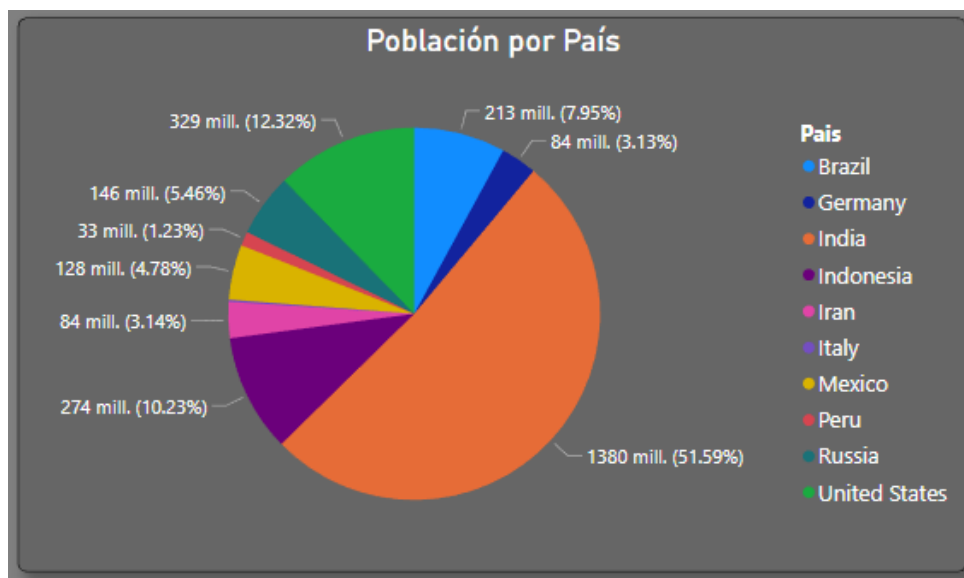
## i) Top 10 País Fallecidos - Cantidad de Vacunados en el Año por País

Este gráfico de columnas agrupadas nos permite saber la cantidad de vacunados a través de los años 2020-2021-222



j) Top 10 - Población por País

Gráfico circular para mostrar la comparativa de la población de cada país con sus respectivos porcentajes.



k) Top 10 - Acumulado de Vacunaciones por País y Año

Treemap para saber el acumulado por año la cantidad de vacunados en cada país, representando desde el país con mayor cantidad de vacunados a la fecha.

