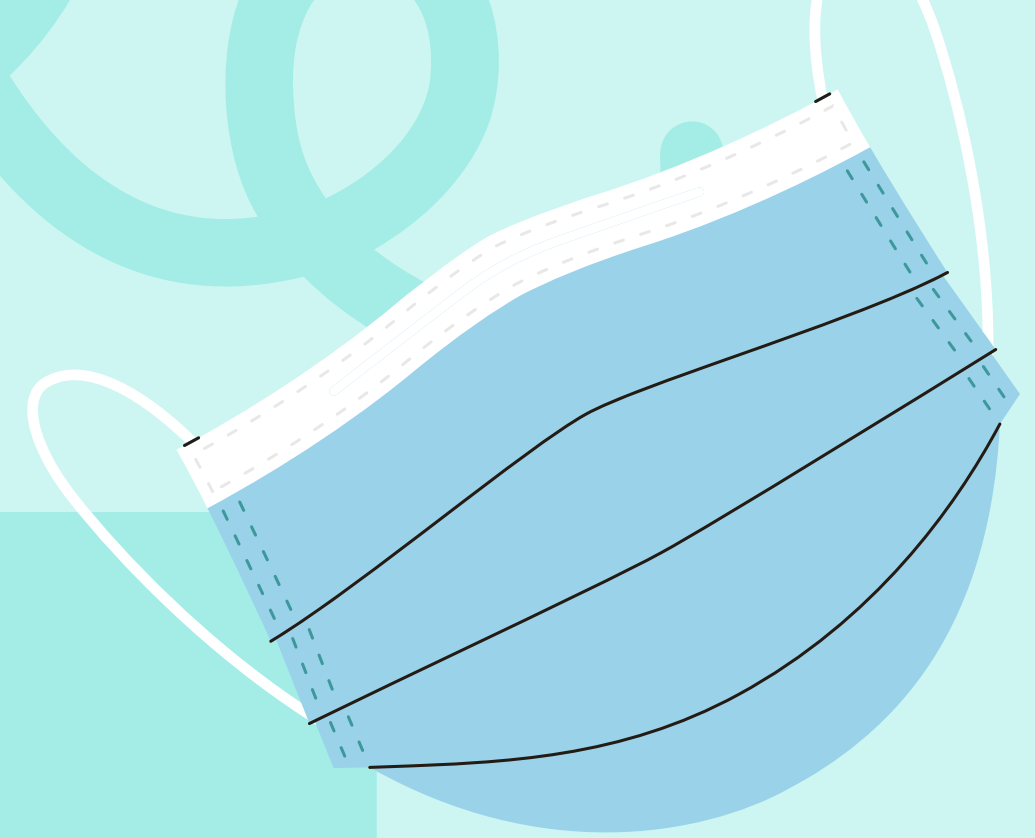
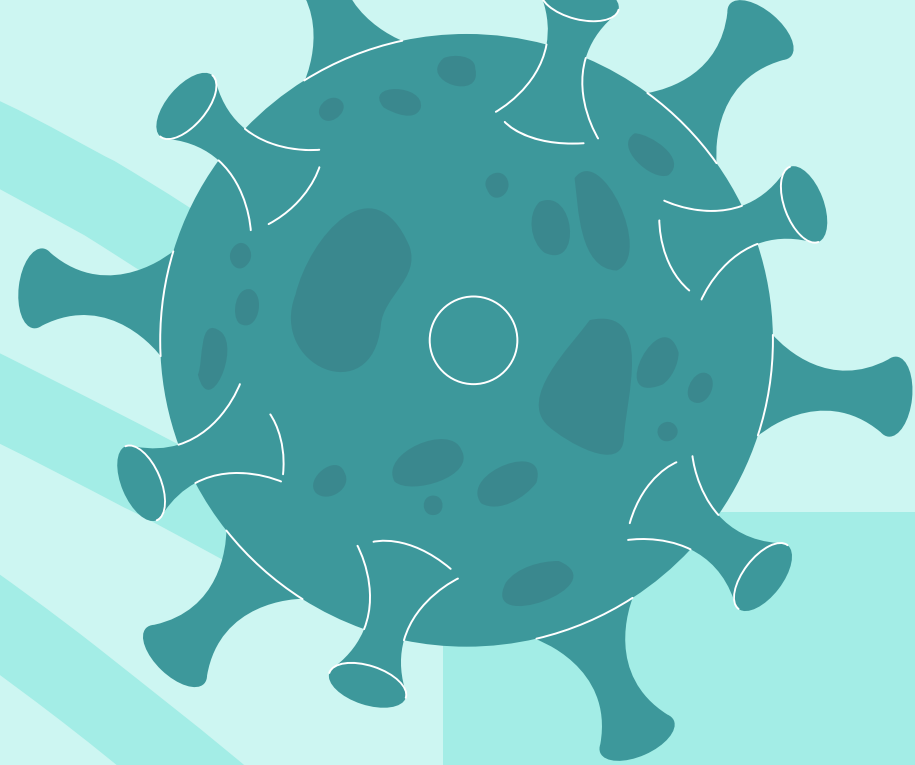


GRUPO 6

TFM | COVID-19

Datos y recomendaciones generales sobre los resultados de la investigación.

Fabián Ascheri Aguerre
José Chavarría Montero
Amaia Miranda Ulloa
Patricia Peña Torres
Juan Carlos Valcuende Aláez



EL EQUIPO



Fabián Ascheri



José Chavarría



Amaia M. Ulloa



Patricia A. Peña



**Juan Carlos
Valcuende**



INTRODUCCIÓN

1. Motivación

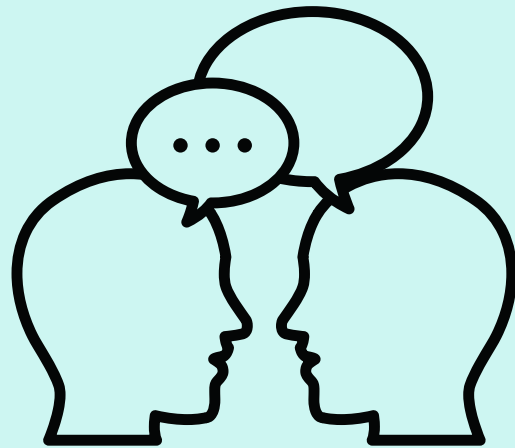
- Comprender y abordar el impacto global del COVID-19.

2. Finalidad

- Transformar datos de la pandemia en conocimiento valioso para nuestros usuarios.



OBJETIVOS PRINCIPALES



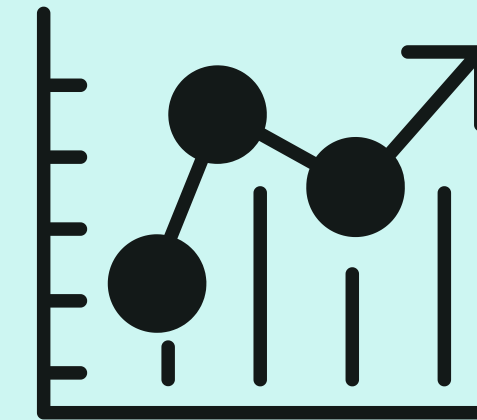
Explorar las necesidades

1. Mapa de empatía, post-motorola, team alignment map
2. Datos OMS, WDI y otras fuentes oficiales
3. Metodologías AGILE + SCRUM



Extracción, transformación y carga

1. Datos OMS, WDI y otras fuentes oficiales
2. Data wrangling
3. Base de datos MySQL
4. Deployment (Azure)



Visualizar y analizar

1. Realización de dashboards en PowerBI
2. Enriquecimiento de datos
3. Extracción de insights

NECESIDADES

El usuario/cliente serían profesionales de la salud, hablamos con ellos para conocer sus intereses y preocupaciones.



¿COVID persistente?
¿Vacunación? ¿Factores que
inciden?

Fue necesario adaptarnos a
nuestros recursos y datos
disponibles pero intentando seguir
dentro de los intereses del target

FOCOS DE INTERÉS

Influencia de la vacunación

Cantidad de dosis recibidas, % de población vacunada, tendencia relativas casos y muertes

Indicadores económicos sociodemográficos y

PIB, población urbana/rural, acceso a internet, acceso a servicios básicos, saneamiento y densidad poblacional

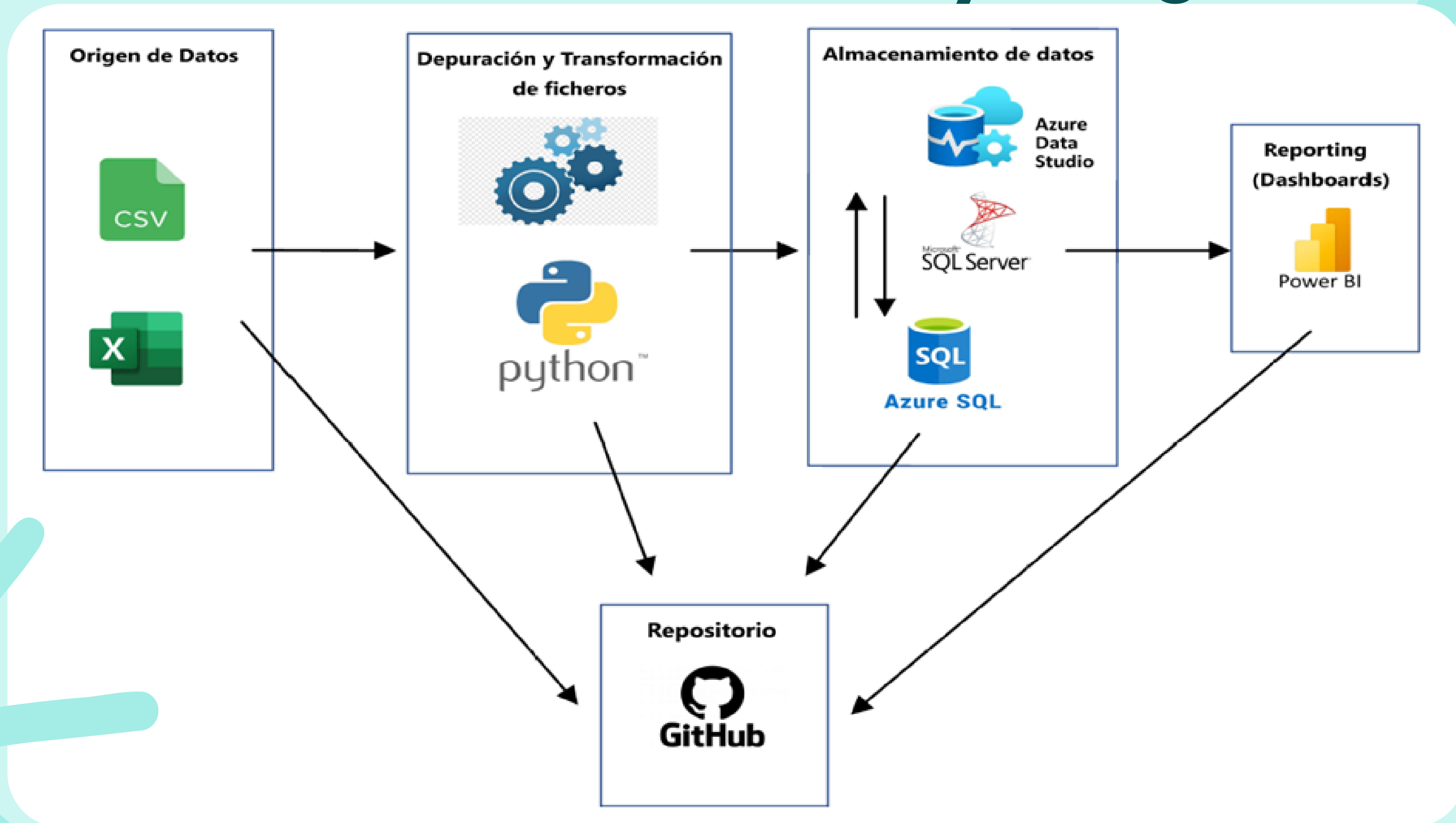
Factores ambientales y geográficos

Temperatura y altitud sobre el nivel del mar



ETL

(Extracción, transformación y carga de datos)



DATA WRANGLING

Extracción de datos

Centro Europeo para la Prevención y el Control de Enfermedades, Organización Mundial de la Salud, Universidad John Hopkins

Análisis exploratorio

Visualización con matplotlib, folium, plotly, seaborn + Resúmenes estadísticos + test de hipótesis + heatmaps clusterizados

Eliminación de duplicados

Tras verificar que efectivamente los valores fueran duplicados se procedía a su eliminación

Tratamiento de atípicos

Se corregirán datos que iban más allá de lo virtualmente posible tras investigar cada caso.



Imputación de valores nulos

Se abordarán los valores nulos mediante análisis y agrupación (group by) para determinar su origen y patrón. Dependiendo del caso, se completarán, eliminarán registros o columnas según sea necesario.

Ingeniería de atributos

Se crearán nuevas características o variables derivadas de las existentes, por ejemplo ratios por 100k habitantes

Normalización de formato

Se reorganizarán las tablas y se renombrarán para conseguir una estructura consistente

Exportación

Finalmente se exportarán los dataframes.

CARGA DE DATOS

Base de datos local

Creada inicialmente con MySQL

Almacenamiento en la nube

Se buscará hacer una migración a una BBDD 100% virtual en Azure

Creación de servidor

Se creará `servidortfmcovid-v2`, se configurarán las reglas del firewall para garantizar el acceso desde distintas direcciones IP

Base de datos SQL Server

Se sobrestimaré el almacenamiento necesario puesto que acabarán siendo menos de 40 megabytes



Administración de coste

Fue necesario asesorarse con el profesor Juan Luis Bermúdez para reducir el coste y dejase de incrementar descontroladamente.

Azure Data Studio

Se utilizará ADS para que interactuase con la BBDD MySQL

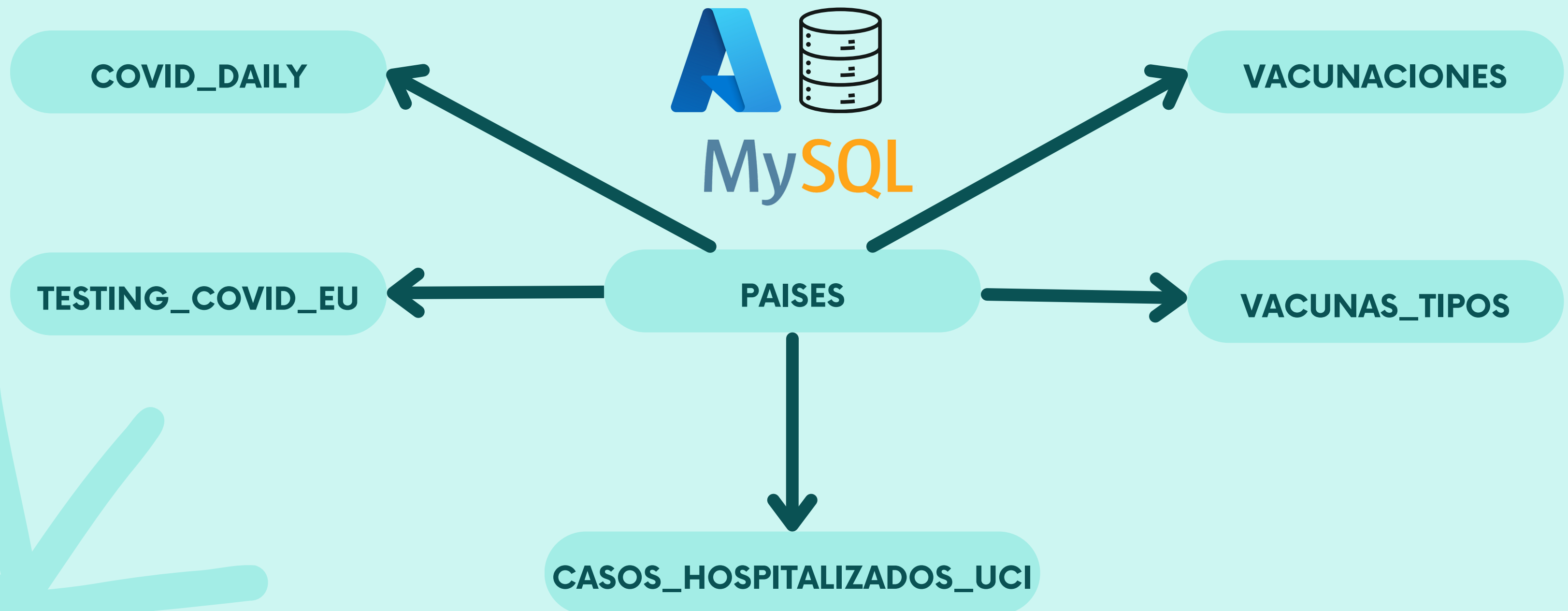
Semi-automatización

Dentro del servidor están alojados los notebooks así que de forma semiautomática se pueden mantener la base de datos, y consecuentemente las visualizaciones, actualizadas

Conexión con PowerBI

Se integrará PowerBI con Azure de forma que obtendrá los datos del servidor

ESTRUCTURA DE LA BBDD



VISUALIZACIÓN

Dashboards

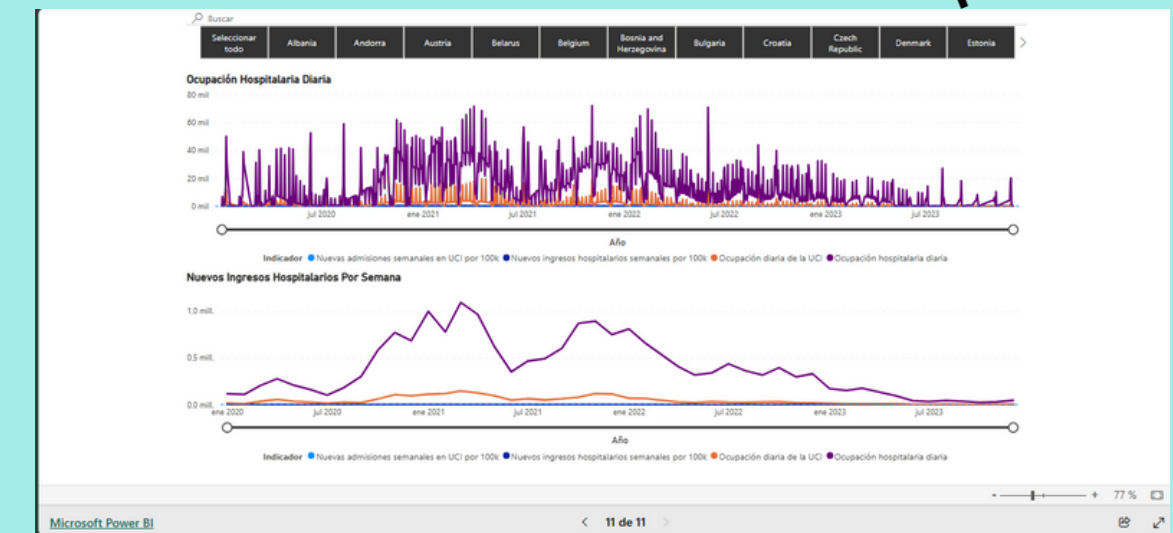
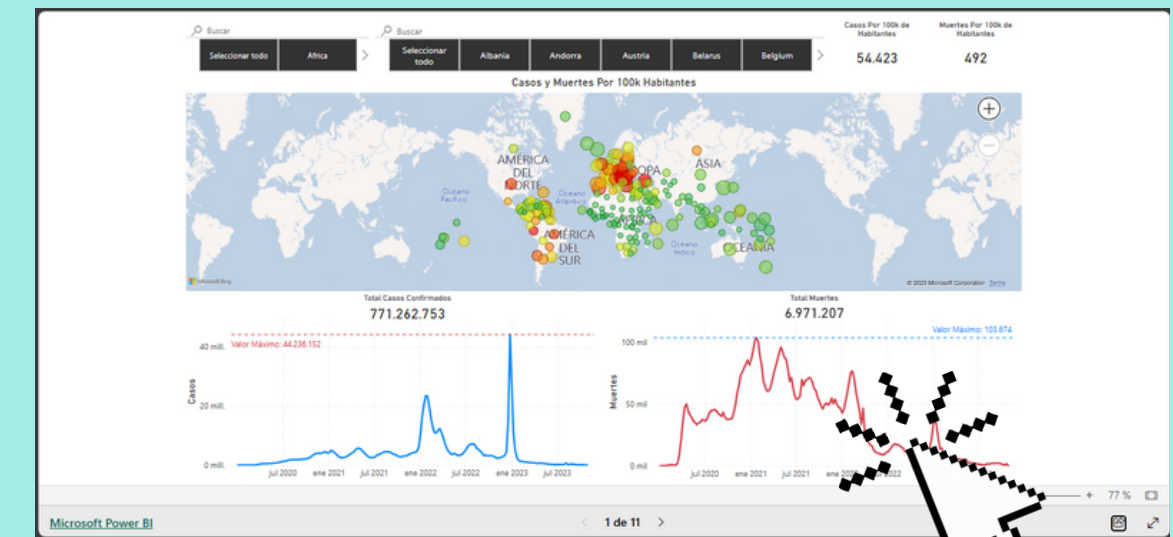
Elaborados en PowerBI

Enriquecidos

Integración de indicadores sociodemográficos, económicos y geográficos

Intención

Se buscará llegar una comprensión global y holística de la pandemia por COVID-19





CONCLUSIONES

Nuestra investigación fue dinámica y adaptable, permitiéndonos ajustar estrategias a medida que avanzábamos. Esto nos llevó a importantes conclusiones, como la falta de correlación directa entre casos y muertes por COVID-19, la importancia crítica de la vacunación, el impacto de la densidad urbana, y la relación con el acceso a servicios e internet.

Además, las condiciones ambientales, como la temperatura y la altitud, influyen en la propagación y mortalidad del virus, lo que sugiere la necesidad de investigaciones adicionales. Este enfoque flexible no solo optimizó nuestro proyecto sino también contribuyó a la comprensión de la pandemia de COVID-19 en un contexto global.



GRACIAS

