**Expansión estratégica Biogenesys.**

**Nombre del autor: Edwin santiago**

**Email: edwinsantiago4224@gmail.com**

**Cohorte: DA-FT13**

**Fecha de entrega:** 10/05/2025

**Institución:**

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# Introducción

En el presente documento se detallan los logros alcanzados en el marco del proyecto de análisis de datos relacionados con la pandemia de COVID-19. Se recopilaron y analizaron datos de diversos países con el objetivo de identificar oportunidades de expansión y desarrollo para la empresa. Este análisis busca proporcionar información estratégica que facilite la toma de decisiones respecto a posibles regiones donde implementar nuevos centros de salud, contribuyendo así a la contención de la pandemia.

El enfoque del proyecto se basa en un equilibrio entre la rentabilidad y el impacto social, procurando que las acciones de expansión sean económicamente viables y, al mismo tiempo, logren extender los servicios de la empresa a la mayor cantidad de personas posible.

# 

# 

# Desarrollo del proyecto

# El proyecto fue desarrollado en cuatro etapas, denominadas "avances", cada una con objetivos y actividades específicas. A continuación, se detalla lo realizado en cada una de ellas:

**Avance 01**

En esta primera etapa, se inició el tratamiento de los datos. Se utilizó un archivo en formato CSV, el cual fue cargado en *Visual Studio Code* mediante el lenguaje de programación Python. Para el procesamiento de los datos, se instalaron y emplearon librerías especializadas como **pandas** y **numpy**.

Las tareas realizadas incluyeron:

* Limpieza y depuración del conjunto de datos, eliminando columnas y filas con más del 90% de valores nulos.
* Corrección de valores negativos en columnas donde estos no eran válidos.
* Cálculo de estadísticas descriptivas como la **mediana** y la **varianza**, con el fin de comprender la distribución de los datos.

Además, se comenzaron a generar métricas clave para el análisis, tales como:

* Cálculo de casos activos.
* Proporción de casos por cada 100 mil habitantes.
* Tasa de recuperación expresada en porcentaje.

Estas métricas sentaron las bases para realizar un análisis más profundo en las siguientes etapas del proyecto.

| **País** | **Tasa Recuperación (%)** | **Tasa Incidencia Casos Nuevos (por 100k)** | **Tasa Mortalidad por 100k** | **Proporción Hombres/Mujeres** | **Casos Activos** | **Tasa Mortalidad Diaria (por 100k)** | **Relación Nuevos Recuperados / Nuevos Confirmados** | **Tasa de Confirmación (%)** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Argentina | 49.53 | 35.34 | 0.45 | 0.96 | 16224.49 | 0.45 | 196.22 | 0.04 |
| Brazil | 49.30 | 70.78 | 1.14 | 1.04 | 15520.14 | 1.14 | 154.57 | 0.07 |
| Chile | 60.34 | 36.60 | 0.77 | 0.99 | 9722.76 | 0.77 | 233.58 | 0.04 |
| Colombia | 54.34 | 56.02 | 1.09 | 1.00 | 10947.55 | 1.09 | 187.07 | 0.06 |
| México | 72.46 | 147.95 | 3.23 | 0.89 | 5940.04 | 3.23 | 262.39 | 0.15 |
| Perú | 83.41 | 226.34 | 2.52 | 0.94 | 2026.18 | 2.52 | 172.48 | 0.23 |

**Avance 02**

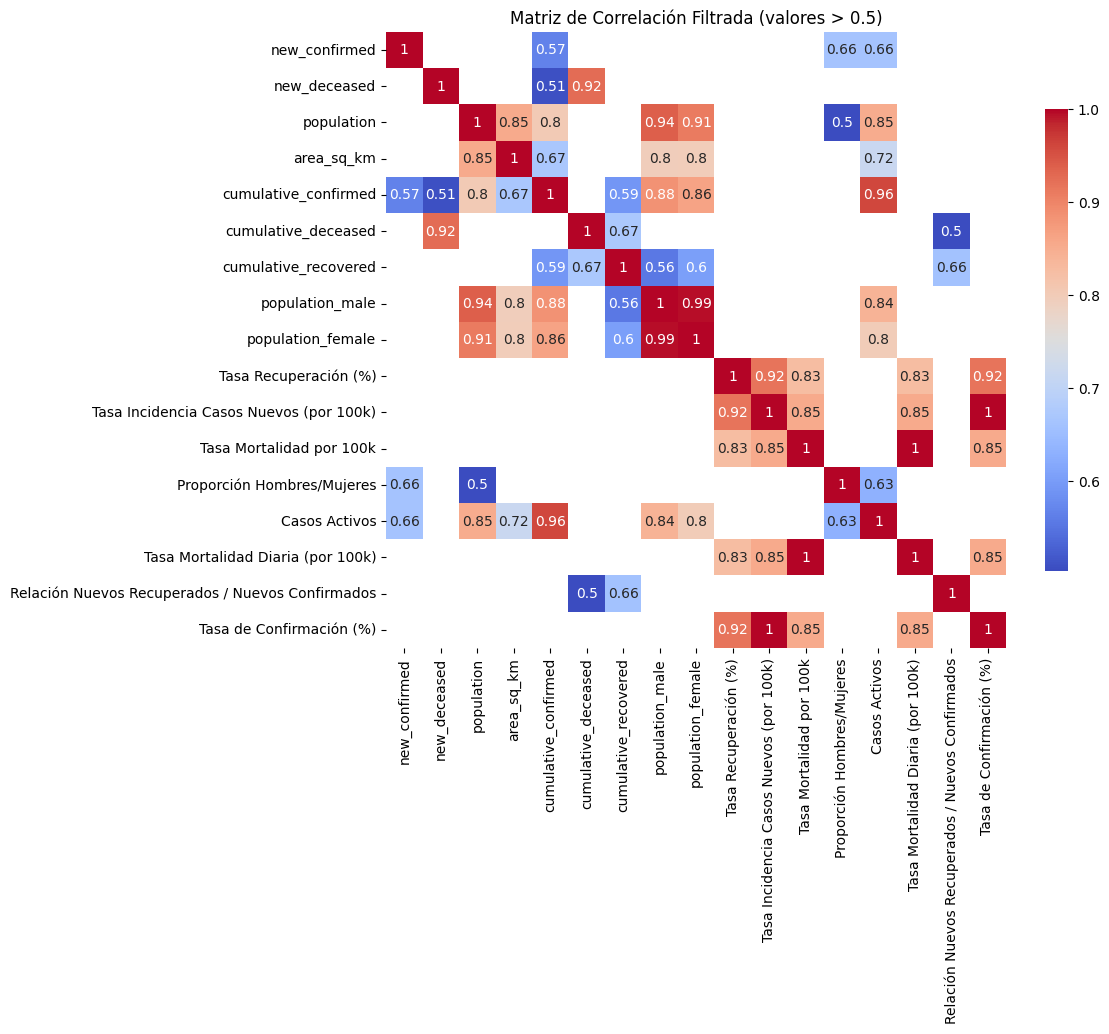
En la segunda etapa del desarrollo se realizó un análisis exploratorio de datos utilizando las librerías pandas y numpy, las cuales permiten efectuar cálculos aritméticos y estadísticos sobre grandes volúmenes de información.

El objetivo principal de esta fase fue comprender la distribución de los datos y analizar las correlaciones entre las distintas variables. Para facilitar la interpretación visual y obtener *insights* más profundos, se emplearon herramientas de visualización como matplotlib y seaborn.

Entre las acciones destacadas de este avance se incluyen:

* Análisis de correlación entre variables clave del conjunto de datos.
* Generación de gráficos para observar la evolución de variables como casos recuperados y fallecimientos, desagregados por país.
* Identificación de relaciones significativas entre las métricas propias calculadas en la etapa anterior.

Estos análisis permitieron visualizar cómo se comportan las variables entre sí y qué implicancias podrían tener en relación con los pacientes afectados. Asimismo, los gráficos por país ofrecieron una visión más clara sobre la distribución de recursos médicos y servicios sanitarios en cada región.



# Avance 03

# En esta etapa del proyecto nos enfocamos en realizar un análisis temporal con el objetivo de entender cómo evolucionaron los principales indicadores a lo largo del tiempo. Utilizamos la columna "date" para descomponer la información en componentes temporales: días, meses y años. Esta transformación nos permitió analizar las tendencias y variaciones en diferentes períodos.

# A partir de nuestras métricas previamente calculadas, llevamos a cabo diversos análisis como:

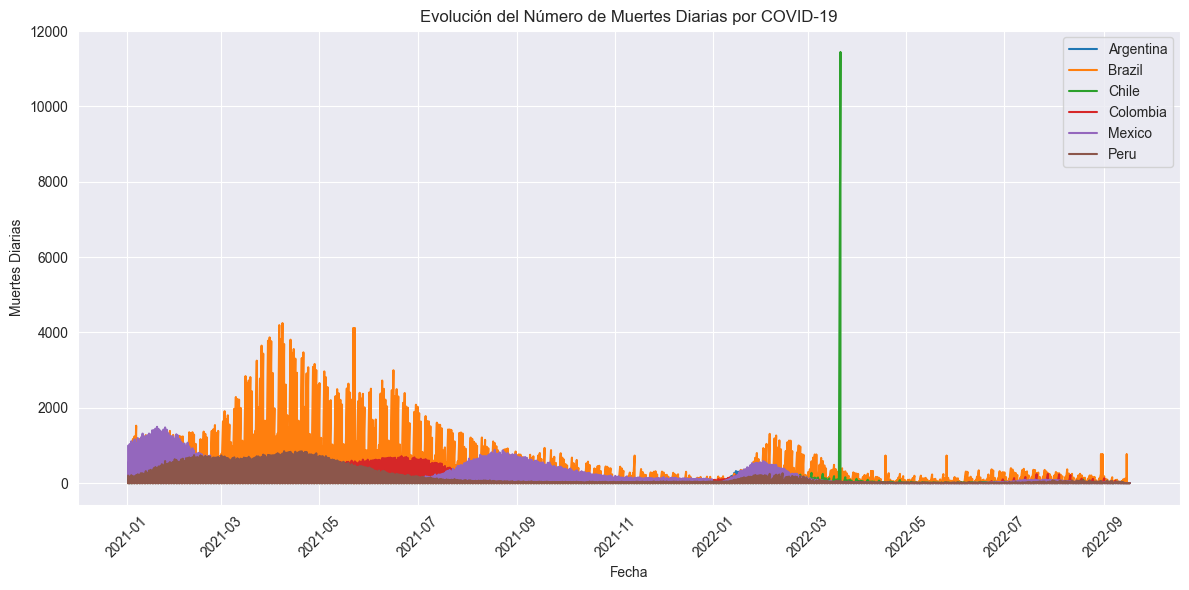
# Evolución de los casos confirmados y casos recuperados a lo largo del tiempo.

# Comportamiento de los casos activos en distintos intervalos temporales.

# Distribución de los casos confirmados en relación con la población en diferentes momentos.

# Este análisis temporal reveló patrones relevantes y permitió identificar momentos críticos durante la pandemia, así como también observar la dinámica de recuperación en los distintos países.

# 



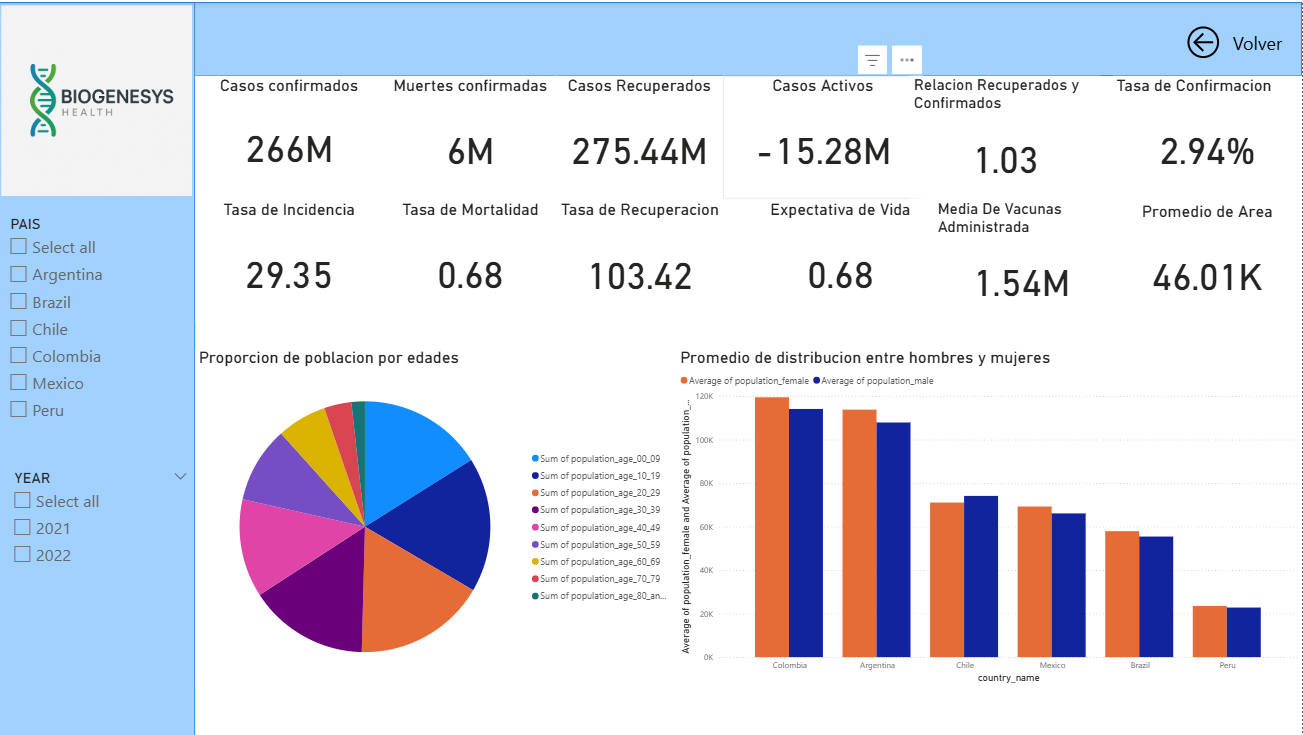
**Avance 04**

En esta última etapa del proyecto se llevó a cabo la creación del dashboard en Power BI. Tras finalizar la limpieza y preparación de los datos en Python, se exportaron en un archivo CSV, el cual fue posteriormente importado a Power BI para su visualización y análisis interactivo.

Dentro de Power BI se desarrollaron diversas métricas, indicadores y visualizaciones que permitieron representar de forma clara y dinámica la información más relevante. Entre los elementos incorporados se incluyen:

* **Gráficos dinámicos** para visualizar métricas clave por país y por período.
* **Slicers** para facilitar la segmentación por variables como país, año o estado de los pacientes.
* **Indicadores personalizados** para destacar datos críticos como la tasa de recuperación o la proporción de casos activos.

Adicionalmente, se integraron al dashboard algunos gráficos generados previamente en Python, utilizando scripts embebidos. Esto permitió incluir visualizaciones más avanzadas que no podían ser replicadas directamente en Power BI.



# Análisis general

# En términos generales, se observaron ciertas inconsistencias y sesgos en los datos. Por ejemplo, en algunos casos se registró una cantidad de casos recuperados superior al número de casos confirmados, lo cual sugiere posibles errores en la recolección o procesamiento de la información.

# Algunos hallazgos destacados del análisis por país fueron los siguientes:

# Argentina y Chile presentaron una de las mayores cantidades de casos activos, sin embargo, también fueron los países con menor número de muertes registradas. En el caso de Argentina, se observó una baja distribución de dosis de vacunas, lo cual podría estar asociado a una menor disponibilidad de personal sanitario o a una infraestructura limitada para la aplicación de vacunas.

# Brasil fue el país con mayor número de muertes y también con la mayor cantidad de dosis administradas. Esta situación podría reflejar una presión significativa sobre el sistema de salud o un posible declive temporal en la expectativa de vida de la población. Además, se identificó un valor atípico: el número de casos recuperados supera ampliamente a los casos confirmados, lo cual indica una probable inconsistencia en los registros.

# Perú mostró el mejor índice de recuperación, incluso superando el número de casos confirmados, lo cual también podría estar vinculado a errores en los datos. Sin embargo, fue uno de los países con menor cantidad de casos confirmados y una alta tasa de vacunación, lo que podría haber contribuido a contener la propagación del virus.

# México registró una alta cantidad de casos confirmados, pero también demostró un buen control de la situación sanitaria, con una tasa de recuperación del 93%. Esta recuperación eficiente podría estar relacionada con la elevada cantidad de dosis administradas en el país.

# Estos hallazgos proporcionan una visión general del impacto de la pandemia en los países analizados, así como posibles áreas de oportunidad para la expansión de servicios sanitarios y análisis de mercado en el sector salud.

# 

**Conclusión.**

A partir del análisis de los datos, es posible realizar una evaluación de la situación sanitaria en los países estudiados, lo que permite definir áreas de intervención clave para la expansión de los servicios de la empresa. Los factores clave a considerar para la toma de decisiones son los siguientes:

**1. Distribución de Casos y Recursos Sanitarios**

En Argentina y Chile, aunque los casos activos son elevados, la baja tasa de mortalidad y la baja distribución de dosis de vacunas sugieren que la intervención podría centrarse en mejorar el acceso a la vacunación y en la ampliación de la infraestructura sanitaria.

Brasil, a pesar de tener un alto número de muertes y una gran cantidad de dosis administradas, presenta un posible error de datos, lo cual podría distorsionar la verdadera magnitud de la crisis. Sin embargo, el país tiene una alta demanda de servicios sanitarios, lo que lo convierte en una opción clave para la expansión, especialmente en áreas de alta densidad poblacional.

Perú, con su alto índice de recuperación y buena distribución de dosis, muestra un modelo de respuesta eficaz. A pesar de tener menos casos confirmados, la alta tasa de vacunación sugiere que el país está en una fase más avanzada de control de la pandemia. Esta información puede ser útil para una posible expansión centrada en la atención post-pandemia.

México ha mostrado un control adecuado de la situación, con una tasa de recuperación destacada y una alta cantidad de casos confirmados. Esto indica que, aunque el país enfrenta una carga considerable, hay un buen control de los recursos y una infraestructura sólida que podría soportar una expansión de servicios sanitarios.

**2. Condiciones Socioeconómicas y Logísticas**

La capacidad de distribución de recursos como las vacunas y la presencia de personal médico son factores fundamentales. Países con menor distribución de dosis, como Argentina, podrían requerir mayores inversiones en infraestructura y personal.

La disponibilidad de datos fiables es crucial para asegurar que las intervenciones se basen en información precisa. En este sentido, es importante tener en cuenta los posibles errores de los datos en **Brasil** al tomar decisiones sobre la expansión.

Con base en los hallazgos y la evaluación anterior, las siguientes recomendaciones pueden ser clave para las decisiones estratégicas de expansión de la empresa:

**1. Expansión en Brasil**

A pesar de los posibles errores de los datos, Brasil representa una gran oportunidad debido a la alta demanda de servicios sanitarios y la gran cantidad de muertes. Se recomienda iniciar un análisis más detallado de la situación sanitaria en las regiones con mayor densidad poblacional y evaluar la posibilidad de establecer centros de salud en esas áreas.

**2. Refuerzo de infraestructura en Argentina y Chile**

Argentina y Chile tienen una alta cantidad de casos activos pero una baja tasa de mortalidad. Estos países podrían beneficiarse de un refuerzo en la distribución de vacunas y la expansión de infraestructura sanitaria. Se recomienda explorar asociaciones con organismos gubernamentales y privados para mejorar el acceso a las vacunas y a servicios médicos.

**3. Modelo de atención post-pandemia en Perú**

Perú, con su alta tasa de recuperación, podría ser un buen lugar para ofrecer servicios centrados en la recuperación post-COVID-19. Dado su éxito con la vacunación, el enfoque aquí podría ser el establecimiento de centros de salud especializados en la atención post-pandemia, en línea con las necesidades de la población.

**4. Fortalecimiento de la infraestructura en México**

México muestra un buen control de la pandemia, pero la alta cantidad de casos confirmados sugiere que aún hay trabajo por hacer. Se recomienda fortalecer la infraestructura en áreas con alta incidencia de casos confirmados y priorizar la distribución equitativa de dosis y atención médica en estas regiones.

**Reflexión personal**

A través de este proyecto de análisis de datos sobre la pandemia de COVID-19, he podido comprender con mayor profundidad no solo la complejidad de los datos sanitarios, sino también las enormes oportunidades y desafíos que enfrentan las empresas en el sector salud. Este análisis nos ha mostrado que la información, cuando es procesada correctamente, puede servir como una poderosa herramienta para tomar decisiones informadas que impactan directamente en la vida de las personas.

Lo más revelador de este ejercicio ha sido observar cómo, a pesar de la magnitud de la crisis sanitaria global, existen patrones comunes en varios países, pero también realidades muy particulares que deben ser consideradas en cualquier estrategia de expansión o intervención. He podido apreciar cómo la diferencia en la distribución de recursos, la infraestructura sanitaria y la efectividad en la vacunación tienen un impacto significativo en la evolución de la pandemia.

A nivel personal, este proyecto refuerza mi creencia de que el análisis de datos va más allá de simples números y fórmulas; se trata de entender y responder a las necesidades de las personas. Las decisiones que tomemos hoy en base a estos análisis no solo influyen en el futuro de las empresas, sino también en la salud y el bienestar de millones de personas.

Por último, he aprendido que, en un contexto tan incierto como el de la pandemia, la adaptabilidad es clave. Los datos pueden cambiar, las circunstancias pueden evolucionar, pero lo que siempre debe permanecer constante es nuestro compromiso con el bienestar social y la mejora continua.

# 