**EXPANSION ESTRATÉGICA DE BIOGENESYS**

**Nombre del autor: Fabian Arias**

**Email:** fabiancarias0111@gmail.com

**Cohorte:** DA-FT11

**Fecha de entrega:** 24/02/25

**Institución:** Biogenesys Institute

# Introducción

El propósito de este proyecto es determinar los lugares ideales para el crecimiento de los laboratorios farmacéuticos Biogenesys, basándose en el estudio de datos vinculados a la pandemia del COVID-19 en seis países de América Latina: Colombia, Argentina, Brasil, Chile y Perú. Este análisis se enfoca en el tiempo que va desde enero de 2021 hasta septiembre de 2022. Los objetivos logrados por la organización comprenden la creación de estrategias fundamentadas en datos que faciliten una toma de decisiones basada en evidencia, potenciando el impacto de la compañía en mercados críticos.

# Desarrollo del proyecto

La metodología empleada para este proyecto incluyó las siguientes etapas:

1. **Recopilación y selección de datos:** Se utilizaron bases de datos públicas sobre casos confirmados, muertes, vacunación, y variables demográficas y climáticas de los seis países analizados.
2. **Transformaciones y limpieza de datos:**
   * Se eliminaron valores duplicados e inconsistentes.
   * Se gestionaron valores nulos aplicando imputación estadística donde fue necesario.
   * Se crearon columnas derivadas para calcular indicadores clave, como tasas de fatalidad y porcentaje de vacunación.
3. **Graficacion:** Se utilizó la base de datos limpia para empezar a desarrollar gráficos que nos permitieran encontrar insights sobre la pandemia del COVID-19 en Latinoamérica.
4. **Dashboard:** Se realizo un informe detallado en Power BI con parte de los insights enocntrados en el tercer avance,

**Conclusión:** Este proceso aseguró que los datos estuvieran preparados para un análisis exploratorio robusto y la generación de insights significativos.

# EDA e insights

En el Análisis Exploratorio de Datos (EDA) se obtuvieron los siguientes insights clave:

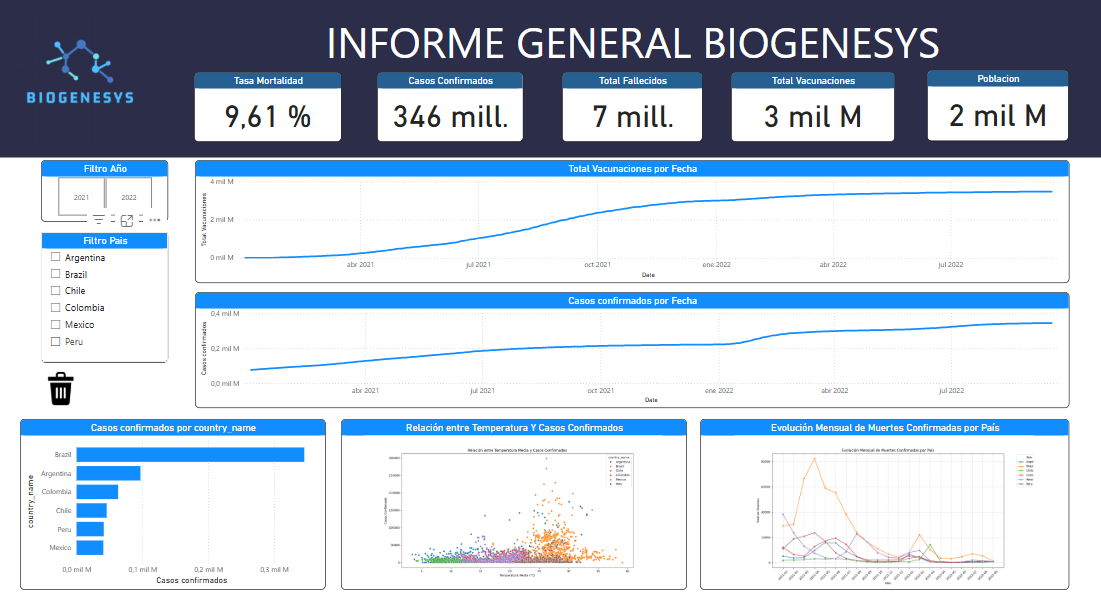
1. Chile es el país con el mayor porcentaje de población anciana.
2. Brasil, debido a su alta población, fue también el país con el mayor número de casos confirmados de COVID-19.
3. El pico de muertes ocurrió entre marzo y mayo de 2021, con Brasil liderando las cifras.
4. Los hombres fueron más propensos a morir por COVID-19 que las mujeres.
5. A partir de marzo de 2021, el número de dosis de vacunas administradas comenzó a incrementarse de manera significativa en todos los países.
6. Entre enero y marzo de 2022 se registró la mayor cantidad de casos confirmados.
7. Las temperaturas promedio entre 25 y 30 grados Celsius estuvieron asociadas al mayor número de contagios.
8. Entre noviembre de 2021 y enero de 2022, los casos activos superaron a los casos recuperados.
9. A pesar de un bajo desempeño en vacunación, Colombia tuvo un impacto proporcionalmente menor en la propagación del virus debido a su densidad poblacional.
10. Se observó una tasa de fatalidad del 2% al 10% entre los seis países analizados.
11. Se observó un aumento critico en los casos confirmados a inicio de año debido a la temporada de vacaciones

# Análisis del dashboard

El dashboard desarrollado en Power BI permite una navegación interactiva por las siguientes visualizaciones clave:

1. **Casos confirmados:** Muestra el total de casos confirmados.
2. **Total Fallecidos:** Muestra total de fallecidos acumulados.
3. **Total Vacunaciones:** Muestra el total de Vacunaciones acumuladas.
4. **Población:** Muestra el Total de población.
5. **Tasa Mortalidad:** Muestra la Tasa de Mortalidad por casos confirmados
6. **Filtro por año:** Una segmentación para filtrar los datos por el año.
7. **Filtro por País:** Segmentación para filtrar datos por País.
8. **Total Vacunas por Fecha:** Esta visualización muestra el acumulado de las vacunas administradas a lo largo del tiempo.
9. **Casos Confirmados por Fecha:** Esta Grafica muestra el acumulado de los casos confirmados por fecha.
10. **Casos Confirmados por País:** Esta Grafica muestra el total de los casos confirmados por País.
11. **Relación entre temperatura media y casos confirmados:** Se observa que las temperaturas promedio entre 25 y 30 grados Celsius estuvieron asociadas al mayor número de contagios, lo que sugiere una posible relación climática con la propagación del virus.
12. **Evolución Mensual de Muertes Confirmadas por País:** Esta visualización Muestra el total de muertes por país a lo largo del tiempo.





Conclusiones extraídas del dashboard:

* La distribución etaria y la densidad poblacional son factores clave para comprender el impacto del COVID-19 en cada país.
* La evolución de las políticas de vacunación fue heterogénea, con un impacto directo en la propagación del virus.

# Conclusiones y Recomendaciones

**Conclusión general:** Los resultados del análisis muestran cómo la dinámica del COVID-19 estuvo influenciada por factores demográficos, climáticos y políticos. Los países con mayor población y menor acceso a vacunas fueron los más afectados.

**Recomendaciones estratégicas:**

* **Brasil:** Debido a su alta población y alto número de casos y muertes, representa un mercado con gran necesidad de soluciones farmacéuticas.
* **México:** Con un sistema de salud que enfrenta retos significativos en cobertura y vacunación, es una oportunidad clave.
* **Colombia:** A pesar de su baja densidad poblacional, el bajo porcentaje de vacunación y sus retos sociales lo convierten en un país estratégico.

# Reflexión personal

Este proyecto me permitió desarrollar habilidades clave en limpieza de datos en phyton y sus diferentes librerías, visualización y creación de dashboards interactivos. Si tuviera que reiniciar el proyecto, me enfocaría desde el principio en optimizar el proceso de transformación de datos y buscaría explorar otras fuentes de información para enriquecer el análisis. Este aprendizaje ha sido invaluable para mi crecimiento como analista de datos y para entender cómo traducir insights en estrategias de impacto.

Analista de Datos

**Fabian Arias**