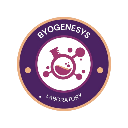
**EXPANSION ESTRATÉGICA DE BIOGENESYS**

**Nombre del autor:** Anderson Arias

**Email:** anderson10arias@gmail.com

**Cohorte:** DA-FT10

**Fecha de entrega:** 27/01/25

**Institución:** Biogenesys Institute 

# Introducción

# El presente proyecto tiene como objetivo identificar las ubicaciones óptimas para la expansión de los laboratorios farmacéuticos Biogenesys, a partir del análisis de datos relacionados con la pandemia del COVID-19 en seis países de América Latina: Argentina, Brasil, Chile, Colombia, México y Perú. Este estudio se centra en el período comprendido entre enero de 2021 y septiembre de 2022. Los objetivos organizacionales alcanzados incluyen el desarrollo de estrategias basadas en datos que permitan una toma de decisiones fundamentada, maximizando el impacto de la empresa en mercados clave.

# 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 graficos que nos permitieran encontrar insights sobre la pandemia del covid en latinoamerica.
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:

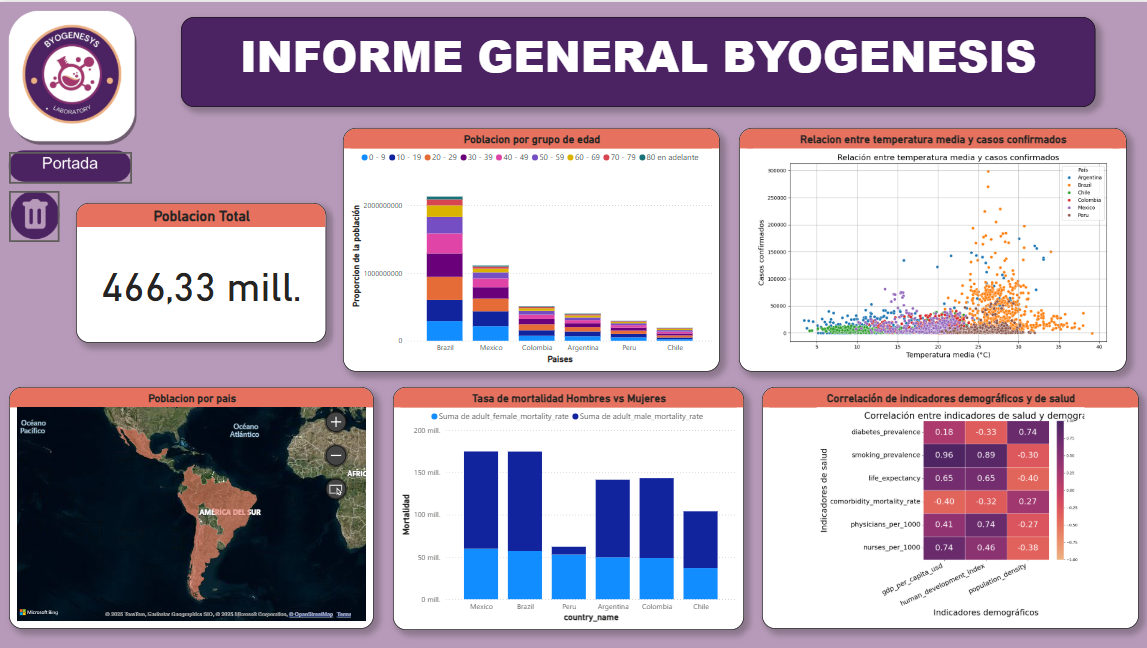
1. Chile es el país con el mayor porcentaje de población anciana.
2. Brasil, debido a su alta población, fue también el país con el mayor número de casos confirmados de COVID-19.
3. El pico de muertes ocurrió entre marzo y mayo de 2021, con Brasil liderando las cifras.
4. Los hombres fueron más propensos 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. Las personas con diabetes y prevalencia de tabaquismo fueron más propensas a fallecer por COVID-19.
9. Entre noviembre de 2021 y enero de 2022, los casos activos superaron a los casos recuperados.
10. A pesar de un bajo desempeño en vacunación, Colombia tuvo un impacto proporcionalmente menor en la propagación del virus debido a su densidad poblacional.
11. Se observó una tasa de fatalidad del 2% al 10% entre los seis países analizados.

# Análisis del dashboard

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

1. **Población por grupo de edad:** Este gráfico clasifica a la población en rangos etarios (0 a 9, 10 a 19, y así sucesivamente hasta 80 años), permitiendo identificar patrones demográficos. Chile destaca por tener el mayor porcentaje de población anciana.
2. **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.
3. **Población total de Latinoamérica:** Esta visualización proporciona una perspectiva comparativa entre los países, destacando a Brasil como el más poblado.
4. **Tasa de mortalidad de hombres versus mujeres:** El análisis muestra que los hombres fueron más propensos a morir por COVID-19 que las mujeres, un patrón consistente en todos los países analizados.
5. **Correlación de indicadores demográficos y de salud:** Este gráfico explora cómo variables como tabaquismo, diabetes y densidad poblacional influyen en la propagación del virus y las tasas de mortalidad.





Conclusiones extraídas del dashboard:

* La distribución etaria y la densidad poblacional son factores clave para comprender el impacto del COVID-19 en cada país.
* Existe una correlación notable entre ciertos indicadores de salud (como prevalencia de diabetes y tabaquismo) y las tasas de fatalidad.
* 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 pyhton 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

**Anderson Arias Bermúdez**