**[Título del proyecto]**

**Nombre del autor:**

**Email:**

**Cohorte:**

**Fecha de entrega:** cuándo se completó y entregó el informe

**Institución: BIOGENESYS laboratory**

# Introducción

# La empresa farmacéutica BIOGENESYS busca identificar las ubicaciones óptimas para la expansión de laboratorios farmacéuticos, basándose en el análisis de datos de incidencia de COVID-19, tasas de vacunación, y la disponibilidad de infraestructuras sanitarias. La meta es optimizar la respuesta a los efectos de la pandemia y postpandemia con el fin de mejorar el acceso a las vacunas.

# Desarrollo del proyecto

primero se utilizo la libreria de pandas para leer el archivo csv proporcionado para el analisis,luego se filtro por los paises y las fechas pedidas. se limpiaron los datos faltantes y se normalizaron segun sus caracteristicas, algunos con la media y otros con datos mas dispersos con la mediana. se utilizaron funciones como fillna, dropna, ffill y bfill para tratar los nulos y rellenar los espacios. luego se utilizaron las librerias de matplotlib y seaborn para visualizaciones. y por ultimo se personalizaron los graficos para que sea mas estetico.

# EDA e insights

# Alcance de datos: 2021–2022 para AR, BR, CL, CO, MX, PE. Se saneó la serie (ffill/bfill y cummax() en acumulados), se obtuvieron tasas por 100k y métricas mensuales; el daily\_vax se derivó como diff() del acumulado.

# Tendencia temporal: picos marcados en 2021-Q1/Q2; onda corta en 2022-Q1 (ómicron) y desaceleración sostenida el resto del año. Los KPIs del último mes muestran caídas m/m fuertes en casos (~−81%) y muertes (~−72%) y leve alza en recuperados (~+21%).

# Carga por país (absoluta): Brasil domina en casos y muertes mensuales; México es el segundo gran contribuyente. Colombia y Perú muestran picos intermitentes; Argentina y Chile con volúmenes más contenidos.

# Correlaciones (heatmap):

# vax\_coverage\_per\_100 vs deaths\_per\_100k: negativa débil (≈ −0.15).

# vax\_coverage\_per\_100 vs cases\_per\_100k: negativa débil (≈ −0.13).

# cases\_per\_100k vs deaths\_per\_100k: positiva moderada (≈ +0.42). Lectura: mayor cobertura se asocia con menor severidad, y los casos anticipan muertes (coherente epidemiológicamente).

# 

# 

# Análisis del dashboard

# Slicers (izquierda):

# year para filtrar 2021/2022.

# location\_key para aislar uno o varios países. *Sugerencia de navegación*: empieza por “Select all” y luego quita/añade países para comparar.

# Fila superior:

# KPI de casos y KPI de muertes con variación m/m y mes de referencia.

# KPI de recuperados (útil como proxy de salida del evento).

# Heatmap de correlación (cobertura, casos, muertes por 100k) para una vista causal rápida.

# Barras anuales de casos por país (compara 2021 vs 2022).

# Centro–izquierda:

# Casos nuevos mensuales por país (barras apiladas por país y mes). Útil para identificar meses pico y país líder en carga absoluta.

# Inferior–izquierda:

# Muertes nuevas mensuales por país. Mismo análisis pero sobre severidad.

# Derecha:

# multiples (líneas) por país. Excelente para leer timing de olas y su magnitud relativa.

# Conclusiones típicas al navegar: al filtrar 2022, cae la incidencia y mortalidad en todos los países; al filtrar países de alta población (BR, MX) se concentran los volúmenes; Chile muestra buena contención relativa en muertes/100k con amplia cobertura.

# 

# Conclusiones y Recomendaciones

# Síntesis general. El ecosistema 2021–2022 muestra olas pronunciadas con descenso marcado en 2022. La cobertura de vacunación guarda relación inversa (aunque débil, como es esperable a nivel agregado) con muertes/100k; la relación casos→muertes es clara. En términos de oportunidad, combinando tamaño de mercado, carga por 100k, cobertura, urbanización/densidad y capacidad sanitaria (nurses/physicians por 1000) se identifican zonas donde nueva infraestructura (diagnóstico, fill-finish, control de calidad, hubs logísticos de biológicos) aportaría mayor valor.

Respuesta principal: ubicaciones óptimas para la expansión de laboratorios farmacéuticos

Prioridad alta

Brasil

Evidencia en el dashboard: Lidera los volúmenes absolutos en *Casos nuevos por año* (2021 y también alto en 2022) y en las barras mensuales de casos/muertes; en los small multiples se ven olas prolongadas y picos intensos.

Lectura operativa: Mercado grande + carga histórica alta ⇒ tiene sentido fill-finish, control de calidad y hubs de biológicos (capacidad para absorber escala).

México

Evidencia en el dashboard: Segundo en volumen anual y con barras mensuales relevantes durante 2021–2022; en líneas MM7 se aprecian picos claros y buena persistencia de demanda.

Lectura operativa: Buen candidato para laboratorios de diagnóstico (alto flujo) y centros de distribución (acopla con la carga observada en el panel).

## Prioridad Media

## Colombia

Evidencia en el dashboard: Picos intermitentes visibles en monthly bars y en el small multiple; carga menor que BR/MX pero no trivial.

Lectura operativa: Reforzar diagnóstico y cadena de frío para sostener cobertura (tu heatmap sugiere que más cobertura ↓ muertes/100k).

Perú

Evidencia en el dashboard: Picos marcados en 2021 y repunte corto en 2022-Q1; luego caída pronunciada (mensuales y MM7).

Lectura operativa: Hubs de distribución y laboratorios regionales para responder a demanda concentrada por olas y mejorar tiempo de respuesta.

## Prioridad Baja

Argentina

Evidencia en el dashboard: Volumen moderado en anual y mensual; en MM7 las olas están más contenidas y 2022 cae fuerte.

Chile

Evidencia en el dashboard: En líneas y barras muestra menor severidad relativa que sus pares; heatmap indica relación negativa entre cobertura y muertes/100k, coherente con su mejor control.

# Reflexión personal

# Aprendí mucho en este módulo sobre python, pandas, numpy, matplotlib y seaborn. Creo que es la manera más eficiente y rápida de archivos con mucho volumen de datos. Aparte de la facilidad para crear gráficas, correlaciones, estadística, medidas de tendencia central, etc.

# EXTRA CREDIT

# [Aquí detallarás las mejoras de tu proyecto]