**Expansión Estratégica de Biogenesys con Python**

**Nombre del autor: Parra Poma, Natalia Romina**

**Email: rominaparrapoma@gmail.com**

**Cohorte: DAFT-15**

**Fecha de entrega:** Lunes 7 de julio de 2025.

**Institución:**

****

# Introducción

# En el contexto de los desafíos sanitarios globales derivados de la pandemia de COVID-19, la empresa farmacéutica BIOGENESYS se ha propuesto fortalecer su presencia en Latinoamérica mediante una expansión estratégica de sus laboratorios y centros de vacunación. Esta iniciativa tiene como finalidad optimizar la capacidad de respuesta ante crisis sanitarias futuras y mejorar significativamente el acceso a vacunas en regiones clave de Colombia, Argentina, Chile, México, Perú y Brasil.

# Para respaldar esta decisión, se ha desarrollado un análisis integral de datos que contempla indicadores críticos como la incidencia de COVID-19, las tasas de vacunación y la disponibilidad de infraestructura sanitaria. El propósito central de este informe es proporcionar insights confiables y visualizaciones interactivas que faciliten la identificación de zonas con alta demanda sanitaria y baja cobertura, priorizando así ubicaciones óptimas para la expansión.

# Mediante el uso de herramientas de programación en Python y técnicas de análisis exploratorio, limpieza de datos y desarrollo de dashboards dinámicos, se busca dotar a la directiva de BIOGENESYS con una base sólida para la toma de decisiones estratégicas, basadas en evidencia y orientadas al impacto social y operativo de la empresa en el contexto postpandémico.

# 

# Desarrollo del proyecto

# Para llevar adelante el análisis de expansión estratégica propuesto por BIOGENESYS, se desarrolló una metodología basada en el uso de herramientas de programación con Python y técnicas de visualización dinámica mediante dashboards interactivos. Este enfoque permite procesar grandes volúmenes de datos sanitarios y demográficos de manera eficiente, garantizando la obtención de insights confiables.

# Metodología de Recopilación y Selección de Datos

# Importación y Carga Inicial Se utilizaron librerías de Python especializadas en análisis de datos y visualización, incluyendo NumPy, Pandas, Matplotlib.pyplot y Seaborn. El conjunto de datos original fue importado desde un archivo .csv, obteniendo inicialmente un volumen de 12.216.057 registros distribuidos en 50 columnas.

# Filtrado por Regiones Relevantes De acuerdo con los objetivos estratégicos del proyecto, se realizó un primer filtrado sobre la columna location\_key para seleccionar exclusivamente los países de interés: Argentina, Brasil, Chile, Colombia, México y Perú. Esto redujo la muestra a 5.946 registros.

# Filtrado Temporal Siguiendo los lineamientos de análisis post-pandémico, se eliminaron los registros anteriores al 1 de enero de 2021, quedando una base ajustada de 3.750 registros y manteniendo las 50 variables originales.

# Evaluación y Limpieza de Datos Se realizaron diversas etapas de saneamiento para asegurar la calidad del dataset:

# Se verificó que no existieran columnas ni filas completamente vacías.

# Se detectaron dos columnas con más del 50% de valores nulos (new\_recovered, cumulative\_recovered), las cuales se eliminaron para evitar sesgos analíticos.

# En columnas restantes con valores nulos, se aplicaron criterios diferenciados:

# Las columnas new\_confirmed, new\_deceased y cumulative\_vaccine\_doses\_administered fueron completadas con ceros, dado que la ausencia de registros podría representar falta de reporte y no necesariamente ausencia de casos.

# Las variables climáticas y acumulativas (new\_confirmed, new\_deceased, minimum\_temperature\_celsius, maximum\_temperature\_celsius, average\_temperature\_celsius, rainfall\_mm, relative\_humidity) fueron completadas con la media de cada variable, al tratarse de fluctuaciones mínimas y no impactar sustancialmente las métricas finales.

# Depuración Adicional y Validación Final

# Se identificó un valor negativo en la columna new\_confirmed, que fue eliminado por considerarse un posible error de carga.

# Se aplicaron métricas descriptivas (median, variance, min, max) para validar la consistencia de las variables numéricas.

# Se confirmó la ausencia total de nulos en el dataset final.

# La base procesada fue exportada bajo el nombre DatosFinalesFiltrados.csv, conformada por 3.749 registros y 48 variables limpias.

# 

# EDA e insights

# 

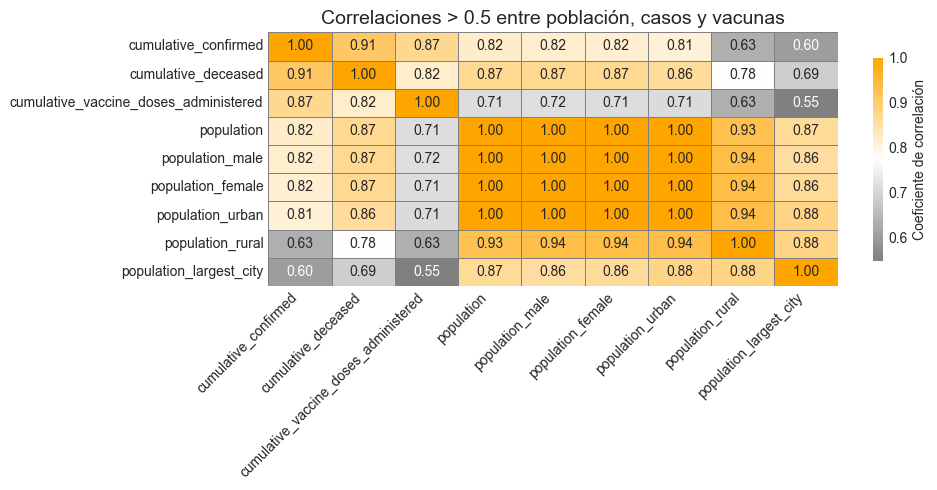
**Análisis de Nuevos Casos y Nuevos Decesos por País**

Se desarrolló una visualización comparativa en formato de barras para evaluar la cantidad de **nuevos casos confirmados** y **nuevos decesos** por COVID-19 en los seis países latinoamericanos seleccionados. Esta gráfica permite detectar rápidamente las regiones con mayor impacto reciente de la pandemia, facilitando la priorización estratégica.

Principales Insights:

* **Brasil** presenta la **mayor cantidad de nuevos casos confirmados**, lo que sugiere una alta demanda sanitaria potencial y la necesidad de reforzar la capacidad de respuesta.
* Sorprendentemente, **Chile** lidera en el número de **nuevos decesos**, lo que podría indicar limitaciones en la infraestructura de atención crítica o demoras en el acceso a vacunas.
* **Argentina**, **México** y **Perú** se posicionan como regiones con impacto intermedio en ambas métricas, lo cual las convierte en candidatas clave para expansión si se confirma baja cobertura sanitaria.
* **Colombia** muestra cifras moderadas en ambos indicadores, aunque requiere análisis adicional para correlacionar con infraestructura y cobertura de vacunación.

Conclusión del Análisis:Este gráfico permite identificar desequilibrios entre **nuevos contagios** y **letalidad**, lo que puede sugerir fallas en los sistemas de atención primaria o desigualdades en la administración de vacunas.



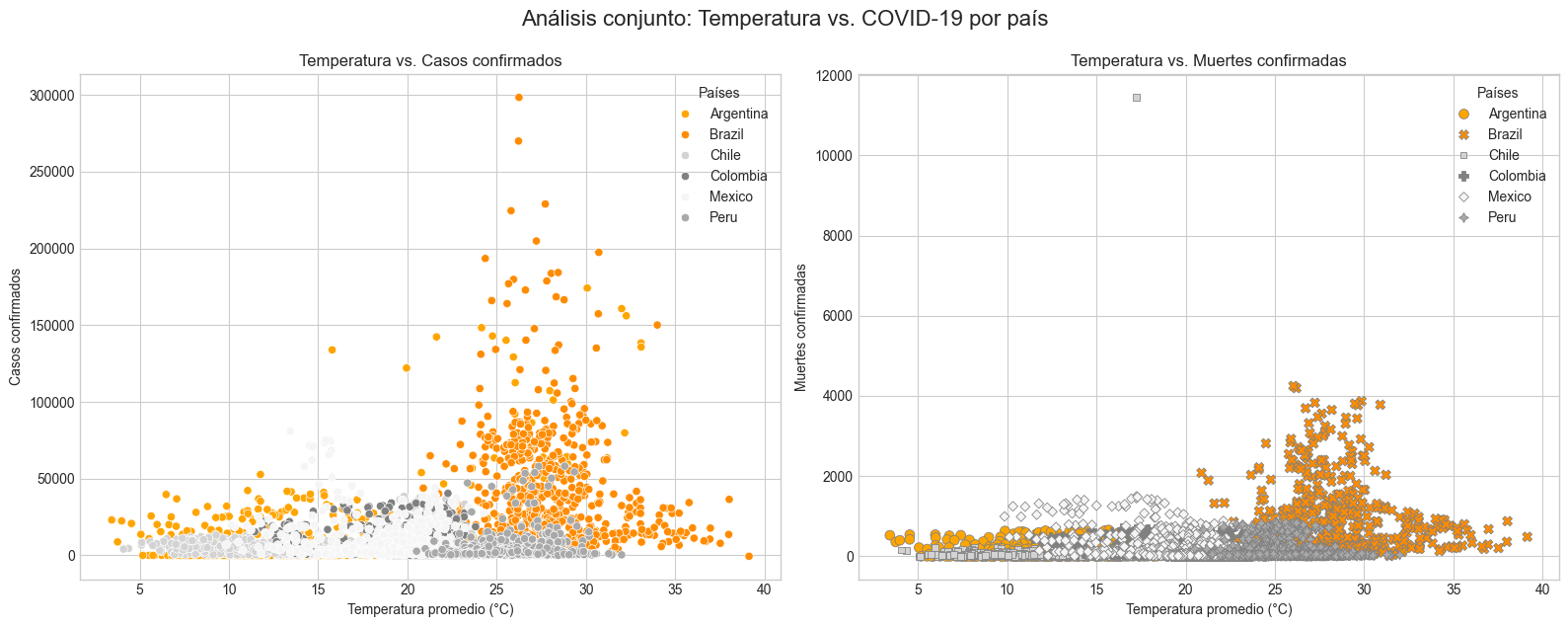
**Matriz de Correlaciones entre Población, Casos y Vacunas**

Se elaboró una matriz de correlación para evaluar relaciones significativas entre variables demográficas (población total, urbana, rural, por género) y sanitarias (casos confirmados, decesos, dosis administradas de vacunas). El análisis se centró exclusivamente en coeficientes superiores a 0.5, filtrando correlaciones débiles para destacar únicamente las relaciones más relevantes.

Principales Insights:

* Existe una **fuerte correlación entre la población total** y el **número acumulado de casos confirmados**, decesos y dosis de vacunas administradas. Esto confirma que los países con mayor densidad poblacional concentran mayor carga sanitaria y mayor actividad logística en vacunación.
* La variable **población urbana** también mostró alta correlación con los indicadores sanitarios, lo que refuerza la idea de priorizar **centros urbanos** en la expansión de laboratorios, por su mayor exposición y capacidad de absorción logística.
* La correlación entre **dosis administradas** y **casos confirmados** fue significativa, lo que podría reflejar una adecuada reacción sanitaria frente al aumento de casos en ciertas regiones.
* Las variables de **género** y **entorno rural** no mostraron correlaciones destacables por encima del umbral, lo cual sugiere que estos factores podrían no ser diferenciales para la selección de ubicación en esta etapa del análisis.

Conclusión del Análisis:Este enfoque permite identificar los ejes demográficos más vinculados a los indicadores de impacto sanitario. Al evidenciar fuertes asociaciones entre población urbana y casos/vacunación, se justifica concentrar los esfuerzos estratégicos en zonas densamente pobladas con alta exposición, favoreciendo así decisiones más eficientes y focalizadas en la expansión territorial.



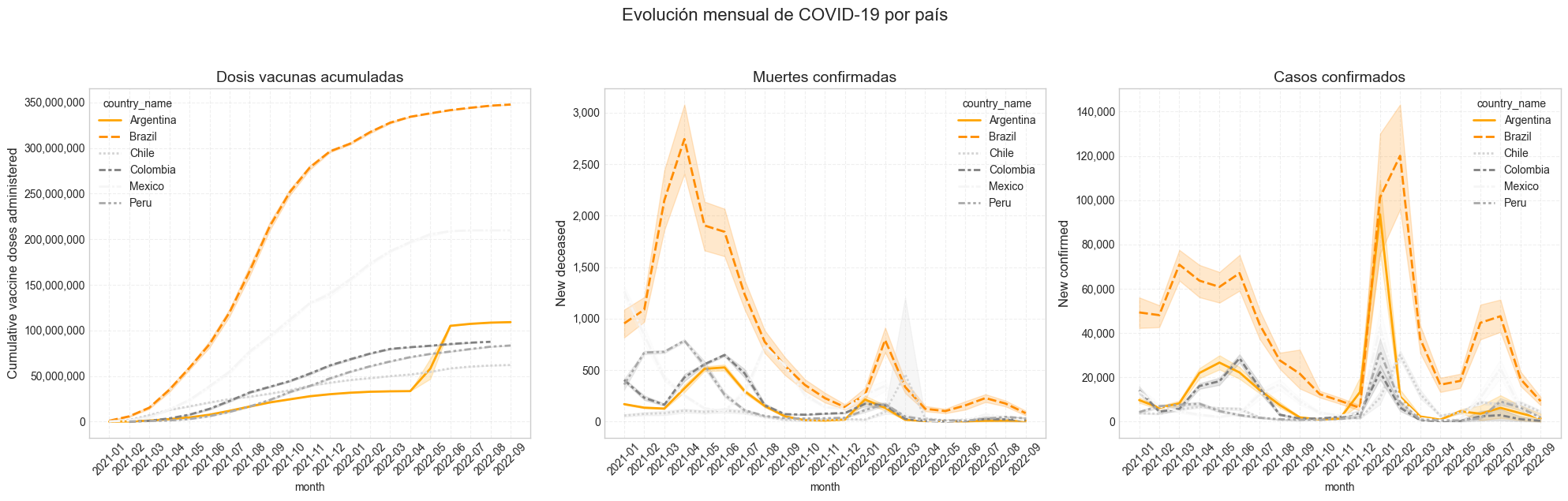
**Relación entre Temperatura Promedio y Casos/Muertes Confirmadas**

Se utilizaron scatterplots segmentados por país para visualizar cómo la temperatura media podría influir en la cantidad de nuevos casos confirmados y nuevos decesos por COVID-19. Esta visualización ayuda a explorar una posible relación entre condiciones climáticas y la propagación o gravedad del virus.

Principales Insights:

* **Temperaturas más elevadas no garantizan menor incidencia**. Países con climas cálidos como Brasil y México siguen mostrando niveles altos de casos confirmados, lo que sugiere que el clima por sí solo no es un factor determinante para la propagación.
* **Chile y Argentina**, con temperaturas más frías en ciertos momentos del año, presentan también cifras relevantes tanto en nuevos casos como en decesos, sin una clara disminución asociada al clima.
* No se observa una **correlación lineal evidente** entre la temperatura promedio y el número de casos o fallecimientos. Esto indica que variables como densidad poblacional, políticas sanitarias o acceso a servicios de salud probablemente tengan mayor peso en los resultados.
* La dispersión por país resalta diferencias estructurales: **Brasil** se destaca en volumen de casos, mientras **Chile** y **Perú** presentan datos más concentrados y variables en muertes.

Conclusión del Análisis:La temperatura promedio no puede considerarse un predictor confiable por sí solo en la planificación de nuevos laboratorios. Sin embargo, al combinar esta variable con otras (densidad poblacional, vacunación, cobertura sanitaria), puede contribuir al diseño de estrategias más integradas y adaptadas al contexto regional.



**Evolución Mensual de Casos, Muertes y Vacunación por País**

Se desarrollaron tres gráficos de líneas que permiten observar la evolución mensual de **casos confirmados**, **muertes** y **dosis acumuladas de vacunas** en los seis países seleccionados. Esta visualización múltiple facilita la comparación temporal entre métricas clave, ayudando a comprender cómo cada país enfrentó la pandemia y en qué momento se reforzaron las estrategias sanitarias.

Insights por cada métrica:

Dosis de Vacunas Acumuladas

* **Brasil** lidera con la mayor cantidad de dosis administradas a lo largo del periodo analizado, mostrando una estrategia de vacunación sostenida.
* **México** y **Argentina** también reflejan curvas de crecimiento consistentes, aunque en menor escala.
* El ritmo de vacunación se incrementa notablemente a partir de mediados de 2021, coincidiendo con las olas más graves del virus.

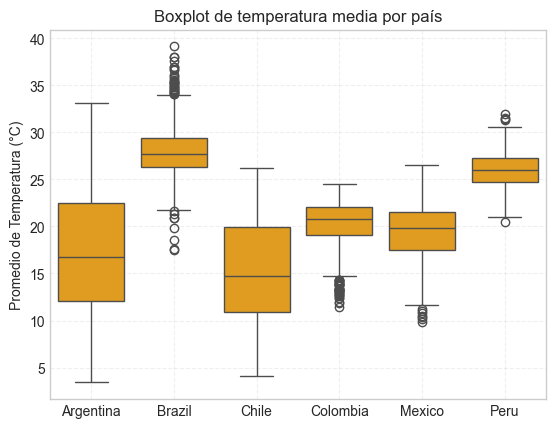
Muertes Confirmadas

* Se observa un pico generalizado de muertes alrededor de mediados de 2021, especialmente en **Brasil**, **Argentina** y **Chile**.
* Estas tendencias revelan momentos críticos de saturación sanitaria que deberían considerarse al elegir zonas para refuerzo de infraestructura.

Casos Confirmados

* Las curvas de nuevos contagios reflejan una dinámica similar a las muertes, con **picos pronunciados** durante 2021 y comportamientos más estables hacia 2022.
* **Brasil** vuelve a destacar con la mayor incidencia, lo que sugiere alta exposición poblacional y necesidad continua de vigilancia epidemiológica.

Conclusión del análisis:La evolución mensual de las tres métricas muestra una relación directa entre aumento de casos y el refuerzo de campañas de vacunación. La visualización resalta puntos de inflexión donde la respuesta sanitaria se intensificó, permitiendo a BIOGENESYS identificar momentos críticos y regiones con mayor presión logística. Estos hallazgos pueden guiar la instalación de nuevos laboratorios en zonas con historial de alta demanda y necesidad persistente de apoyo estructural.



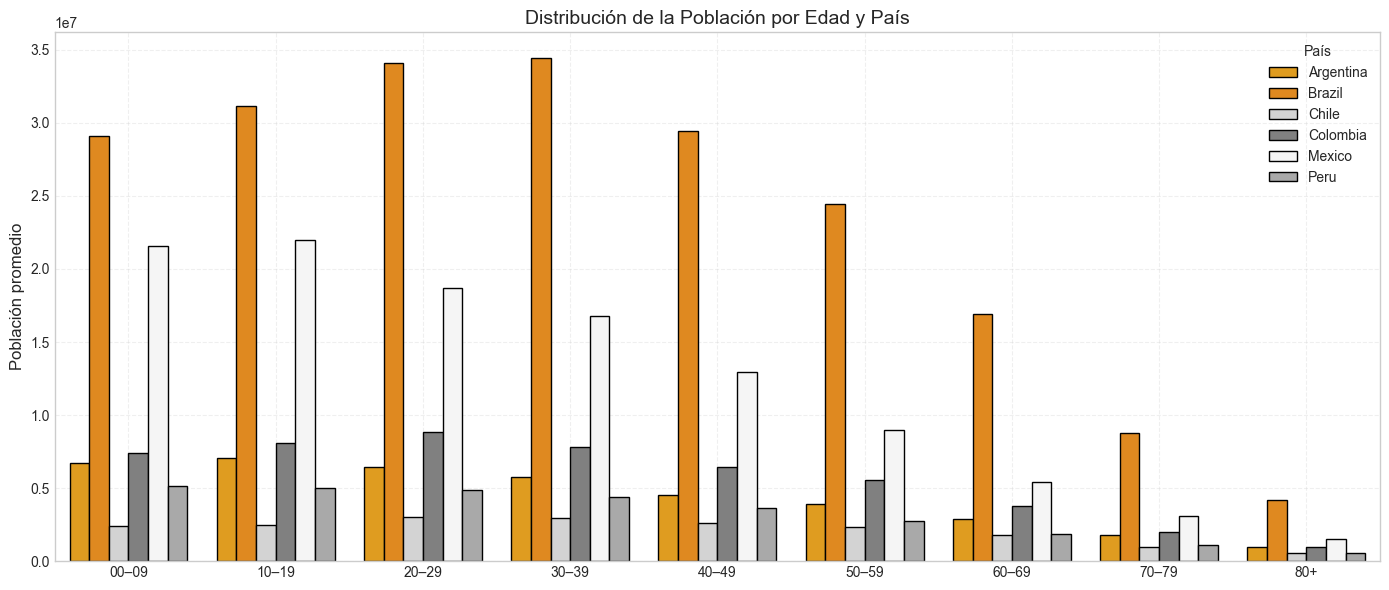
**Distribución de la Temperatura Media por País**

Se construyó un gráfico tipo *boxplot* para representar la variabilidad de la **temperatura promedio diaria** registrada en cada uno de los países analizados. Esta visualización permite identificar medianas, rangos intercuartiles y valores extremos, aportando contexto ambiental al estudio epidemiológico.

Principales Insights:

* **Chile** presenta las temperaturas promedio más bajas del conjunto, con baja dispersión, lo cual podría correlacionarse con mayores tasas de mortalidad observadas en etapas previas del análisis.
* **México** y **Colombia** muestran distribuciones más cálidas, con medianas cercanas a los 20 °C, sugiriendo una exposición climática más homogénea.
* **Brasil** evidencia una dispersión térmica amplia, posiblemente por su variedad geográfica, lo cual refuerza la necesidad de segmentar análisis por regiones internas si se considera expansión allí.
* La **mediana de temperatura en Argentina y Perú** se sitúa en rangos moderados, pero con presencia de outliers fríos, lo que podría influir en la distribución local del virus o en la planificación logística.

Conclusión del análisis: Si bien la temperatura por sí sola no mostró correlación directa con la incidencia del COVID-19 en visualizaciones anteriores, esta distribución térmica resulta útil para contextualizar aspectos **logísticos y operativos** de la expansión. Por ejemplo, zonas con climas extremos podrían requerir consideraciones adicionales en infraestructura, conservación de vacunas o transporte especializado.



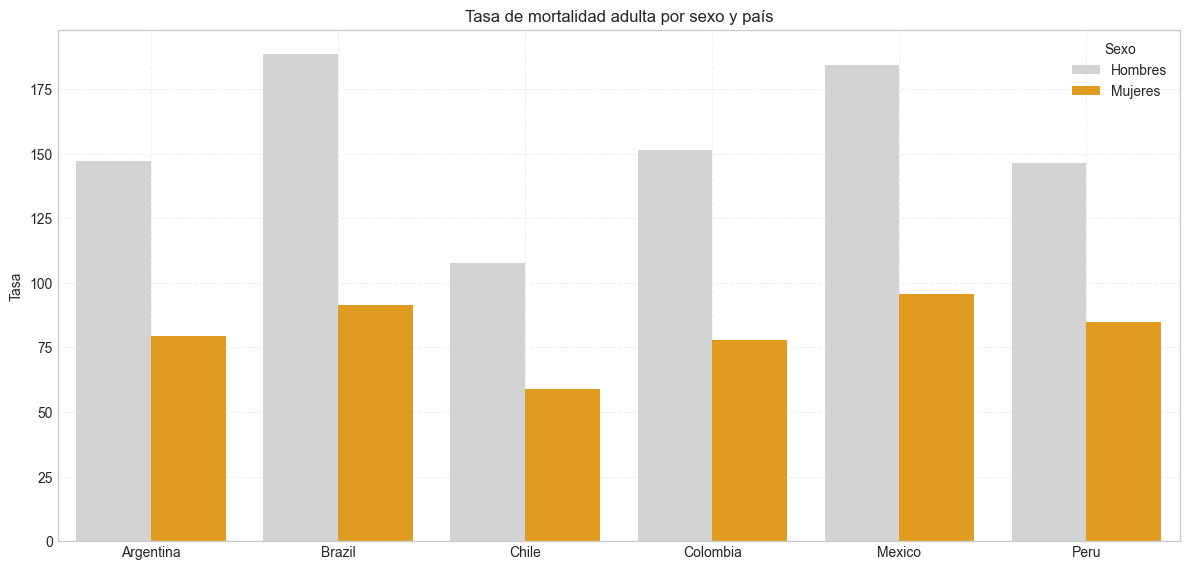
**Comparación de la Población por Grupo Etario y País**

Se elaboró un gráfico de barras agrupadas que presenta la **distribución promedio de la población por grupo etario** en los países analizados. Este enfoque permite identificar tendencias demográficas relevantes para la planificación de centros sanitarios y campañas de vacunación, según las características poblacionales de cada país.

Principales Insights:

* Los grupos **infantiles y adolescentes (00–09, 10–19)** tienen una presencia destacada en **Brasil, México y Colombia**, lo que indica regiones con alta concentración de población joven. Esto sugiere la necesidad de **campañas de vacunación escolar y pediátrica** reforzadas en esas zonas.
* Los grupos adultos **(20–49 años)** son predominantes en todos los países, particularmente en **México, Brasil y Argentina**, representando una fracción vital para la economía y la transmisión comunitaria del virus. Este segmento debe ser considerado prioritario en la expansión de laboratorios para asegurar rápida atención en contextos productivos.
* La población mayor de **60 años**, si bien más reducida en volumen, tiene **alto impacto sanitario**. **Chile** y **Argentina** presentan proporciones destacadas en los rangos **60–69**, **70–79** y **80+**, lo que justifica reforzar allí los servicios de salud y vacunación geriátrica.
* La visualización revela una estructura demográfica **asimétrica** entre los países, con diferentes proporciones según grupo etario, lo que **exige una estrategia adaptada** por país y no uniforme para toda la región.

Conclusión del análisis:Este gráfico permite ajustar la estrategia de expansión territorial de BIOGENESYS en función de la demanda potencial por edad. Las regiones con mayor cantidad de niños y adolescentes requieren infraestructura accesible y móvil, mientras que aquellas con población envejecida demandan centros especializados con mayor capacidad asistencial. Incorporar el enfoque demográfico fortalece la toma de decisiones basada en evidencia y maximiza el impacto social del proyecto.



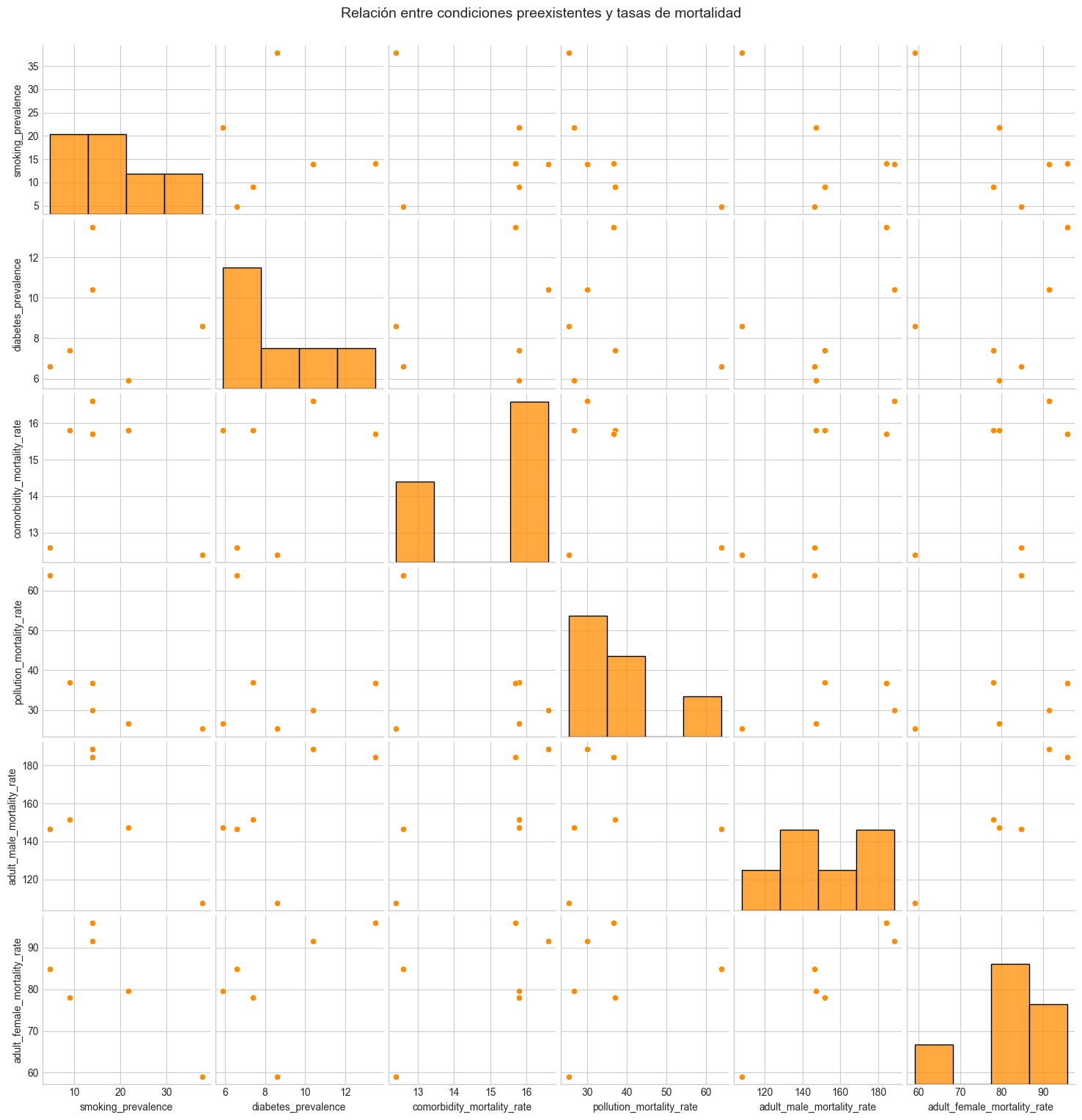
**Tasa de Mortalidad Adulta por Sexo y País**

Se generó un gráfico de barras agrupadas que compara la tasa de mortalidad adulta en hombres y mujeres en los seis países analizados. Este enfoque revela desigualdades de género en términos de impacto sanitario, lo cual puede influir en la priorización de infraestructura médica y campañas preventivas segmentadas.

Principales Insights:

* En todos los países, la tasa de mortalidad adulta **es consistentemente mayor en hombres** que en mujeres, lo que sugiere una posible vulnerabilidad estructural que podría abordarse mediante iniciativas específicas de salud masculina.
* **México** presenta la **mayor tasa de mortalidad masculina**, mientras que **Brasil** tiene el valor más alto en **mujeres**. Esto podría indicar diferencias en políticas de atención médica o factores socioeconómicos que impactan la longevidad por género.
* **Chile** muestra las **tasas más bajas en ambos sexos**, destacándose por su perfil sanitario favorable, lo que lo posiciona como un referente para estrategias de prevención.
* Las brechas entre sexos no son homogéneas: países como **Argentina y Perú** presentan diferencias moderadas, mientras que en **Colombia** la distancia es más marcada, lo que invita a considerar enfoques focalizados por país.

Conclusión del análisis:Las desigualdades de género en mortalidad adulta revelan la necesidad de incluir la variable sexo en el diseño de políticas sanitarias y ubicación de recursos médicos. BIOGENESYS puede incorporar este criterio para promover equidad en el acceso a sus servicios, asegurando que hombres y mujeres reciban atención proporcional al nivel de riesgo observado en cada región.



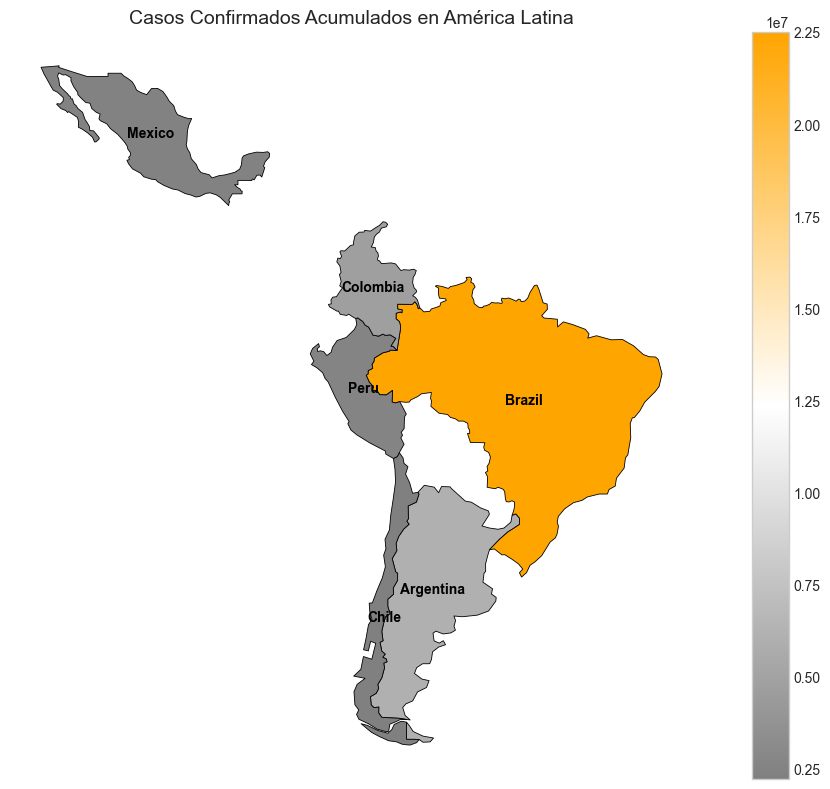
**Relación entre Condiciones Preexistentes y Tasas de Mortalidad Adulta**

Se desarrolló una visualización multivariada (*pairplot*) para evaluar la relación entre factores de riesgo poblacional —como prevalencia de tabaquismo, diabetes, comorbilidades y contaminación— y las tasas de mortalidad adulta (masculina y femenina) en cada país analizado. Esta técnica permite observar correlaciones entre variables de salud pública que influyen directamente en la capacidad de respuesta sanitaria.

Principales Insights:

* La **prevalencia de tabaquismo** se relaciona positivamente con la **mortalidad masculina**, lo que sugiere que hábitos nocivos pueden impactar de forma más pronunciada en este grupo demográfico.
* Las **comorbilidades** y la **prevalencia de diabetes** presentan correlaciones moderadas con ambas tasas de mortalidad, lo que refuerza la importancia de atender enfermedades crónicas como parte de la estrategia sanitaria.
* La **tasa de mortalidad atribuida a contaminación** muestra una dispersión notable, pero en países con altos niveles parece coincidir con tasas elevadas de mortalidad general, lo que indica riesgos estructurales en ambientes urbanos mal gestionados.
* En general, las tasas de mortalidad femenina presentan menos variabilidad asociada a estos factores, lo que podría sugerir diferencias en acceso o respuesta a tratamientos preventivos entre géneros.

Conclusión del análisis:Este enfoque evidencia que factores preexistentes —muchos de ellos prevenibles— tienen una relación directa con las tasas de mortalidad adulta, especialmente en hombres. BIOGENESYS puede utilizar esta información para **priorizar regiones con mayor carga de enfermedades crónicas y ambientales**, integrando componentes de educación y prevención en sus centros de expansión. Incorporar estas dimensiones mejora la eficiencia operativa y fortalece el impacto social de la estrategia.



**Mapa de Casos Confirmados Acumulados en América Latina**

Se desarrolló una visualización geográfica utilizando datos sanitarios agregados por país junto con geometrías oficiales de fronteras nacionales. Este mapa temático representa los **casos confirmados acumulados de COVID-19** en los seis países objetivos del proyecto, reforzando el enfoque territorial de expansión de BIOGENESYS.

Principales Insights:

* **Brasil** destaca por su coloración más intensa (naranja), representando la **mayor cantidad de casos acumulados**. Este dato lo posiciona como un candidato prioritario para el refuerzo sanitario y logístico.
* Los países **Argentina, México y Colombia** se ubican en un rango intermedio de incidencia, sugiriendo potenciales zonas de expansión especialmente en áreas con alta densidad urbana y menor cobertura.
* **Chile y Perú**, con tonalidades más claras, presentan valores menores en casos acumulados. Sin embargo, su análisis debe complementarse con indicadores como tasas de mortalidad, estructura demográfica y cobertura médica para determinar prioridad estratégica.
* La representación geográfica con etiquetas visibles sobre el mapa permite **identificar rápidamente las regiones objetivo** y facilita su incorporación en planes operativos y presentaciones directivas.

Conclusión del análisis:El componente geoespacial potencia la interpretación estratégica del entorno sanitario. Este mapa permite a BIOGENESYS visualizar rápidamente la distribución de carga epidemiológica en Latinoamérica, priorizando la expansión hacia países con alta incidencia acumulada. Se recomienda complementar esta visualización con mapas adicionales de infraestructura médica y distribución poblacional para obtener una perspectiva más integral.

# 

# Análisis del dashboard

# Página 1: Portada Geoespacial del Análisis Biogenesys

# Navegación:

# Esta página funciona como pantalla de bienvenida con enfoque geográfico. Al centro, se muestra un mapa de América Latina con los países estratégicos resaltados: México, Colombia, Brasil, Argentina, Chile y Perú.

# En el lateral derecho hay una barra de navegación vertical con secciones temáticas del dashboard: Portada, General, Correlaciones, Temperaturas y Afecciones. Se puede hacer clic en estas secciones para ir a las visualizaciones específicas.

# Elementos destacados:

# El mapa es interactivo (si lo configuraste con tooltips o filtros vinculados), permitiendo visualizar diferencias regionales o abrir visuales complementarias al seleccionar cada país.

# La portada incluye el logo de Biogenesys y tipografía institucional, dando un marco corporativo al dashboard.

**Conclusión estratégica:** Esta página introduce el proyecto y sitúa visualmente los países objetivo. Es útil para comenzar cualquier presentación ejecutiva, ya que establece de forma clara el alcance regional del análisis y permite navegar intuitivamente entre las distintas secciones temáticas del informe.

# Página 2: Análisis General

# Cómo navegarla:

# Usá el selector de país (arriba a la izquierda) para filtrar los datos y visualizar métricas específicas por región.

# El menú de periodos mensuales (enero 2021 a septiembre 2022) permite explorar la evolución temporal y detectar picos críticos.

# Las tarjetas resumen muestran indicadores agregados: casos confirmados, decesos, dosis administradas y población total.

# Visualizaciones clave:

# Dos gráficos de línea comparan mensualmente:

# *Nuevos casos confirmados*

# *Nuevas muertes*

# *Dosis administradas*

# Están desglosados por país, lo que facilita identificar diferencias en respuesta sanitaria.

# Conclusiones posibles: Esta página permite a BIOGENESYS comparar el comportamiento de la pandemia entre países, entender cómo cada uno respondió con campañas de vacunación, y detectar cuándo se produjo el mayor impacto sanitario. Es ideal para fundamentar decisiones sobre *temporalidad de instalación* y *dimensionamiento de recursos* según el país seleccionado.

# Página 3: Correlaciones > 0.5 entre población, casos y vacunas

# Cómo navegarla:

# Usá el slicer de país y mes (arriba a la izquierda) para filtrar los datos por región específica y temporalidad. Esto ajusta dinámicamente la matriz de correlación para mostrar únicamente los registros que cumplen esa condición.

# El centro de la página muestra una heatmap con coeficientes de correlación entre variables claves: casos confirmados, muertes, vacunas, población total, urbana, rural, hombres, mujeres y residentes en grandes ciudades.

# Qué visualizar primero:

# Fijate en los valores altos de correlación (de 0.70 a 1.00), especialmente entre:

# Población y muertes/casos

# Población urbana y vacunación

# Género y carga poblacional

# Los colores cálidos (naranja intenso) indican relaciones fuertes; los grises representan correlaciones moderadas.

# Conclusiones posibles: Esta página muestra cómo ciertas variables demográficas están directamente asociadas a la incidencia sanitaria. Es útil para justificar decisiones de expansión territorial en zonas densamente pobladas, con alta carga epidemiológica o fuerte actividad de vacunación. También permite identificar qué indicadores pueden servir como proxy en zonas con datos incompletos.

# Página 4: Relación entre Temperatura Promedio y COVID-19

# Cómo navegarla:

# Utilizá los botones superiores para seleccionar país (Argentina, Chile, México, Brasil, Colombia o Perú) y período mensual (de enero 2021 a febrero 2022). Esto actualiza los gráficos a la derecha según la región y mes elegidos.

# La página incluye dos scatterplots comparativos:

# A la izquierda: temperatura promedio vs. nuevos casos confirmados.

# A la derecha: temperatura promedio vs. muertes confirmadas.

# Qué observar:

# Cada país se representa con un color diferente, permitiendo identificar visualmente diferencias de comportamiento frente a la temperatura.

# Al filtrar, podés ver si hay patrones estacionales en la evolución del virus y si la temperatura afecta directa o indirectamente los indicadores sanitarios.

# Conclusiones posibles: Esta página permite analizar si el clima influye en la propagación o letalidad del COVID-19 en cada país. Aunque no se observa una correlación clara, sí sirve como complemento contextual para la toma de decisiones logísticas (ej. transporte de vacunas, diseño de instalaciones médicas).

# Página 5: Afecciones Preexistentes y Mortalidad

# Cómo navegarla:

# En el panel izquierdo podés filtrar por país (Argentina, Brasil, Chile, Colombia, México, Perú) y por mes (de enero 2021 a abril 2022). Esto actualiza el gráfico central para mostrar datos específicos de cada contexto.

# La visualización principal es un scatter matrix (pairplot) que compara variables como:

# Prevalencia de tabaquismo

# Prevalencia de diabetes

# Tasa de mortalidad atribuida a comorbilidades

# Tasa de mortalidad atribuida a contaminación

# Mortalidad adulta masculina y femenina

# Qué observar:

# En cada intersección podés identificar patrones de correlación. Por ejemplo, si hay concentración de puntos en diagonal o agrupamientos en ciertas regiones del gráfico, pueden revelar asociaciones fuertes entre factores de riesgo y mortalidad.

# Conclusiones posibles: Esta página permite detectar relaciones estructurales entre condiciones de salud poblacional y las tasas de mortalidad. Es útil para anticipar necesidades médicas en regiones con alta carga de enfermedades crónicas o contaminación ambiental. BIOGENESYS puede usar esta información para priorizar zonas donde las causas de mortalidad son evitables mediante prevención y seguimiento.

# 

# Conclusiones y Recomendaciones

# El análisis integral de los datos epidemiológicos, demográficos y estructurales en los seis países objetivo de Latinoamérica —Argentina, Brasil, Chile, Colombia, México y Perú— ha permitido identificar patrones relevantes para la expansión estratégica de laboratorios y centros de vacunación de BIOGENESYS.

# Las visualizaciones revelan que:

# Brasil, por volumen de casos, extensión geográfica y alta densidad poblacional, representa el mayor foco sanitario y logístico, con fuerte necesidad de respuesta estructural.

# Las regiones urbanas concentran los valores más altos de incidencia y vacunación, indicando que los entornos densamente poblados deben ser priorizados.

# La estructura etaria varía por país: México y Brasil presentan alta presencia de población joven, mientras que Chile y Argentina concentran adultos mayores, lo cual demanda estrategias diferenciadas.

# Existe una desigualdad marcada en las tasas de mortalidad por género y condiciones preexistentes (como tabaquismo y diabetes), lo que habilita una política sanitaria segmentada.

# Las variables climáticas y ambientales no mostraron una relación directa con el aumento de casos, pero deben considerarse en la logística de transporte y conservación de insumos médicos.

# Recomendaciones Estratégicas para BIOGENESYS

# A partir de los hallazgos anteriores, se proponen las siguientes acciones estratégicas:

| Objetivo | Recomendación |
| --- | --- |
| Priorizar zonas de alto impacto epidemiológico | Expandir operaciones en Brasil, México y Argentina, donde se registran altos niveles de casos confirmados, mortalidad y población activa. |
| Atender estructuras etarias específicas | Diseñar campañas especializadas en función de la edad: vacunación pediátrica en países con alta población joven y refuerzo geriátrico en Chile y Argentina. |
| Considerar desigualdades de género en salud | Incorporar unidades con enfoque en salud masculina, especialmente en México y Brasil, donde se detectan mayores tasas de mortalidad en hombres. |
| Integrar el componente ambiental y logístico | Asegurar infraestructura adaptable a condiciones climáticas y ambientales adversas, especialmente en regiones de alta contaminación. |
| Optimizar la red de distribución | Instalar centros en nodos urbanos estratégicos con buena accesibilidad para maximizar cobertura y eficiencia en el suministro de vacunas. |
| Fomentar análisis continuo y monitoreo regional | Mantener dashboards interactivos actualizados que permitan a la directiva seguir tendencias y ajustar decisiones de manera ágil. |

# 

# Ubicaciones Óptimas por País

| País | Ciudad/Región Sugerida | Justificación Estratégica |
| --- | --- | --- |
| Brasil | São Paulo y Manaus | Alta incidencia acumulada, densidad poblacional elevada, infraestructura médica desarrollada en São Paulo y necesidad de refuerzo en zonas amazónicas como Manaus. |
| México | Ciudad de México y Guadalajara | Alta concentración de población joven, elevada tasa de mortalidad masculina, buena conectividad logística. |
| Argentina | Buenos Aires y Córdoba | Alta proporción de adultos mayores, infraestructura sanitaria sólida pero con desigualdades regionales. |
| Chile | Santiago y Valparaíso | Alta tasa de mortalidad, población envejecida, buena cobertura urbana pero necesidad de expansión periférica. |
| Colombia | Bogotá y Medellín | Alta densidad urbana, desigualdad en acceso sanitario, brechas por género y comorbilidades. |
| Perú | Lima y Arequipa | Dispersión geográfica, desafíos en cobertura rural, alta carga de enfermedades crónicas y ambientales. |

# Criterios de Selección Aplicados

# Carga epidemiológica: Casos confirmados y decesos acumulados.

# Demografía: Estructura etaria, densidad poblacional y distribución urbana/rural.

# Infraestructura sanitaria: Acceso a centros médicos, cobertura pública y privada.

# Condiciones preexistentes: Prevalencia de comorbilidades y factores de riesgo.

# Logística y conectividad: Accesibilidad terrestre y aérea para distribución de insumos.

# 

# Reflexión personal

# Si tuviera que volver a comenzar este proyecto desde cero, mantendría la mayoría de las decisiones metodológicas, ya que permitieron asegurar la limpieza, estructura y relevancia de los datos para obtener insights confiables. La planificación por etapas —importación, filtrado, tratamiento de nulos, visualización y análisis— fue eficaz y permitió desarrollar un marco sólido para la toma de decisiones estratégicas.

# Sin embargo, existe un aspecto que reconsideraría: la eliminación de las columnas new\_recovered y cumulative\_recovered. Si bien en ese momento se justificó por el elevado porcentaje de valores nulos, reflexionando sobre el impacto de estas métricas en el contexto sanitario, resulta evidente que la recuperación es un indicador crucial para evaluar la efectividad del sistema de salud, la evolución de la pandemia y la capacidad de atención de cada país.

# En proyectos futuros, propondría conservar esas columnas y explorar métodos alternativos para su tratamiento, como imputación estadística o integración con fuentes complementarias, con el fin de enriquecer el análisis y proporcionar una visión más completa del ciclo epidemiológico.