

EXPANSIÓN ESTRATÉGICA DE BIOGENESYS CON PYTHON

Por: Valeria Lozano Castañeda

Email: valerialozano0600@gmail.com

Cohorte: DA-FT12

Fecha de entrega: 07/04/2025

Institución: Laboratorios Biogenesys

INTRODUCCIÓN

Propósito del Proyecto:

Este proyecto tuvo como finalidad desarrollar una herramienta analítica en Power BI para la empresa BIOGENESYS, con el fin de apoyar la toma de decisiones estratégicas relacionadas con la expansión de sus laboratorios y centros de vacunación en América Latina. Para ello, se empleó la visualización de datos clave en salud, demografía y aspectos socioeconómicos.

Logros a Nivel Organizacional:

- **Detección de Zonas Estratégicas para Nuevos Laboratorios:** Se identificaron regiones ideales para la expansión, basándose en distintos criterios de análisis.
- **Evaluación de Infraestructura Médica:** Se analizaron la cantidad de médicos y enfermeros disponibles por cada 1.000 habitantes.
- **Capacidad Hospitalaria:** Se examinó la infraestructura de los hospitales en diferentes regiones para establecer su nivel de adecuación.
- **Mejora de Estrategias de Vacunación:**
 - Se exploró la relación entre la cobertura de vacunación y la disminución de casos y muertes.
 - Se identificaron los enfoques de vacunación más eficaces por país.
- **Asignación de Recursos y Presupuesto en Salud:**
 - Se estudió la distribución del personal sanitario para orientar la inversión en recursos humanos.
 - Se abordaron problemáticas socioeconómicas que podrían dificultar la implementación de nueva infraestructura.
- **Análisis Demográfico y de Salud Pública:**
 - Se determinaron las áreas con poblaciones más vulnerables según el rango de edad.
 - Se observaron tendencias temporales en los casos y muertes para facilitar una planificación más precisa.

El proyecto logró cumplir con los objetivos de BIOGENESYS, al entregar una solución analítica sólida que respalda decisiones estratégicas. El dashboard desarrollado en Power BI facilitó la identificación

de ubicaciones clave para la expansión, el diseño de políticas de vacunación más efectivas y una mejor asignación de inversiones en salud y talento humano.

DESARROLLO DEL PROYECTO

RECOPIACIÓN DE DATOS

Fuentes Utilizadas:

La información utilizada en el análisis fue proporcionada directamente por el laboratorio, complementada con datos adicionales obtenidos desde diversas plataformas.

Métodos de Recolección:

- **APIs Públicas:** Se accedió a datos en tiempo real mediante APIs disponibles libremente, lo cual permitió mantener la información constantemente actualizada y precisa.
- **Archivos Descargables:** Se obtuvieron datasets en formatos como CSV, Excel y JSON desde portales de datos abiertos y sitios oficiales.
- **Web Scraping:** Se aplicaron técnicas de extracción de datos desde páginas web cuando estos no estaban disponibles para su descarga directa.

Selección de Información Relevante

1. Criterios de Relevancia:

- **Variables Principales:** Se priorizaron indicadores directamente relacionados con los objetivos del proyecto, como casos confirmados, muertes recientes, vacunas aplicadas, población total y disponibilidad de personal médico por cada 1.000 habitantes.
- **Factores Socioeconómicos:** Se integraron variables como el Índice de Desarrollo Humano (IDH) y el Producto Interno Bruto (PIB) per cápita para entender el contexto económico y social de cada región.

2. Evaluación de la Calidad de los Datos:

- **Fuentes Confiables:** Solo se trabajó con datos provenientes de entidades reconocidas para asegurar su precisión y fiabilidad.
- **Integridad de los Datasets:** Se seleccionaron conjuntos de datos que estuvieran completos y sin vacíos significativos, facilitando un análisis más robusto.

3. Aspectos Temporales:

- **Actualización:** Se garantizó que los datos utilizados correspondieran a las fechas más recientes disponibles.

- **Análisis Histórico:** En ciertos casos, se incluyeron registros pasados para identificar patrones, tendencias y realizar proyecciones temporales.

Limpieza y Transformación de Datos

- **Depuración de Información:**
 - **Eliminación de Duplicados:** Se detectaron y removieron registros repetidos para evitar errores en el análisis.
 - **Tratamiento de Valores Faltantes:** Se aplicaron métodos como imputación (relleno con promedios o medianas) o eliminación de registros cuando los vacíos eran significativos.
 - **Corrección de Errores:** Se revisaron y corrigieron datos atípicos o inconsistentes, tanto de forma manual como automática.
- **Transformación de Datos:**
 - **Estandarización:** Se aplicaron técnicas de normalización y escalamiento para homogeneizar los valores y permitir comparaciones entre diferentes variables.
 - **Cálculo de Nuevas Métricas:** Se crearon columnas derivadas para obtener indicadores como tasas de vacunación, mortalidad y densidad poblacional.
 - **Ajuste de Formatos:** Se verificó que todas las variables estuvieran en el tipo de dato adecuado, como fechas en formato datetime y valores numéricos en float o int.

Todo este proceso de recopilación, depuración y transformación fue clave para asegurar la calidad del análisis. Gracias a esta metodología, fue posible generar visualizaciones e indicadores que facilitaron la toma de decisiones estratégicas para BIOGENESYS. La preparación adecuada de los datos permitió construir una base sólida sobre la cual se pudieron desarrollar análisis detallados y con alto valor informativo.

ANÁLISIS EXPLORATORIO DE DATOS (EDA) E INSIGHTS OBTENIDOS

1. Comparativa entre Casos Activos y Casos Recuperados

1.1 Evaluación de la Propagación y Control del Virus

- **Casos Activos:** Representan la cantidad de personas actualmente infectadas, por lo que un alto volumen puede reflejar zonas donde el virus aún no está controlado.
- **Casos Recuperados:** Indican cuántas personas superaron la enfermedad, lo cual puede ser una señal de inmunidad comunitaria creciente y de la eficacia del sistema de salud local.

1.2 Distribución de Recursos Humanos y Sanitarios

- **Asignación de Personal:** Las regiones con mayor cantidad de casos activos requieren más personal médico, camas hospitalarias y equipos de protección personal (EPP).

- **Instalación de Centros de Vacunación:** Aquellas áreas con alta incidencia podrían beneficiarse de nuevos centros que aceleren el proceso de vacunación.

1.3 Ubicación Estratégica de Nuevos Laboratorios

- **Intervención en Zonas Prioritarias:** BIOGENESYS puede enfocar la instalación de nuevos laboratorios en áreas con mayor número de casos activos, facilitando el análisis de variantes y el desarrollo de tratamientos específicos.
- **Prevención de Rebrotos:** Las zonas con una combinación alta de casos activos y recuperados pueden requerir laboratorios destinados al monitoreo y la prevención a largo plazo.

1.4 Estrategias de Vacunación Más Efectivas

- **Campañas Dirigidas:** Las regiones con alta propagación y bajas tasas de recuperación pueden ser candidatas para intervenciones vacunales más agresivas.
- **Evaluación Comparativa:** Analizar los índices de recuperación en zonas vacunadas frente a las no vacunadas permite medir la efectividad de las campañas y rediseñar estrategias.

1.5 Logística e Impacto Económico

- **Planificación de Insumos:** Comprender la dinámica entre casos activos y recuperados ayuda a optimizar la distribución de vacunas e insumos médicos.
- **Visión a Futuro:** Las proyecciones basadas en esta relación también permiten tomar decisiones sobre la sostenibilidad y expansión de infraestructuras sanitarias en el largo plazo.

Resultados del Análisis por País

1. Argentina:

- **Casos Activos:** Se identificaron varias olas de contagio, con un pico destacado entre diciembre de 2021 y enero de 2022.
- **Casos Recuperados:** Mantienen un comportamiento estable, aunque en niveles más bajos que los casos activos.
- **Conclusión:** El desbalance sugiere que la situación sanitaria no está completamente controlada, lo que resalta la necesidad de reforzar las estrategias de contención.

2. Brasil:

- **Casos Activos:** Presenta una evolución con múltiples olas y una tendencia más irregular.
- **Casos Recuperados:** Acompañan en forma más cercana la evolución de los casos activos, mostrando un mayor nivel de recuperación.
- **Conclusión:** La correlación entre casos activos y recuperados puede ser reflejo de una respuesta sanitaria más eficaz, aunque la fluctuación sugiere la necesidad de mantener medidas de vigilancia y control.

3. Chile

- **Casos Activos:** Se identificó un pico importante al inicio de 2022.
- **Casos Recuperados:** Permanecen bajos y relativamente estables frente al alto número de casos activos.
- **Análisis:** Al igual que en Argentina, esto indica que el control del virus no ha sido del todo efectivo, lo que podría requerir una mayor asignación de recursos y fortalecimiento del sistema sanitario.

4. Colombia

- **Casos Activos:** El país ha atravesado varias olas de contagio con picos marcados a lo largo del tiempo.
- **Casos Recuperados:** Muestran una correlación más cercana con los casos activos, similar al comportamiento observado en Brasil.
- **Análisis:** Aunque el sistema de salud parece responder adecuadamente, las constantes olas indican que la amenaza sigue presente y se requiere mantener la vigilancia.

5. México

- **Casos Activos:** Varias fases de incremento, con un pico destacado a comienzos de 2022.
- **Casos Recuperados:** Se mantienen bajos en comparación, con escasa variación.
- **Análisis:** La brecha entre casos activos y recuperados sugiere limitaciones en el sistema de recuperación y control, lo que plantea la necesidad de fortalecer las estrategias de salud pública.

6. Perú

- **Casos Activos:** Comportamiento similar al de otros países, con varios picos de contagio.
- **Casos Recuperados:** Se mantienen en niveles bajos sin grandes variaciones.
- **Análisis:** Al igual que México y Argentina, el país muestra una limitada capacidad de recuperación, lo que indica la necesidad de una gestión más eficiente de la pandemia.

Conclusiones Generales

A partir del análisis comparativo entre los casos activos y recuperados en distintos países, se identifican necesidades específicas en términos de recursos sanitarios:

- **Alta demanda de recursos médicos:** En países como Argentina, Chile, México y Perú, donde los casos activos superan ampliamente a los recuperados, se requiere una mayor inversión en personal médico, infraestructura hospitalaria y campañas de inmunización.

- **Priorización estratégica:** Estas regiones deben ser consideradas prioritarias para la instalación de nuevos laboratorios y centros de vacunación por parte de BIOGENESYS, con el objetivo de contener la propagación del virus y monitorear posibles variantes.
- **Optimización en la toma de decisiones:** Comprender la relación entre los casos activos y recuperados permite evaluar la eficacia de las medidas implementadas, identificar áreas críticas y orientar mejor la distribución de los recursos.

Este enfoque analítico resulta clave para BIOGENESYS en su estrategia de expansión en América Latina. No solo facilita la identificación de zonas con mayores necesidades, sino que también contribuye al fortalecimiento del sistema de salud pública mediante decisiones informadas y basadas en evidencia.

2) Análisis de Series Temporales y Autocorrelación

Descomposición de Series Temporales

Para cada país analizado, se realizó una descomposición de la serie temporal de casos confirmados de COVID-19 en tres componentes fundamentales: la **tendencia (trend)**, que muestra la dirección general del comportamiento de los casos a lo largo del tiempo; la **estacionalidad (seasonal)**, que representa los patrones recurrentes; y los **residuos (resid)**, que capturan las variaciones no explicadas por los dos componentes anteriores.

- **Argentina:** La tendencia refleja un incremento notable de casos a finales de 2021 y principios de 2022. La estacionalidad muestra picos periódicos, mientras que los residuos destacan fluctuaciones no atribuibles a patrones regulares.
- **Brasil:** Se observan variaciones marcadas en la tendencia, con picos similares a los de Argentina. La estacionalidad es clara y regular, y los residuos presentan mayor amplitud, indicando posibles eventos atípicos.
- **Chile:** Exhibe un comportamiento similar al de Argentina y Brasil, con una tendencia al alza en diciembre de 2021. La estacionalidad es evidente, y los residuos también presentan variaciones no explicadas.
- **Colombia:** La tendencia marca varias olas sucesivas, acompañadas de un patrón estacional definido. Los residuos reflejan anomalías comparables a los demás países.
- **México:** La serie muestra incrementos significativos hacia el cierre de 2021 y el inicio de 2022. La estacionalidad está bien definida, y los residuos indican fluctuaciones inesperadas.
- **Perú:** Presenta picos de contagio relevantes a finales de 2021. Los patrones estacionales son consistentes, mientras que los residuos sugieren variaciones no explicadas por los otros componentes.

Análisis de Autocorrelación (ACF)

Las gráficas de autocorrelación permiten examinar cómo los valores actuales de la serie están relacionados con sus valores pasados, lo cual es útil para identificar ciclos o patrones temporales repetitivos.

- **Argentina:** La autocorrelación decrece rápidamente, revelando ciclos que sugieren estacionalidad.
- **Brasil:** Después de una caída inicial, aparecen ciclos recurrentes, evidenciando componentes estacionales.
- **Chile:** Comportamiento similar al de Argentina, con caídas rápidas seguidas de patrones cíclicos.
- **Colombia:** Presenta una disminución inicial y luego ciclos sostenidos, similar a Brasil.
- **México:** Muestra una fuerte señal cíclica tras la caída inicial, lo que reafirma la estacionalidad.
- **Perú:** Sigue el patrón general, con un descenso rápido seguido de ciclos marcados.

Aplicación Estratégica para BIOGENESYS

El análisis de series temporales y autocorrelación aporta una base sólida para la planificación y toma de decisiones por parte de BIOGENESYS, especialmente en cuanto a:

- **Detección de patrones estacionales:** Estos análisis permiten anticipar períodos críticos de incremento en casos, facilitando la planificación eficiente de recursos y personal sanitario.
- **Optimización de campañas de vacunación:** Conociendo los momentos de mayor incidencia, se pueden calendarizar campañas que maximicen su alcance y efectividad.
- **Identificación de tendencias a largo plazo:** Observar la evolución de los casos permite prever futuros picos y adaptar la expansión de laboratorios a las necesidades proyectadas.
- **Reconocimiento de anomalías:** Las irregularidades detectadas pueden ser indicativas de eventos extraordinarios, como brotes focalizados o fallas en los sistemas de salud, que requieren atención inmediata.
- **Monitoreo de intervenciones:** A través del seguimiento de autocorrelaciones, es posible evaluar la eficacia de las medidas implementadas y realizar ajustes en función de los resultados obtenidos.

La incorporación de herramientas como la descomposición de series temporales y el análisis de autocorrelación ha sido clave para profundizar en la dinámica del COVID-19 en los distintos países evaluados. Esta perspectiva permite a BIOGENESYS actuar con anticipación, optimizar la asignación de recursos y diseñar estrategias más efectivas para mejorar la salud pública en América Latina.

3) Relación entre Cobertura de Vacunación y Nuevos Casos Confirmados

Este análisis explora cómo varía la cantidad de nuevos casos de COVID-19 a medida que se incrementa la cobertura de vacunación en cada país. Se busca identificar correlaciones y comportamientos que permitan diseñar estrategias más eficaces de salud pública.

1. Argentina

- **Hallazgos:** Los datos muestran una alta concentración de puntos en el extremo inicial del eje de dosis administradas, con un pico de nuevos casos incluso con una baja cobertura inicial.
- **Interpretación:** No se observa una relación clara entre el avance de la vacunación y la disminución de casos. Esto sugiere la influencia de factores externos como variantes del virus, políticas públicas, o el comportamiento social.

2. Brasil

- **Hallazgos:** A medida que aumentan las dosis administradas, se observa una tendencia a la baja en los nuevos casos, aunque con fluctuaciones y picos esporádicos.
- **Interpretación:** La vacunación parece estar teniendo un efecto positivo, aunque los datos reflejan que otros factores también afectan la evolución de la pandemia.

3. Chile

- **Hallazgos:** Se identifica una disminución general de nuevos casos conforme avanza la vacunación, aunque con picos inesperados incluso en momentos de alta cobertura.
- **Interpretación:** La inmunización juega un papel clave, pero elementos estacionales y la implementación de medidas sanitarias pueden incidir en los resultados.

4. Colombia

- **Hallazgos:** Hay una reducción visible de los casos nuevos conforme avanza la vacunación, aunque con ciertas oscilaciones importantes.
- **Interpretación:** La vacunación es efectiva, pero las fluctuaciones indican la necesidad de mantener un enfoque integral que combine medidas preventivas y vigilancia.

5. México

- **Hallazgos:** Se registra una tendencia descendente en nuevos casos con el incremento en dosis administradas, aunque persisten algunos picos significativos.
- **Interpretación:** Aunque la vacunación contribuye a contener los casos, otros factores — como nuevas variantes o relajación en las medidas— influyen en la aparición de brotes.

6. Perú

- **Hallazgos:** Se observa un comportamiento similar a los otros países, con una disminución general de casos a medida que progresa la vacunación, acompañada de picos aislados.
- **Interpretación:** La vacunación es un factor clave en la reducción de contagios, pero se requiere complementar con otras acciones sanitarias y de prevención.

Aplicación Estratégica para BIOGENESYS

Efectividad de la Vacunación

- El análisis evidencia que la vacunación tiene un impacto positivo en la reducción de nuevos contagios, aunque no suficiente por sí sola. Este hallazgo respalda la necesidad de intensificar la cobertura, especialmente en zonas con alta transmisión.

Planificación y Distribución de Recursos

- Al identificar regiones con baja cobertura vacunal y alta incidencia de casos, BIOGENESYS puede priorizar la asignación de recursos, personal médico y puntos de vacunación.

Monitoreo y Ajuste Dinámico de Estrategias

- La relación variable entre vacunación y casos sugiere la importancia de un seguimiento continuo para adaptar rápidamente las medidas a las dinámicas locales.

Enfoque Multidimensional

- Las fluctuaciones indican que las vacunas deben ir acompañadas de políticas públicas sostenidas: uso de mascarillas, distanciamiento social, vigilancia epidemiológica, y campañas educativas.

Investigación y Desarrollo Continuos

- Comprender los picos de contagio en contextos de alta vacunación es fundamental. Esto abre la puerta a estudios adicionales sobre el comportamiento del virus, la efectividad de las vacunas frente a nuevas variantes, y posibles necesidades de refuerzo o reformulación.

La exploración de la relación entre cobertura vacunal y nuevos casos confirma que, si bien la vacunación es esencial para el control de la pandemia, debe integrarse con medidas complementarias. Para BIOGENESYS, este tipo de análisis proporciona una base sólida para la toma de decisiones estratégicas, la planificación eficiente de recursos y el desarrollo de respuestas adaptativas ante los cambios en la dinámica de la pandemia.

4) Análisis de Variables Clave: Densidad Poblacional, Profesionales de la Salud y Desarrollo Humano

Esta sección examina indicadores relevantes que afectan la capacidad de respuesta sanitaria de los países analizados, incluyendo el Índice de Desarrollo Humano (IDH), la densidad de enfermeras y médicos por cada 1,000 habitantes. Estos datos son esenciales para orientar la estrategia de expansión de BIOGENESYS en América Latina.

1. Índice de Desarrollo Humano (IDH)

- **Hallazgos:**
Todos los países bajo análisis presentan niveles de IDH considerados altos. Argentina y Chile destacan por tener los valores más elevados, mientras que Brasil, Colombia, México y Perú se ubican levemente por debajo, pero aún dentro del rango alto.
- **Interpretación:**
Un IDH elevado generalmente implica una población con mejor acceso a educación, salud y calidad de vida, lo cual constituye una ventaja para la implementación de infraestructuras

como laboratorios y centros de salud. Estos entornos podrían facilitar una mayor aceptación social y cooperación institucional en la instalación de nuevos servicios.

2. Densidad de Enfermeras por Cada 1,000 Habitantes

- **Hallazgos:**
Chile lidera con la mayor densidad de enfermeras, seguido por Brasil. En contraste, Argentina, Colombia, México y Perú muestran niveles considerablemente más bajos.
- **Interpretación:**
La alta densidad de enfermeras sugiere una mayor capacidad operativa frente a emergencias sanitarias. En los países con menor densidad, la instalación de laboratorios por parte de BIOGENESYS podría requerir inversiones adicionales en capacitación de personal y recursos humanos para garantizar una respuesta adecuada.

3. Densidad de Médicos por Cada 1,000 Habitantes

- **Hallazgos:**
Argentina encabeza la lista en cuanto a densidad médica, seguido por Chile y México. Brasil, Colombia y Perú presentan valores más bajos.
- **Interpretación:**
Un mayor número de médicos por habitante indica una capacidad instalada más fuerte para ofrecer servicios de salud. Este contexto favorece la implementación de centros especializados sin necesidad de una inversión inicial elevada en personal médico. En países con menor densidad, será clave fortalecer las redes de atención a través de capacitación, soporte técnico y colaboración institucional.

Implicaciones Estratégicas para BIOGENESYS

1. Planificación de la Expansión

- **Fortalezas Regionales:**
Países como Chile, Argentina y Brasil presentan condiciones sanitarias robustas que podrían facilitar la implementación ágil de laboratorios y centros de vacunación.
- **Intervención en Regiones Vulnerables:**
En naciones con menor densidad de profesionales sanitarios —como Colombia, México y Perú— se recomienda adoptar un enfoque progresivo, orientado a fortalecer capacidades locales y mejorar la infraestructura existente.

2. Optimización de Recursos

- **Asignación Inteligente:**
El uso estratégico de estos datos permitirá distribuir los recursos humanos y materiales de manera más eficiente, enfocándose en las regiones más necesitadas.
- **Formación Continua:**

Invertir en el desarrollo profesional del personal de salud en áreas con escasez puede incrementar significativamente la capacidad de respuesta ante futuras crisis sanitarias.

3. Estrategias de Colaboración

- **Alianzas Institucionales:**

BIOGENESYS puede beneficiarse de convenios con universidades, hospitales y centros de investigación en países con alto IDH, fomentando programas conjuntos de capacitación y transferencia tecnológica.

- **Redes Regionales de Apoyo:**

La creación de programas de intercambio entre países con diferentes niveles de cobertura médica puede facilitar una distribución más equitativa de la carga de trabajo y mejorar la resiliencia del sistema sanitario en toda la región.

La evaluación del IDH, junto con la densidad de médicos y enfermeras, proporciona un marco integral para la toma de decisiones estratégicas. Al priorizar la expansión de infraestructura en zonas con capacidades consolidadas y reforzar las regiones con mayor vulnerabilidad, BIOGENESYS puede maximizar su impacto y contribuir de manera efectiva al fortalecimiento de los sistemas de salud pública en América Latina.

5) Análisis de Correlaciones: Implicaciones Estratégicas para BIOGENESYS

Este análisis se basa en mapas de calor que visualizan la matriz de correlación entre variables epidemiológicas, demográficas y sanitarias clave. La información obtenida permite a BIOGENESYS identificar patrones relevantes para optimizar su intervención en salud pública y priorizar acciones en territorios estratégicos.

Gráfico 1: Mapa de Calor General – Correlaciones entre Variables Clave

Principales Hallazgos

Nuevos Casos Confirmados

- **Correlaciones positivas:**
 - Con nuevos fallecidos (0.52), población total y subgrupos urbanos/rurales (~0.50), dosis de vacunas administradas (0.31) y recuperaciones acumuladas (0.43).
- **Correlaciones negativas:**
 - Con densidad poblacional (-0.14), lo que sugiere que la propagación del virus no depende exclusivamente del número de personas por área.

Nuevos Fallecidos

- **Correlaciones positivas:**
 - Con nuevos casos confirmados (0.52) y variables demográficas (población general y urbana).

- **Correlaciones negativas:**
 - Mínima con la densidad poblacional (-0.0062), reflejando una influencia marginal.

Dosis de Vacunas Administradas

- **Correlaciones positivas:**
 - Con población total y sus subgrupos, nuevos casos y recuperaciones.
- **Correlaciones negativas:**
 - Con densidad poblacional (-0.13), lo que puede deberse a desafíos logísticos en áreas densamente pobladas o urbanas.

Gráfico 2: Mapa de Calor Filtrado – Correlaciones Significativas (> 0.5)

Principales Hallazgos

Correlaciones Positivas Relevantes

- **Entre subgrupos de población** (masculina, femenina, urbana, rural): extremadamente altas (>0.9), como se espera por su composición demográfica.
- **Entre recuperaciones acumuladas y disponibilidad de médicos** (con un coeficiente >0.5): lo que refuerza la idea de que una mayor capacidad médica mejora significativamente los resultados clínicos.

Correlaciones Negativas Relevantes

- **Entre el Índice de Desarrollo Humano (IDH) y población rural (-0.62):** sugiere que un mayor desarrollo se asocia con mayor urbanización.
- **Entre el IDH y la densidad poblacional (-0.52):** puede reflejar que las zonas más densamente pobladas enfrentan más desafíos estructurales, afectando su nivel de desarrollo.

Utilidad Estratégica para BIOGENESYS

1. Priorización de Intervenciones Sanitarias

- Las altas correlaciones entre nuevos casos y variables demográficas indican que las regiones con mayor población —especialmente urbana— deben ser priorizadas para campañas intensivas de vacunación, tratamiento y diagnóstico.
- El seguimiento continuo de las recuperaciones, junto con los niveles de vacunación, permitirá evaluar en tiempo real la efectividad de las medidas implementadas.

2. Optimización de Recursos

- Dada la correlación entre el número de médicos y la recuperación, es crucial reforzar la capacidad médica en zonas con menor densidad de profesionales de salud.

- Las áreas con mayor población deben ser el foco de estrategias logísticas eficientes para la distribución de vacunas, equipos y personal sanitario.

3. Planificación Territorial Basada en el IDH

- La relación negativa del IDH con la población rural y la densidad sugiere que estas zonas presentan mayores vulnerabilidades estructurales.
- BIOGENESYS podría enfocar sus esfuerzos en fortalecer la infraestructura sanitaria y educativa en regiones con bajo IDH para elevar su resiliencia a futuras crisis de salud.

4. Sostenibilidad y Equidad

- El análisis de correlaciones permite diseñar planes de intervención más equitativos, priorizando no solo los territorios con mayor impacto actual, sino también aquellos que, por sus condiciones estructurales, son más propensos a sufrir en el futuro.

La matriz de correlación ofrece una visión integral de cómo interactúan variables críticas en el contexto de la pandemia. Para BIOGENESYS, esta herramienta representa una base sólida para la toma de decisiones basadas en datos. La implementación estratégica de recursos en función de estas correlaciones permitirá mejorar la capacidad de respuesta sanitaria, optimizar la expansión de servicios, y avanzar hacia un modelo de intervención sostenible y adaptativo.

6) Análisis del Dashboard: Navegación, Visualización y Aplicaciones Estratégicas

El dashboard desarrollado para BIOGENESYS constituye una herramienta interactiva fundamental para la toma de decisiones estratégicas en torno a la expansión de laboratorios y centros de vacunación contra el COVID-19. A través de visualizaciones dinámicas y estructuradas, permite explorar los datos clave de salud pública y demografía a nivel regional, facilitando una planificación informada y focalizada.

Navegación del Dashboard

1. Portada

- **Sección Principal:** Introducción al objetivo del dashboard, centrado en guiar la expansión estratégica de la infraestructura sanitaria de BIOGENESYS.
- **Botones de Navegación:**
 - **PORTADA:** Regresa a la página inicial.
 - **INFO GENERAL:** Acceso a los principales indicadores epidemiológicos y de vacunación.
 - **INFRAESTRUCTURA:** Visualización de recursos sanitarios y características demográficas por país.

2. INFO GENERAL: Impacto Epidemiológico y Vacunación

Visualizaciones Clave:

- **Tasa de Mortalidad, Vacunación y Nuevos Casos:** Brinda una visión comparativa entre países sobre el impacto del COVID-19.
- **Gráficos de Barras y Torta:**
 - Nuevos fallecimientos por país.
 - Comparación entre nuevos casos confirmados y fallecidos.
 - Porcentaje de población vacunada.
 - Nuevos casos confirmados.
 - Distribución etaria: población menor y mayor de 50 años.

Estas representaciones permiten identificar de manera clara las regiones con mayor carga sanitaria y eficacia en la cobertura de vacunación.

3. INFRAESTRUCTURA: Recursos Sanitarios y Variables Demográficas

Indicadores Relevantes:

- **Proporción de Enfermeras y Médicos por cada 1,000 habitantes.**
- **Índice de Desarrollo Humano (IDH)** por país.

Gráficos Destacados:

- **Barras Comparativas:** Cantidad de enfermeras y médicos por país.
- **Diagramas de Caja:** Distribución de población rural vs. urbana.
- **Comparación de Mortalidad en Menores de 10 y Mayores de 70 Años:** Indicador clave de vulnerabilidad etaria.
- **Gráfico de Dispersión:**
 - Densidad poblacional vs. cantidad de médicos.
 - Correlaciones entre variables sanitarias y demográficas.

Estos gráficos facilitan la detección de patrones críticos para la planificación estratégica de BIOGENESYS.

Conclusiones Clave Derivadas del Dashboard

1. Identificación de Zonas Críticas

- **Perú y Chile** presentan altos niveles de nuevos fallecimientos, lo cual resalta la urgencia de reforzar su infraestructura sanitaria.
- **Argentina** reporta un elevado número de casos confirmados, lo que indica la necesidad de intensificar acciones de prevención y control epidemiológico.

2. Distribución de Recursos

- **Vacunación:**
 - Chile lidera en porcentaje de población vacunada.
 - Colombia y Argentina muestran rezagos, señalando una distribución desigual y áreas prioritarias para intervención.
- **Personal Médico:**
 - La disponibilidad de médicos y enfermeras varía significativamente.
 - Perú, por ejemplo, destaca por una proporción más alta, lo que puede reflejar una mejor capacidad de respuesta sanitaria.

3. Impacto Demográfico

- **Población Mayor de 70 Años:**
 - Elevadas tasas de mortalidad en esta franja etaria indican la necesidad de estrategias de protección específicas.
- **Densidad Poblacional:**
 - Zonas densamente pobladas con escaso personal médico deben ser priorizadas para instalación de laboratorios y centros de atención.

4. Toma de Decisiones Estratégicas

- **Infraestructura Sanitaria:**
 - La combinación de IDH, densidad poblacional y personal de salud proporciona una base sólida para definir dónde invertir con mayor impacto.
- **Patrones de Vacunación:**
 - Analizar la relación entre tasas de vacunación y disminución de casos permite mejorar la focalización de campañas futuras.

Valor Estratégico del Dashboard

El dashboard brinda a BIOGENESYS una **visión integral, interactiva y basada en datos reales**, lo cual:

- Agiliza la **identificación de zonas vulnerables**.
- Facilita una **distribución eficiente de recursos**.
- Permite diseñar **intervenciones específicas según el perfil demográfico y epidemiológico** de cada país.

Esta herramienta se convierte así en un soporte esencial para **garantizar una expansión efectiva, equitativa y sostenible** de los laboratorios y centros de vacunación de BIOGENESYS en América Latina.

CONCLUSIONES Y RECOMENDACIONES

El análisis realizado a través de las visualizaciones y del dashboard interactivo en Power BI permitió obtener una comprensión profunda y estructurada de la situación sanitaria y demográfica en seis países de América Latina: **Argentina, Brasil, Chile, Colombia, México y Perú**. A través de indicadores clave como incidencia de COVID-19, tasas de vacunación, densidad de personal médico, distribución etaria y desarrollo humano, se identificaron los principales desafíos y oportunidades para la expansión de laboratorios y centros de vacunación por parte de BIOGENESYS.

Visualizaciones Clave Analizadas

- **Nuevos Fallecimientos por País:** Perú y Chile presentan las tasas más altas, lo cual evidencia la urgencia de reforzar la infraestructura sanitaria en estas regiones.
- **Tasa de Vacunación:** Chile lidera en vacunación, mientras que Colombia y Argentina presentan porcentajes más bajos, señalando desigualdades en la distribución y efectividad de las campañas.
- **Distribución de Población por Edad:** Si bien una gran parte de la población es menor de 50 años, la mortalidad es considerable en personas mayores de 70 años, lo cual indica una alta vulnerabilidad de este grupo.
- **Densidad de Personal Médico:** Existen diferencias significativas. Perú, por ejemplo, tiene una mayor proporción de personal médico por habitante, lo cual podría estar asociado a una mejor preparación en términos de respuesta sanitaria.
- **Densidad Poblacional vs. Cantidad de Médicos:** Las zonas con alta densidad poblacional pero baja proporción de médicos requieren atención prioritaria para ampliar la cobertura y capacidad de respuesta.

Conclusiones Estratégicas

1. Políticas de Vacunación por País

- **Chile:** Fuerte cobertura; mantener y expandir las estrategias hacia zonas rurales.
- **Argentina y Colombia:** Reforzar la logística y eficiencia de las campañas de vacunación para mejorar la cobertura.

2. Inversión en Salud Pública

- **Perú y Chile:** Priorizar la inversión en infraestructura sanitaria para reducir la mortalidad y ampliar la cobertura.
- **Brasil:** Requiere fortalecimiento en políticas sanitarias y aumento del personal médico capacitado.

3. Dinámica Poblacional y Urbanización

- Las ciudades densamente pobladas de **Brasil, México y Colombia** (São Paulo, Ciudad de México, Bogotá) demandan mayor presencia de centros de salud y vacunación para controlar la propagación del virus.

4. Factores Socioeconómicos

- **Argentina y Perú** enfrentan limitaciones económicas que repercuten en la salud pública. Se recomienda implementar políticas de apoyo financiero para mejorar la capacidad del sistema sanitario.

Recomendación Principal: Ubicación Estratégica para Nuevos Laboratorios BIOGENESYS

A partir del análisis realizado, se identifican las siguientes ubicaciones prioritarias para la expansión de laboratorios y centros de vacunación:

- **Perú:** Instalación urgente en **Lima** y otras zonas con alta mortalidad.
- **Chile:** Expansión en **zonas rurales** para mejorar el alcance de la vacunación.
- **Brasil:** Apertura de centros en **São Paulo** y **Río de Janeiro** para responder a la alta densidad e incidencia.
- **Argentina y Colombia:** Invertir en **Buenos Aires** y **Bogotá**, respectivamente, como centros clave de atención y distribución de vacunas.

El análisis de datos realizado con apoyo de Power BI facilitó la identificación de **necesidades diferenciadas** en los países analizados. Estas conclusiones permiten a BIOGENESYS **dirigir estratégicamente sus esfuerzos** hacia regiones con mayor impacto sanitario, menor cobertura médica o desafíos logísticos más marcados.

Tomar decisiones **basadas en datos confiables y visualizaciones claras** es esencial para fortalecer la salud pública, optimizar los recursos disponibles y prepararse de manera más efectiva frente a futuras emergencias sanitarias.

REFLEXIÓN PERSONAL

Este proyecto ha sido una experiencia profundamente enriquecedora, permitiéndome aplicar y fortalecer mis habilidades como analista de datos en un contexto de gran relevancia como lo es la salud pública. A través del análisis del impacto del COVID-19 y la evaluación de la infraestructura sanitaria en distintos países de América Latina, desarrollé una visión integral sobre cómo los datos pueden respaldar decisiones estratégicas para la expansión de laboratorios y centros de vacunación.

Habilidades Desarrolladas

1. Análisis Exploratorio de Datos (EDA): Aprendí a realizar un análisis profundo y metódico de las variables clave, identificando patrones, tendencias y anomalías relevantes. El uso de técnicas avanzadas con librerías como *pandas* y *numpy* me permitió manipular y analizar datos de series temporales con eficiencia.

2. Visualización de Datos: Desarrollé competencias en la creación de visualizaciones efectivas utilizando *Power BI*, *matplotlib* y *seaborn*. Implementé diversos tipos de gráficos —barras, tortas, diagramas de caja, dispersión, líneas, mapas de calor— para comunicar de manera clara y persuasiva los hallazgos del análisis.

3. Limpieza y Preparación de Datos: Adquirí experiencia en la transformación y normalización de conjuntos de datos grandes y complejos, aplicando técnicas de imputación de valores faltantes y tratamiento de outliers. Esta etapa fue crítica para garantizar la calidad de los datos analizados.

4. Desarrollo de Dashboards Interactivos: Creé dashboards dinámicos y funcionales en Power BI que facilitaron la interpretación de datos por parte de stakeholders no técnicos. Utilicé medidas y columnas calculadas para entregar insights relevantes y enfocados en la toma de decisiones.

5. Scripting en Python e Integración con BI: Incorporé scripts en Python dentro de Power BI para realizar análisis más avanzados, como correlaciones y descomposición de series temporales, lo cual enriqueció la profundidad del análisis y las visualizaciones personalizadas.

6. Documentación del Proceso: Reconocí la importancia de documentar cada etapa del proceso analítico, desde la limpieza de datos hasta la toma de decisiones, lo que permite mayor trazabilidad, revisión y replicabilidad del proyecto.

Reflexión Final

Si tuviera que comenzar este proyecto nuevamente, mantendría muchos de los enfoques aplicados, pero también realizaría ciertos ajustes clave para optimizar el proceso:

Aspectos que Mantendría:

- Un **análisis exploratorio exhaustivo** al inicio, que permitió comprender a fondo los datos y formular hipótesis fundamentadas.
- La **integración entre Power BI y Python**, que resultó ser una combinación poderosa para el análisis técnico y visual.
- La **iteración constante** de visualizaciones y métricas para refinar la calidad del dashboard y mejorar la comunicación de los hallazgos.

Cambios que Consideraría:

- **Mayor planificación inicial**, definiendo con más precisión los objetivos, cronograma y etapas del proyecto.
- **Automatización del proceso de limpieza de datos**, con scripts más robustos desde el inicio, para ahorrar tiempo y asegurar la consistencia del dataset.
- **Mayor énfasis en la documentación técnica**, incluyendo justificaciones detalladas para cada decisión tomada durante el análisis.

Este proyecto me permitió consolidar conocimientos técnicos y estratégicos, y reafirmó la importancia de una mirada crítica y estructurada al trabajar con datos. La posibilidad de transformar información compleja en herramientas de decisión reales y concretas ha sido una de las experiencias más gratificantes de mi formación. Me siento mejor preparada para abordar proyectos futuros con una visión analítica, rigurosa y colaborativa.

EXTRA CREDIT – MEJORAS IMPLEMENTADAS Y RECOMENDACIONES PARA FUTUROS PROYECTOS

Durante el desarrollo de este proyecto, se identificaron diversas oportunidades de mejora que permitieron optimizar los procesos analíticos, elevar la calidad de los resultados y facilitar la toma de decisiones estratégicas. A continuación, se detallan las principales **mejoras implementadas** y las **sugerencias clave** para futuros proyectos similares.

Mejoras Implementadas

1. Automatización del Proceso de Limpieza de Datos

- **Script de Limpieza Automática:** Se desarrolló un script en Python para automatizar la limpieza y preparación del dataset. Esta automatización redujo considerablemente el tiempo requerido para esta etapa, permitiendo centrarse en el análisis.
- **Estandarización de Formatos:** Se implementaron reglas uniformes para estandarizar los formatos de fechas, nombres de columnas y valores categóricos, facilitando la integración de múltiples fuentes de datos y reduciendo errores.

2. Visualizaciones Mejoradas

- **Gráficos Interactivos en Power BI:** Se diseñaron visualizaciones dinámicas que permiten al usuario filtrar y explorar los datos desde diferentes ángulos, obteniendo insights más precisos.
- **Uso de Paletas de Colores Consistentes:** Se aplicó un diseño visual uniforme, con colores estratégicamente seleccionados para destacar la información crítica y mejorar la legibilidad.

3. Análisis de Correlación Avanzado

- **Submáscara en Mapas de Calor:** Se aplicaron submáscaras para resaltar únicamente las correlaciones estadísticamente más relevantes, facilitando la identificación de relaciones clave entre variables.
- **Descomposición de Series Temporales:** Esta técnica permitió analizar las variables a lo largo del tiempo, separando componentes de tendencia, estacionalidad y ruido, para una mejor interpretación de los patrones.

4. Medidas y Columnas Calculadas Específicas

- **Medidas Clave en Power BI:** Se desarrollaron medidas personalizadas para calcular tasas de vacunación, mortalidad y cobertura sanitaria, brindando indicadores claros para la toma de decisiones.
- **Columnas Calculadas:** Se incorporaron indicadores derivados como densidad poblacional y personal médico por cada 1.000 habitantes, que enriquecieron el análisis multidimensional.

Sugerencias para Futuros Proyectos

1. Mejora en la Planificación del Proyecto

- **Definición de Objetivos y Etapas:** Es recomendable dedicar más tiempo al diseño inicial del proyecto, definiendo claramente los objetivos, fases, entregables y responsables, para asegurar una ejecución más ordenada y eficiente.

2. Optimización del Rendimiento en Power BI

- **Uso de Agregaciones y Filtrados Eficientes:** Para proyectos con grandes volúmenes de datos, se sugiere aplicar técnicas de modelado y filtrado avanzado que optimicen el tiempo de carga y respuesta del dashboard.

3. Documentación Exhaustiva del Proceso

- **Registro Detallado de Decisiones:** Llevar un seguimiento continuo de las decisiones tomadas, los criterios aplicados y los obstáculos enfrentados facilita el trabajo colaborativo, la revisión del proyecto y futuras iteraciones.

4. Capacitación y Desarrollo Continuo

- **Formación en Herramientas Avanzadas:** Se recomienda seguir invirtiendo en el aprendizaje de nuevas herramientas, técnicas estadísticas y de visualización que fortalezcan el perfil analítico y permitan abordar retos más complejos.

Las mejoras implementadas durante el desarrollo de este proyecto incrementaron significativamente la precisión, eficiencia y valor del análisis realizado. Gracias a estas optimizaciones, BIOGENESYS cuenta ahora con una herramienta poderosa para tomar decisiones estratégicas basadas en datos concretos y visualmente comprensibles. Al adoptar un enfoque de mejora continua, estructurado y basado en buenas prácticas de análisis de datos, se sientan las bases para enfrentar con éxito futuros desafíos en contextos similares, garantizando análisis más robustos y recomendaciones más acertadas.