

Expansión estratégica de Biogenesys con Python

Nombre del autor: Ruben Hernán Pared

Email: hernan2014@hotmail.com

Cohorte: DA-PT01

Fecha de entrega: 29-07

Institución: Laboratorios Biogenesys



Introducción

Propósito del Proyecto:

El proyecto tuvo como objetivo proporcionar a BIOGENESYS una herramienta analítica en Power BI para facilitar la toma de decisiones estratégicas en la expansión de laboratorios y centros de vacunación en América Latina. Esto se logró mediante la visualización de datos críticos de salud, demografía y socioeconomía.

Objetivos Organizacionales Alcanzados:

- Identificación de Ubicaciones Óptimas para la Expansión de Laboratorios
- Infraestructura Médica: Evaluación de la disponibilidad de médicos y enfermeras por cada 1,000 habitantes.
- Capacidad Hospitalaria: Análisis de la infraestructura hospitalaria para determinar la idoneidad de cada región.
- Optimización de Políticas de Vacunación
- Cobertura de Vacunación y Reducción de Casos: Visualización de la relación entre tasas de vacunación y disminución de nuevos casos y fallecimientos.
- Efectividad de Vacunación por País: Identificación de las estrategias de vacunación más efectivas.
- Inversión en Salud y Recursos Humanos:
- Distribución de Recursos Sanitarios: Decisiones informadas sobre la inversión en personal médico.
- Problemas Sociales y Económicos: Análisis de indicadores socioeconómicos para comprender los problemas que afectan la implementación de nuevas infraestructuras.
- Análisis Demográfico y Epidemiológico:
- Población por Edad: Identificación de regiones con mayor vulnerabilidad y necesidad de atención médica.
- Tendencias y Patrones Temporales: Visualización de la evolución de nuevos casos y fallecimientos para una mejor planificación.

El proyecto cumplió con los objetivos de BIOGENESYS, proporcionando una herramienta robusta para la toma de decisiones estratégicas. El dashboard en Power BI permitió identificar las mejores ubicaciones para la expansión de laboratorios y centros de vacunación, optimizar políticas de vacunación e invertir eficientemente en salud y recursos humanos.

Desarrollo del proyecto

Recopilación de Datos:

Fuentes de Datos:

Los datos se obtuvieron de un Dataset provisto por el laboratorio

Métodos de Recopilación:

- **API:** Utilización de APIs públicas para acceder a datos en tiempo real, asegurando que la información sea actualizada y precisa.
- **Descarga de Archivos:** Obtención de conjuntos de datos en formato CSV, Excel y JSON desde portales de datos abiertos y sitios web oficiales.
- **Web Scraping:** Implementación de técnicas de scraping para extraer datos de páginas web cuando no estaban disponibles para descarga directa.

Selección de Datos:

1. Relevancia:

- **Variables Clave:** Se seleccionaron variables directamente relacionadas con los objetivos del estudio, como nuevos casos confirmados, nuevos fallecimientos, dosis de vacunas administradas, población total y personal médico por cada 1,000 habitantes.
- **Indicadores Socioeconómicos:** Se incluyeron variables como el Índice de Desarrollo Humano (IDH) y el Producto Interno Bruto (PIB) per cápita para analizar el contexto socioeconómico de cada país.

2. Calidad de los Datos:

- **Veracidad y Confiabilidad:** Se priorizaron fuentes oficiales y reconocidas para garantizar la precisión y la confiabilidad de los datos.
- **Compleción:** Se prefirieron conjuntos de datos completos que no presentaran valores faltantes significativos, facilitando el análisis posterior.

3. Temporalidad:

- **Actualidad:** Se aseguraron de que los datos estuvieran actualizados hasta la fecha más reciente posible, proporcionando una visión actual del panorama de salud y demografía.
- **Historial:** En algunos casos, se consideraron datos históricos para análisis de tendencias y proyecciones temporales.

Transformaciones y Limpieza de Datos:

- **Limpieza de Datos:**

Módulo 4

- **Eliminación de Duplicados:** Se identificaron y eliminaron registros duplicados para evitar redundancias y sesgos en el análisis.
- **Manejo de Valores Faltantes:** Los valores faltantes se abordaron mediante técnicas como la imputación (rellenando con valores promedio o medianos) o la eliminación de registros cuando los datos faltantes eran significativos y afectaban la integridad del análisis.
- **Corrección de Errores:** Se realizaron correcciones manuales y automáticas de errores tipográficos y valores anómalos que no eran coherentes con el resto del conjunto de datos.

Transformaciones de Datos:

- **Normalización y Escalamiento:** Se aplicaron técnicas de normalización y escalamiento para estandarizar las variables y facilitar las comparaciones entre diferentes conjuntos de datos.
- **Creación de Columnas Calculadas:** Se generaron columnas calculadas para obtener métricas específicas como la tasa de vacunación, la tasa de mortalidad y la densidad poblacional, proporcionando insights más detallados y relevantes.
- **Conversión de Tipos de Datos:** Se aseguraron de que todas las variables estuvieran en el formato correcto (por ejemplo, fechas en formato datetime, números en formato float o int) para permitir un análisis eficiente y preciso.

El proceso de recopilación, selección, limpieza y transformación de datos fue esencial para asegurar la calidad y relevancia del análisis realizado. La metodología aplicada garantizó que los datos fueran precisos, actualizados y pertinentes para los objetivos del proyecto. Las transformaciones realizadas permitieron generar visualizaciones y métricas que ofrecen una visión clara y accionable para BIOGENESYS, facilitando la toma de decisiones estratégicas informadas. La limpieza de datos y las transformaciones adecuadas son fundamentales para cualquier proyecto de análisis de datos, ya que proporcionan una base sólida sobre la cual se pueden construir análisis más complejos y valiosos.

EDA e insights

1) **Análisis de Casos Activos Vs Casos Recuperados:**

1. Evaluación de la Propagación y Control del Virus

- **Casos Activos:** Indican la cantidad de personas que actualmente están enfermas y pueden potencialmente transmitir el virus. Un alto número de casos activos puede señalar regiones donde la pandemia no está bajo control.

- **Casos Recuperados:** Indican la cantidad de personas que han superado la enfermedad, lo cual puede ser un indicador del nivel de inmunidad en la población y la efectividad del sistema de salud.

2. Planificación de Recursos y Personal Sanitario

- **Distribución de Personal:** Conocer las regiones con altos casos activos ayuda a determinar dónde se necesita más personal médico y recursos, como camas de hospital y equipos de protección personal (EPP).

Módulo 4

- Centros de Vacunación: Las áreas con altos casos activos pueden necesitar más centros de vacunación para acelerar la inmunización y controlar el brote.

3. Decisión sobre Ubicaciones de Nuevos Laboratorios

- Foco en Áreas Críticas: BIOGENESYS puede priorizar la apertura de laboratorios en regiones con altos casos activos para investigar nuevas variantes del virus y desarrollar tratamientos específicos.

- Seguimiento y Prevención: Regiones con un alto número de recuperados, pero también con casos activos pueden necesitar laboratorios enfocados en el seguimiento a largo plazo y la prevención de rebrotes.

4. Estrategias de Vacunación

- Inmunización Focalizada: Áreas con muchos casos activos y bajos casos recuperados pueden indicar la necesidad de campañas de vacunación más intensivas y focalizadas.

- Evaluación de la Efectividad: Comparar la recuperación en áreas vacunadas vs. no vacunadas puede ayudar a evaluar la efectividad de las campañas de vacunación actuales y ajustar estrategias según sea necesario.

5. Impacto Económico y Logístico

- Gestión de Suministros: Comprender las dinámicas de casos activos y recuperados puede ayudar en la planificación logística de suministros médicos y vacunas, asegurando que las regiones más afectadas reciban lo necesario.

- Planeación a Largo Plazo: Proyecciones basadas en la relación entre casos activos y recuperados pueden informar decisiones sobre la expansión a largo plazo y la sostenibilidad de nuevas instalaciones.

Teniendo en cuenta el análisis realizado se obtuvieron los siguientes resultados:

Interpretación de las Gráficas

1. **Argentina**

- Casos Activos: Los casos activos muestran varias olas, con un pico significativo alrededor de diciembre de 2021 y enero de 2022.
- Casos Recuperados: Se mantienen constantes y bajos en comparación con los casos activos.
- Interpretación: Un alto número de casos activos en comparación con los recuperados sugiere que la pandemia no está completamente bajo control.

2. **Brasil**

- Casos Activos: Similar a Argentina, Brasil muestra varias olas, pero con un patrón más fluctuante.
- Casos Recuperados: Los casos recuperados siguen un patrón similar al de los casos activos, lo que sugiere una mejor tasa de recuperación.
- Interpretación: Aunque hay fluctuaciones, la similitud entre casos activos y recuperados puede indicar una capacidad de respuesta más efectiva del sistema de salud.

Módulo 4

3. Chile

- Casos Activos: Se observa un pico significativo a principios de 2022.
- Casos Recuperados: Se mantienen constantes y bajos en comparación con los casos activos.
- Interpretación: Similar a Argentina, sugiere que la pandemia no está completamente bajo control, y podría necesitar más recursos.

4. Colombia

- Casos Activos: Varias olas con picos notables a lo largo del tiempo.
- Casos Recuperados: Hay una mejor correspondencia entre casos activos y recuperados, similar a Brasil.
- Interpretación: La similitud sugiere un buen sistema de recuperación, pero las olas continuas indican que la pandemia sigue siendo una preocupación.

5. México

- Casos Activos: Varias olas con un pico significativo a principios de 2022.
- Casos Recuperados: Se mantienen constantes y bajos en comparación con los casos activos.
- Interpretación: Necesita mejorar la capacidad de recuperación y control de la pandemia.

6. Perú

- Casos Activos: Similar a otros países con picos notables.
- Casos Recuperados: Se mantienen constantes y bajos.
- Interpretación: Similar a México y Argentina, sugiere la necesidad de una mejor gestión de la pandemia.

En base a los análisis anteriores podemos deducir lo siguiente:

Con respecto a las necesidades de recursos, en las regiones donde los casos activos son elevados en comparación con los casos recuperados como Argentina, Chile, Colombia, México, se pueden requerir mayores recursos médicos, vacunas y personal sanitario. La baja recuperación de los pacientes puede señalar la necesidad de mayores campañas de vacunación y apertura de más centros de tratamientos.

Las zonas con altos casos activos y baja recuperación como las analizadas anteriormente, deben ser tratadas como prioritarias para la instalación de laboratorios y centros de vacunación.

La comprensión de la relación entre casos activos y recuperados es crucial para BIOGENESYS. Proporciona información vital sobre el control de la pandemia, la efectividad de las medidas de salud pública y la necesidad de recursos adicionales. Al utilizar esta información, BIOGENESYS puede optimizar su estrategia de expansión y contribuir de manera efectiva a la mejora de la salud pública en América Latina.

Módulo 4**2) Análisis de Descomposición de Series Temporales y Autocorrelación****Gráficas de Descomposición de Series Temporales**

Para cada país, las gráficas muestran la descomposición de la serie temporal de casos confirmados de COVID-19 en tres componentes principales: tendencia (Trend), estacionalidad (Seasonal) y residuales (Resid).

1) Argentina

- Trend: Muestra un aumento significativo en casos alrededor de diciembre de 2021 y enero de 2022.
- Seasonal: Patrón estacional con varios picos recurrentes.
- Resid: Variaciones y anomalías que no se explican por la tendencia ni la estacionalidad.

2) Brasil

- Trend: Fluctuaciones con picos notables similares a Argentina, pero con variaciones más pronunciadas.
- Seasonal: Patrones estacionales regulares con picos significativos.
- Resid: Variaciones y anomalías como en Argentina, pero con mayor amplitud.

3) Chile

- Trend: Similar a Argentina y Brasil, con un aumento en diciembre de 2021.
- Seasonal: Patrones estacionales claros.
- Resid: Anomalías y variaciones no explicadas por otros componentes.

4) Colombia

- Trend: Fluctuaciones con varios picos, indicando diferentes olas de la pandemia.
- Seasonal: Patrones estacionales regulares.
- Resid: Variaciones y anomalías similares a los otros países.

5) México

- Trend: Variaciones con picos significativos a finales de 2021 y principios de 2022.
- Seasonal: Patrones estacionales claros.
- Resid: Anomalías y variaciones como en otros países.

6) Perú

- Trend: Picos significativos a finales de 2021.
- Seasonal: Patrones estacionales regulares.
- Resid: Variaciones y anomalías no explicadas por la tendencia ni la estacionalidad.

Gráficas de Autocorrelación

Las gráficas de autocorrelación (ACF) muestran cómo los valores de una serie temporal están correlacionados consigo mismos en diferentes retardos.

1) Argentina

La autocorrelación disminuye rápidamente, con ciclos que sugieren un patrón estacional.

2) Brasil

Módulo 4

Muestra una disminución inicial seguida de un patrón cíclico, indicando componentes estacionales.

3) Chile

Similar a Argentina, con una disminución rápida y ciclos sugerentes de estacionalidad.

4) Colombia

Disminución inicial seguida de un patrón cíclico, similar a Brasil.

5) México

Patrón cíclico claro después de una disminución inicial, indicando estacionalidad.

6) Perú

Disminución rápida seguida de ciclos, similar a otros países.

Utilidad para BIOGENESYS

Identificación de Patrones Estacionales

- Estrategias de Respuesta: Comprender los patrones estacionales ayuda a BIOGENESYS a anticipar olas de infecciones y planificar la distribución de recursos y personal médico de manera eficiente.
- Campañas de Vacunación: Ajustar el tiempo de las campañas de vacunación según los picos estacionales puede maximizar su efectividad.

Análisis de Tendencias

- Planificación a Largo Plazo: La identificación de tendencias permite a BIOGENESYS planificar la expansión de laboratorios en áreas donde se espera un aumento en casos.
- Prevención: Implementar medidas preventivas en base a tendencias detectadas para mitigar futuros picos de infecciones.

Análisis de Anomalías

- Investigación: Las variaciones no explicadas (residuales) pueden indicar eventos específicos que requieren investigación adicional, como brotes localizados o problemas en la infraestructura sanitaria.
- Respuesta Rápida: Detectar anomalías a tiempo permite a BIOGENESYS responder rápidamente a brotes inesperados.

Evaluación de la Efectividad

- Monitoreo Continuo: Utilizar autocorrelaciones para monitorear la efectividad de intervenciones a lo largo del tiempo y ajustar estrategias según los resultados observados.

Las descomposiciones de series temporales y los análisis de autocorrelación proporcionan a BIOGENESYS una visión detallada de las dinámicas de COVID-19 en diferentes países. Esta información es crucial para la toma de decisiones estratégicas, planificación de recursos y desarrollo de intervenciones efectivas para controlar la pandemia y mejorar la salud pública en América Latina.

Módulo 4

3) Análisis de la Relación entre Cobertura de Vacunación y Nuevos Casos Confirmados

1. Argentina

- Observaciones: La mayoría de los puntos están concentrados al principio del eje horizontal (dosis administradas) y hay un pico de nuevos casos confirmados incluso con pocas dosis administradas.
- Interpretación: La correlación entre el número de dosis administradas y la disminución de nuevos casos no es clara. Picos altos de casos nuevos pueden deberse a otras variables como variantes del virus, medidas de control, y comportamiento social.

2. Brasil

- Observaciones: Muestra una disminución en nuevos casos confirmados con el aumento de dosis administradas, aunque hay fluctuaciones significativas y algunos picos altos.
- Interpretación: La vacunación parece tener un impacto en la reducción de nuevos casos confirmados, pero las fluctuaciones indican que otros factores también juegan un papel importante.

3. Chile

- Observaciones: Hay un patrón similar de disminución de nuevos casos confirmados con el aumento de la cobertura de vacunación, aunque con picos inesperados incluso con alta vacunación.
- Interpretación: La vacunación ayuda a reducir los casos, pero otros factores como la temporada del año y las medidas de salud pública pueden afectar los resultados.

4. Colombia

- Observaciones: Disminución visible en nuevos casos confirmados con el aumento de dosis administradas, aunque también con picos y fluctuaciones significativas.
- Interpretación: La vacunación tiene un impacto positivo, pero las fluctuaciones sugieren que es necesario un enfoque integral que incluya medidas adicionales de control y prevención.

5. México

- Observaciones: Patrón de disminución en nuevos casos confirmados con el aumento de dosis administradas, pero con algunos picos altos.
- Interpretación: La vacunación parece reducir los casos confirmados, pero la presencia de picos indica que se deben considerar otros factores como variantes del virus y comportamiento de la población.

6. Perú

- Observaciones: Patrón similar de disminución en nuevos casos confirmados con el aumento de dosis administradas, con algunos picos altos.
- Interpretación: La vacunación es efectiva para reducir los casos, pero se requieren medidas adicionales para manejar las fluctuaciones y picos.

Utilidad para BIOGENESYS

Módulo 4

Efectividad de la Vacunación

Evaluar el Impacto: Los gráficos sugieren que la vacunación tiene un impacto positivo en la reducción de nuevos casos confirmados. Esto respalda la necesidad de expandir la cobertura de vacunación en áreas con altas tasas de infección.

Planificación de Recursos

Distribución Estratégica: Identificar regiones con bajos niveles de vacunación, pero altos casos nuevos puede ayudar a BIOGENESYS a priorizar la distribución de recursos y personal sanitario.

Monitoreo Continuo: Es crucial mantener el monitoreo de la relación entre vacunación y casos nuevos para ajustar estrategias rápidamente según las necesidades emergentes.

Enfoque Integral

Medidas Complementarias: Las fluctuaciones y picos sugieren que, además de la vacunación, se necesitan otras medidas de control y prevención (como el uso de mascarillas, distanciamiento social y restricciones temporales).

Campañas de Concienciación: Promover la importancia de la vacunación y las medidas de salud pública puede aumentar la efectividad general de las estrategias de control de la pandemia.

Investigación y Desarrollo

Estudios Continuos: Realizar estudios continuos para entender mejor las causas de los picos en nuevos casos, incluso con alta vacunación, ayudará a BIOGENESYS a desarrollar intervenciones más específicas y efectivas.

Adaptación de Vacunas: Investigar y adaptar las vacunas para combatir nuevas variantes del virus y mejorar la respuesta inmunológica de la población.

En resumen, la relación entre cobertura de vacunación y nuevos casos confirmados proporciona información valiosa para BIOGENESYS en la toma de decisiones estratégicas, planificación de recursos, y desarrollo de medidas complementarias para controlar la pandemia de manera más efectiva.

4) Análisis de los Datos Relacionados con la Densidad Poblacional, Enfermeras y Médicos por País

1. Índice de Desarrollo Humano (IDH) por País

Observaciones:

Todos los países analizados tienen un IDH relativamente alto.

Argentina y Chile tienen los valores más altos de IDH.

Brasil, Colombia, México y Perú tienen valores ligeramente inferiores, pero aún en un rango alto.

Interpretación:

Un IDH alto generalmente está asociado con mejores condiciones de vida y mayor acceso a servicios de salud.

Módulo 4

Esto sugiere que en estos países, los nuevos laboratorios podrían beneficiarse de una infraestructura sólida y una población con buena educación y salud básica.

2. Densidad de Enfermeras por cada 1,000 Habitantes por País

Observaciones:

Chile tiene la mayor densidad de enfermeras, seguido por Brasil.

Argentina, Colombia, México y Perú tienen una densidad significativamente menor.

Interpretación:

Países con mayor densidad de enfermeras como Chile y Brasil pueden manejar mejor la carga de trabajo relacionada con el COVID-19.

En países con menor densidad de enfermeras, BIOGENESYS podría necesitar invertir más en recursos humanos y formación para asegurar una respuesta efectiva.

3. Densidad de Médicos por cada 1,000 Habitantes por País

Observaciones:

Argentina tiene la mayor densidad de médicos, seguido de Chile y México.

Brasil, Colombia y Perú tienen una densidad menor.

Interpretación:

La alta densidad de médicos en Argentina y Chile sugiere que estos países tienen una buena capacidad para proporcionar atención médica, lo que puede facilitar la implementación de nuevos laboratorios y centros de vacunación.

En países con menor densidad de médicos, es posible que se necesite un enfoque más integrado, incluyendo la capacitación y el apoyo continuo a los profesionales de la salud.

Utilidad para BIOGENESYS

Planificación de la Expansión

Infraestructura Sanitaria: Países con alta densidad de médicos y enfermeras, como Chile, Argentina y Brasil, pueden ser más adecuados para la rápida expansión de laboratorios debido a su infraestructura sanitaria robusta.

Apoyo en Zonas con Menos Recursos: En países como Colombia, México y Perú, donde la densidad de profesionales de la salud es menor, BIOGENESYS puede enfocarse en fortalecer la infraestructura sanitaria y proporcionar capacitación adicional.

Optimización de Recursos

Distribución Eficiente: Utilizar los datos de densidad de médicos y enfermeras para distribuir eficientemente los recursos humanos y materiales en las regiones más necesitadas.

Capacitación y Desarrollo: Invertir en la formación y desarrollo de personal de salud en países con menor densidad para asegurar una respuesta eficaz a futuras crisis de salud.

Estrategias de Colaboración

Módulo 4

Alianzas con Instituciones Locales: Colaborar con universidades y hospitales locales en países con alto IDH para desarrollar programas de investigación y capacitación conjunta.

Programas de Intercambio: Establecer programas de intercambio de profesionales de la salud entre países con alta y baja densidad de médicos y enfermeras para equilibrar la carga de trabajo y mejorar la capacidad de respuesta.

Los gráficos proporcionan información valiosa sobre la capacidad de los diferentes países para manejar la expansión de laboratorios y centros de vacunación. Al analizar la densidad de profesionales de la salud y el IDH, BIOGENESYS puede tomar decisiones informadas sobre dónde y cómo invertir recursos para maximizar el impacto y mejorar la respuesta a las crisis de salud pública.

5) Análisis de la Matriz de Correlación (Mapas de Calor) para los Objetivos de BIOGENESYS

Gráfico 1: Mapa de Calor de Correlaciones entre Variables Clave

New Confirmed (Nuevos Casos Confirmados)

Correlación positiva con:

- new_deceased (0.52): A medida que aumentan los nuevos casos confirmados, también aumentan los fallecidos.
- cumulative_vaccine_doses_administered (0.31): Indica que más casos confirmados pueden llevar a una mayor administración de dosis de vacunas.
- population (0.5), population_male (0.5), population_female (0.5), population_urban (0.49): La población en general y urbana están correlacionadas con nuevos casos confirmados.
- cumulative_recovered (0.43): Más casos confirmados también correlacionan con más recuperaciones.

Correlación negativa con:

- population_density (-0.14): Una mayor densidad poblacional parece correlacionarse con menos nuevos casos confirmados.

New Deceased (Nuevos Fallecidos)

Correlación positiva con:

- new_confirmed (0.52): Como se mencionó, más casos confirmados llevan a más fallecimientos.
- population, population_male, population_female, population_urban: Similar a la correlación con nuevos casos confirmados.
- cumulative_recovered (0.23): Más fallecimientos también correlacionan con más recuperaciones.

Correlación negativa con:

- population_density (-0.0062): Similar a nuevos casos confirmados.

Cumulative Vaccine Doses Administered (Dosis de Vacunas Administradas)

Módulo 4

Correlación positiva con:

- new_confirmed (0.31): Como se mencionó anteriormente.
- population, population_male, population_female, population_urban, cumulative_recovered.

Correlación negativa con:

- population_density (-0.13).
-

Gráfico 2: Mapa de Calor de Correlaciones Filtradas (> 0.5)

Correlaciones Filtradas

Altas correlaciones positivas:

- population_male y population_female (1.0): Alta correlación esperada.
- population con population_male, population_female, population_rural, population_urban: Muestra cómo la población total se distribuye entre diferentes subgrupos.
- cumulative_recovered con population, population_male, population_female, population_rural, population_urban, physicians_per_1000: Más recuperaciones se correlacionan con una mayor población y más médicos.

Correlaciones negativas:

- human_development_index con population_rural (-0.62) y population_density (-0.52): Un mayor desarrollo humano está negativamente correlacionado con la población rural y la densidad poblacional.

Utilidad para BIOGENESYS

Identificación de Áreas Prioritarias para la Vacunación y Tratamiento

Regiones con Alta Correlación de Casos y Población: Las regiones con alta correlación entre nuevos casos confirmados y densidad poblacional pueden ser priorizadas para campañas de vacunación y tratamiento intensivo.

Monitoreo de Recuperaciones: La fuerte correlación entre casos confirmados, dosis de vacunas administradas y recuperaciones sugiere la necesidad de seguimiento continuo de estos indicadores para evaluar la efectividad de las intervenciones.

Desarrollo y Distribución de Recursos

Personal Médico: La alta correlación entre la recuperación y la disponibilidad de médicos indica que aumentar la densidad de profesionales de salud puede mejorar significativamente los resultados en términos de recuperación.

Distribución de Vacunas: La correlación entre la administración de vacunas y la población total sugiere que las áreas densamente pobladas deben ser objetivo de estrategias de distribución eficiente de vacunas.

Planificación Estratégica

Módulo 4

Índice de Desarrollo Humano (IDH): La correlación negativa del IDH con la población rural y la densidad poblacional sugiere que áreas con bajo desarrollo humano y alta densidad pueden requerir más apoyo para mejorar las condiciones de salud pública.

Infraestructura Sanitaria: Invertir en aumentar la densidad de enfermeras y médicos en áreas con baja disponibilidad de estos recursos puede mejorar los resultados de salud y aumentar la capacidad de respuesta ante futuras crisis sanitarias.

El análisis de la matriz de correlación proporciona información valiosa sobre cómo las diferentes variables de salud pública y demográficas están interrelacionadas. Para BIOGENESYS, esta información es crucial para planificar la expansión estratégica de laboratorios y centros de vacunación, optimizar la distribución de recursos, y mejorar la capacidad de respuesta ante la pandemia y futuras emergencias de salud pública.

Análisis del dashboard

Cómo Navegar el Dashboard

- **Portada:**

Sección Principal: Introducción general sobre el propósito del dashboard, centrado en la expansión estratégica de laboratorios para la investigación y vacunación de COVID-19.

- **Botones de Navegación:**

PORTADA: Lleva de vuelta a la página principal.

INFO GENERAL: Presenta información general sobre tasas de mortalidad, vacunación, nuevos casos y densidad poblacional.

INFRAESTRUCTURA: Detalles sobre la infraestructura sanitaria, incluyendo personal médico y datos demográficos.

- **INFO GENERAL:**

Tasa de Mortalidad, Vacunación y Nuevos Casos: Muestra las tasas clave para entender el impacto del COVID-19.

- **Gráficos de Barras y Torta:**

Nuevos Fallecimientos: Comparación de nuevos fallecimientos por país.

Nuevos Casos vs. Fallecidos: Distribución de casos confirmados frente a fallecidos.

% de Vacunación: Porcentaje de población vacunada por país.

Nuevos Confirmados: Nuevos casos confirmados por país.

Distribución de Población por Edad: Comparación entre población menor de 50 años y mayor de 50 años.

- **INFRAESTRUCTURA:**

- **Métricas Clave:**

Nurses / Population: Proporción de enfermeras por cada 1000 habitantes.

Módulo 4

Physicians / Population: Proporción de médicos por cada 1000 habitantes.

IDH: Índice de Desarrollo Humano por país.

- **Gráficos de Barras y Dispersión:**

Cantidad de Enfermeras y Médicos: Comparación de cantidad de personal médico por país.

Población Rural y Urbana: Diagrama de cajas para comparar la población rural y urbana.

Muertes <10 años y >70 años: Comparación de muertes confirmadas en estas poblaciones.

Densidad Poblacional vs. Cantidad de Médicos: Gráfico de dispersión para analizar la relación entre densidad poblacional y cantidad de médicos.

Dispersion de Variables: Gráficos de dispersión de variables clave para identificar patrones y relaciones.

- **Posibles Conclusiones**

Identificación de Zonas Críticas:

Perú y Chile presentan altas tasas de nuevos fallecimientos, lo que sugiere una necesidad urgente de infraestructura adicional en estas áreas.

Argentina muestra un alto número de nuevos casos confirmados, lo que indica la necesidad de mejorar las estrategias de contención y vacunación.

Distribución de Recursos:

Vacunación: Chile tiene el mayor porcentaje de vacunación, mientras que países como Colombia y Argentina están rezagados, sugiriendo una distribución desigual de recursos.

Personal Médico: La densidad de enfermeras y médicos es variable, con algunos países como Perú teniendo una proporción significativamente mayor. Esto puede indicar una mejor preparación sanitaria en ciertas áreas.

Impacto Demográfico:

Población Mayor: Las muertes en poblaciones mayores de 70 años son elevadas en ciertos países, lo que podría indicar una mayor vulnerabilidad de esta población y la necesidad de estrategias de protección dirigidas.

Densidad Poblacional: Las áreas con alta densidad poblacional y baja proporción de personal médico podrían beneficiarse de la expansión de laboratorios y centros de vacunación para mejorar la capacidad de respuesta sanitaria.

Toma de Decisiones Estratégicas:

Infraestructura Sanitaria: La evaluación de la infraestructura sanitaria (personal médico, densidad poblacional, IDH) proporciona una base sólida para decisiones informadas sobre dónde invertir en nuevas instalaciones.

Patrones de Vacunación: Analizar las tasas de vacunación y su relación con la reducción de casos puede ayudar a optimizar las campañas de vacunación.

Módulo 4

El dashboard permite a BIOGENESYS visualizar de manera integral y detallada los datos necesarios para tomar decisiones estratégicas informadas, enfocadas en la expansión efectiva y eficiente de sus laboratorios y centros de vacunación.

Conclusiones y Recomendaciones

El análisis realizado a través de las visualizaciones y el dashboard en Power BI permitió obtener una visión clara y detallada de la situación sanitaria y demográfica en los países analizados (Argentina, Brasil, Chile, Colombia, México y Perú). Las visualizaciones se centraron en aspectos clave como la incidencia de COVID-19, la distribución de vacunas, la densidad de personal médico y la infraestructura sanitaria. Estos elementos fueron esenciales para evaluar la capacidad de cada país para soportar la expansión de laboratorios y centros de vacunación.

Visualizaciones Clave:

Nuevos Fallecimientos por País:

Perú y Chile presentan las tasas más altas de nuevos fallecimientos, indicando una necesidad urgente de reforzar la infraestructura sanitaria.

Tasa de Vacunación:

Chile lidera en porcentaje de vacunación, mientras que Colombia y Argentina muestran rezagos, sugiriendo la necesidad de una mejor distribución de recursos.

Distribución de Población por Edad:

Un alto porcentaje de la población es menor de 50 años en varios países, pero la mortalidad es significativa en poblaciones mayores de 70 años, indicando una mayor vulnerabilidad en esta franja etaria.

Densidad de Personal Médico:

La densidad de enfermeras y médicos varía significativamente entre países, con Perú mostrando una proporción más alta de personal médico, lo que podría reflejar una mejor preparación sanitaria.

Densidad Poblacional vs. Cantidad de Médicos:

Las áreas con alta densidad poblacional y baja proporción de personal médico necesitan urgentemente más infraestructura sanitaria.

Conclusiones Estratégicas

1. Políticas de Vacunación por País:

Chile: Fuerte política de vacunación con altos porcentajes de vacunación. Continuar con las estrategias actuales y expandir la cobertura a áreas menos accesibles.

Argentina y Colombia: Necesitan mejorar la distribución y la eficiencia de las campañas de vacunación para aumentar la cobertura.

2. Inversión en Salud:

Invertir en la infraestructura de salud en Perú y Chile para reducir las tasas de mortalidad y mejorar la respuesta sanitaria.

Módulo 4

Brasil: Reforzar las políticas de salud y aumentar la inversión en personal médico.

3. Habitantes y Densidad:

Las áreas urbanas con alta densidad poblacional en Brasil, México y Colombia requieren más centros de salud y vacunación para manejar mejor la propagación del virus y mejorar la atención médica.

4. Problemas Sociales y Económicos:

Los problemas económicos en Argentina y Perú afectan la capacidad de respuesta sanitaria. Es crucial considerar políticas de apoyo económico para fortalecer el sistema de salud.

Respuesta Principal del Proyecto: Ubicaciones Óptimas para la Expansión de Laboratorios Farmacéuticos

Ubicaciones Óptimas Identificadas:

- Perú: Necesidad urgente de infraestructura sanitaria. Se recomienda instalar laboratorios y centros de vacunación en Lima y áreas con altas tasas de mortalidad.
- Chile: Fortalecer la infraestructura existente y expandir en áreas rurales para mejorar la cobertura de vacunación.
- Brasil: Focalizar en São Paulo y Río de Janeiro, donde la alta densidad poblacional y las tasas de infección justifican la expansión.
- Argentina y Colombia: Incrementar la infraestructura en Buenos Aires y Bogotá, respectivamente, para mejorar la respuesta a la pandemia y la distribución de vacunas.

En resumen, el análisis detallado con visualizaciones y el dashboard permitió identificar las necesidades específicas de cada país y ofrecer recomendaciones estratégicas para la expansión de laboratorios y centros de vacunación. Las decisiones informadas basadas en datos son cruciales para mejorar la salud pública y gestionar mejor la pandemia de COVID-19 en América Latina.

Reflexión personal

Este proyecto ha sido una experiencia enriquecedora y ha proporcionado un profundo entendimiento sobre el análisis de datos en el contexto de la salud pública y la expansión estratégica de infraestructuras sanitarias. A lo largo del proyecto, se han adquirido y reforzado varias habilidades clave:

Análisis Exploratorio de Datos (EDA):

Aprender a realizar un análisis detallado de las variables clave para identificar patrones, tendencias y anomalías.

Uso de técnicas avanzadas de pandas y numpy para la manipulación y análisis de series temporales.

Visualización de Datos:

Desarrollo de habilidades en la creación de visualizaciones efectivas y persuasivas utilizando herramientas como Power BI y librerías de Python como matplotlib y seaborn.

Módulo 4

Implementación de diversos tipos de gráficos, incluidos mapas de calor, diagramas de dispersión, diagramas de caja y bigotes, y gráficos de líneas, para comunicar de manera clara los hallazgos del análisis.

Manejo y Limpieza de Datos:

Experiencia en la preparación de datos, incluyendo la limpieza, transformación y normalización de conjuntos de datos grandes y complejos.

Uso de técnicas de imputación de valores faltantes y manejo de outliers.

Uso de Herramientas de BI:

Desarrollo de dashboards interactivos y dinámicos en Power BI para facilitar la toma de decisiones basadas en datos.

Creación de medidas y columnas calculadas para proporcionar insights específicos y relevantes.

Desarrollo de Scripts en Python:

Integración de scripts de Python en Power BI para realizar análisis avanzados y crear visualizaciones personalizadas.

Aplicación de técnicas de análisis de correlación y descomposición de series temporales para evaluar relaciones entre variables y entender la evolución de la pandemia.

Reflexión Personal

Si tuviera que volver a empezar este proyecto, hay ciertos aspectos que mantendría y otros que consideraría ajustar:

Aspectos a Mantener:

Enfoque en el Análisis Exploratorio:

Continuaría con un análisis exploratorio exhaustivo al inicio del proyecto para entender mejor el conjunto de datos y formular hipótesis claras.

Integración de Herramientas de BI y Python:

Mantendría el uso de Power BI y Python juntos, ya que esta combinación permite un análisis profundo y visualizaciones efectivas.

Iteración y Refinamiento:

La iteración constante para refinar visualizaciones y medidas es crucial y seguiría siendo parte integral del proceso.

Cambios Considerados:

Planificación del Proyecto:

Invertiría más tiempo en la planificación inicial del proyecto para definir claramente las etapas y los objetivos específicos de cada fase.

Automatización del Proceso de Limpieza de Datos:

Módulo 4

Desarrollaría scripts más robustos para la limpieza y preparación de datos desde el inicio, para reducir el tiempo invertido en estas tareas a lo largo del proyecto.

Documentación:

Aumentaría el énfasis en la documentación de cada paso del proceso, incluyendo decisiones tomadas y racionales detrás de ellas, para facilitar la revisión y el seguimiento del proyecto.

Reflexión Final

El proyecto ha sido una oportunidad invaluable para aplicar y expandir mis habilidades como analista de datos. La capacidad de extraer insights valiosos de los datos y traducirlos en recomendaciones estratégicas es crucial para cualquier analista. Este proyecto no solo ha fortalecido mis competencias técnicas, sino que también ha subrayado la importancia de una planificación meticulosa y una documentación exhaustiva.

En conclusión, si tuviera que empezar de nuevo, abordaría el proyecto con una estrategia más estructurada y con mayor automatización en las etapas iniciales, pero mantendría el rigor en el análisis y la creatividad en las visualizaciones que han sido fundamentales para el éxito del proyecto.

EXTRA CREDIT

Durante el desarrollo del proyecto se identificaron varias áreas clave que se podrían mejorar para aumentar la eficiencia, la precisión y el impacto del análisis. A continuación, se presentan las mejoras implementadas y las sugerencias para futuros proyectos similares:

Mejoras Implementadas:

Automatización del Proceso de Limpieza de Datos:

Script de Limpieza Automática: Se desarrolló un script de Python para automatizar la limpieza y preparación de los datos, lo que redujo significativamente el tiempo invertido en estas tareas manuales.

Estandarización de Formatos: Se establecieron reglas claras para la estandarización de formatos de datos, lo que facilitó la integración y análisis de múltiples fuentes de datos.

Visualizaciones Mejoradas:

Gráficos Interactivos en Power BI: Se crearon visualizaciones más interactivas y dinámicas en Power BI, permitiendo a los usuarios explorar los datos en profundidad y obtener insights más detallados.

Uso de Colores y Diseños Consistentes: Se mejoró la estética de los gráficos utilizando una paleta de colores consistente y diseños claros que facilitaron la interpretación de los resultados.

Análisis de Correlación Avanzado:

Submáscara en Mapas de Calor: Se implementó una submáscara en los mapas de calor para resaltar correlaciones significativas, mejorando la capacidad de identificar relaciones clave entre variables.

Módulo 4

Descomposición de Series Temporales: Se utilizó la descomposición de series temporales para analizar la evolución de las variables a lo largo del tiempo, identificando patrones y tendencias ocultas.

Medidas y Columnas Calculadas Específicas:

Medidas de Tasa de Vacunación y Mortalidad: Se crearon medidas específicas en Power BI para calcular tasas de vacunación y mortalidad, proporcionando métricas críticas para la toma de decisiones.

Columnas Calculadas para Indicadores Clave: Se añadieron columnas calculadas para indicadores clave como la densidad poblacional y la disponibilidad de personal médico, facilitando un análisis más profundo.

Sugerencias para Futuros Proyectos:

Mejora en la Planificación del Proyecto:

Definición Clara de Objetivos y Etapas: Dedicar más tiempo a la planificación inicial del proyecto para definir claramente los objetivos, las etapas y las responsabilidades, asegurando una ejecución más fluida y organizada.

Optimización del Rendimiento de Power BI:

Uso de Agregaciones y Filtrados Eficientes: Implementar técnicas avanzadas de agregación y filtrado en Power BI para mejorar el rendimiento y la velocidad de carga de los dashboards, especialmente cuando se trabajan con conjuntos de datos grandes.

Documentación Exhaustiva:

Registro Detallado de Decisiones y Procesos: Mantener una documentación exhaustiva de cada paso del proceso, incluyendo decisiones tomadas, razonamientos y problemas encontrados, facilitando la revisión y la colaboración en equipo.

Capacitación y Desarrollo Continuo:

Entrenamiento en Herramientas y Técnicas Avanzadas: Invertir en capacitación continua para mantenerse actualizado con las últimas herramientas y técnicas de análisis de datos y visualización, asegurando la aplicación de las mejores prácticas en futuros proyectos.

Las mejoras implementadas durante el desarrollo de este proyecto han incrementado significativamente la calidad y la utilidad del análisis, proporcionando a BIOGENESYS insights valiosos para la toma de decisiones estratégicas. Al adoptar un enfoque más estructurado y eficiente, y al mantener un compromiso con la mejora continua, es posible alcanzar niveles aún mayores de excelencia en futuros proyectos de análisis de datos.