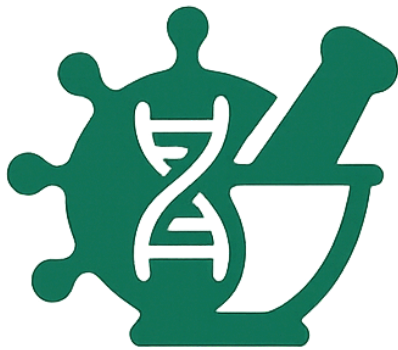


Informe Biogenesys



BIOGENESYS

Nombre del autor: Gian Pisani Arntsen

Email: gianpisani01@gmail.com

Cohorte: DAFT-12

Fecha de entrega: 07/04/2025

Introducción

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

El objetivo es realizar un estudio que ayudará en su estrategia de expansión en Latinoamérica específicamente en Colombia, Argentina, Chile, México, Perú y Brasil.

Este proyecto es vital para que estén preparados y puedan tener una respuesta rápida ante cualquier situación futura que pueda surgir, la directiva ha propuesto que para poder hacer esta inversión necesitan ubicar regiones y recolectar datos que sean de importancia y que les ayuden a tomar decisiones.

Desarrollo del proyecto

1° avance

Carga

1. Creamos un nuevo notebook e importamos las siguientes librerías necesarias para realizar el primer avance del PI.
 - a. Numpy → significa Numerical Python y es una biblioteca de Python fundamental para el cálculo científico.
 - b. Pandas → simplifica el proceso de carga, manipulación, y análisis de datos, proporcionando una forma intuitiva y consistente de trabajar con datos estructurados.
2. A su vez, recomendamos instalar la extensión de DataWrangler que te permite visualizar en formato tabular los DF de Pandas.
3. Leemos el archivo data_latinoamerica.csv a través de la librería Pandas. Pandas proporciona una serie de funciones para cargar datos desde varios formatos de archivo directamente en un DataFrame. En este caso utilizamos la función para cargar el archivo .csv, especificando:
 - a. Parseo de fechas → ya que el archivo contiene la columna "Date" que es una fecha. Se utiliza este argumento para asegurarse de que esta columna se interprete correctamente como fecha.
4. Comprobamos que el dataset cargado tiene la cantidad de registros y columnas especificadas utilizando el método .info() de Pandas.

Filtrado

5. Creamos nueva data frame filtrando solo por los países donde se expandirán: Colombia, Argentina, Chile, México, Perú y Brasil.
 - a. Filtramos con isin() que permite seleccionar filas de un DataFrame donde un valor o conjunto de valores específicos, aparecen en una columna.
6. Además, descubrimos que hay una variable llamada "location_key" que subdivide los países en distintas regiones. Decidimos filtrarnos aquellas subregiones y quedarnos solo con la "location_key" que hace referencia a los países relevantes para el análisis.
7. Filtramos los datos en fechas mayores a 2021-01-01, como indica la consigna.
 - a. Creamos una máscara booleana que es True para todas las filas donde la condición se cumple, es decir, donde la fecha es mayor a 2021-01-01. Luego, usamos esta máscara para indexar la df.

Manejo de nulos

8. Comparamos a nivel de país la cantidad de valores faltantes por columna.
 - a. Para lograr esto, agrupamos por país y utilizamos una función lambda para sumar el total de registros nulos por columna.
 - b. Nos filtramos por aquellas columnas que contengan valores nulos mediante .loc[] que sirve para aplicar condiciones complejas y hacer rebanados más legibles y explícitos.

Módulo 4

9. Eliminamos las filas que todos sus valores sean nulos y también eliminamos las columnas que contienen únicamente valores nulos (100% de sus valores NaN - Not a Number)
 - a. Usamos dropna() que elimina filas o columnas con datos faltantes
10. De todas maneras, siguen quedando valores nulos que debemos manejar porque todavía quedan variables que no están compuestas 100% por valores nulos. Decidimos eliminar aquellas variables que estén compuestas por un total de valores nulos mayor al 60% del total de registros. Esto se debe a que no cumplen con el criterio de completitud donde el conjunto de datos no contiene todos los valores necesarios para un análisis preciso y significativo.
 - a. Creamos una barrera con el 60% del total de registros.
 - b. Generamos una variable con los nombres de las columnas que su cantidad de valores nulos está por encima del 60%.
 - c. Eliminamos estas columnas del data frame.
11. Todavía nos encontramos con columnas que están compuestas por cierta cantidad de valores nulos.
 - a. Evaluamos qué columnas contienen valores nulos y nos quedamos con los nombres de estas columnas alojados en una variable.
 - b. A través de una iteración FOR, para cada una de estas columnas alojadas en la variable, decidimos completar los valores faltantes con la mediana de dicha columna agrupando por país. Es decir, la mediana de la variable pero del país, no general.
 - c. Decidimos completarlos con la mediana ya que representa el valor del medio del total de valores, es decir, el valor que separa la mitad superior de la mitad inferior de una muestra de datos. Usamos este valor para completar en vez de la media porque nos parece más representativo. La media es sensible a outliers y puede generar valores que no sean precisos.
12. Para la columna "new_recovered", que es el conteo de nuevas recuperaciones de un caso positivo de COVID-19 en esta fecha, siguen quedando nulos. Esto se debe a que ciertos países lo contabilizaron, mientras que, otros no. Con el fin de poder desarrollar análisis posteriores sin romper las fórmulas vamos a asignarle valor = 0 para estos países pero entendiendo que representan valor nulo. Es decir, siempre que queramos hacer análisis posteriores con esta variable vamos a tener que categorizar por países para comprender correctamente los datos.

Guardamos nueva base

13. Guardamos los datos filtrados y limpiados en un archivo con el nombre DatosFinalesFiltrado.csv a fin de poder utilizarlo luego y no tener que repetir el proceso de filtrado y limpieza cada vez que referenciamos el df.

Estadísticos descriptivos

14. Finalmente, creamos una función para sacar la media, mediana, varianza y rango de la columna nombrada para el país especificado. Si o si poniendo como parámetro alguna de las columnas que esté compuesta por valores numéricos.
 - a. También, hicimos una iteración para calcular los estadísticos de las columnas numéricas llamando a la función previamente creada.

Módulo 4

Preguntas

15. ¿Qué implican estas métricas y cómo pueden ayudar en el análisis de datos?
 - a. Mediana: Valor central de un conjunto de datos ordenados. Útil para evitar sesgos por valores extremos. Ayuda a entender la tendencia central en datos asimétricos.
 - b. Varianza: Medida de dispersión que indica cuán lejos están los datos de la media. Mide que tan dispersos están los datos. Una varianza alta indica mayor variabilidad, mientras que una baja señala consistencia.
 - c. Rango: Diferencia entre el valor máximo y mínimo en un conjunto de datos. Proporciona una idea rápida de la dispersión total. Limitado por no considerar la distribución interna de los datos.
16. ¿Cómo varía la dispersión de los datos en el conjunto de datos analizado, en términos de la varianza y el rango?
 - a. La dispersión de los datos en el conjunto de datos analizado es alta tanto en término de varianza como en el rango. Los datos están muy dispersos, indicando alta variabilidad y poca consistencia alrededor de la media/mediana. Confirma gran amplitud entre valores mínimos y máximos, reforzando la falta de uniformidad en los datos.
17. Confirma gran amplitud entre valores mínimos y máximos, reforzando la falta de uniformidad en los datos.
 - a. La mediana puede no representar bien los datos debido a la alta dispersión, sugiriendo outliers o una distribución muy extendida.
18. ¿Se muestran todas las estadísticas en todas las columnas durante el análisis?
¿Cuál es la razón de la respuesta anterior y cómo podría afectar la interpretación de los resultados obtenidos?
 - a. No, debido a que existen datos no numéricos y, específicamente, la variable "new_recovered" se le asignaron valores = 0 para representar nan. Lo segundo podría afectar a los resultados si se comprende como que no hubo casos y no como realmente se tiene que comprender que es como si fuesen valores nulos.

2° y 3° avance

Para lograr estos avances nos profundizamos en el análisis exploratorio de datos para extraer insights valiosos que nos orienten en la planificación estratégica de la expansión de laboratorios farmacéuticos. Buscamos descubrir patrones, tendencias y anomalías en los datos de incidencia de COVID-19, tasas de vacunación y disponibilidad de infraestructura sanitaria. Con el objetivo de crear un cuadro completo que permita a los directivos tomar decisiones informadas sobre dónde ubicar los nuevos laboratorios y centros de vacunación para maximizar el impacto positivo en la salud pública.

Logramos obtener todos los resultados requeridos mediante la siguiente estructura para llevar a cabo correctamente el proceso de EDA:

Carga

1. Importamos librerías necesarias para utilizar en este avance
 - a. Pandas y Numpy para hacer análisis estadísticos.
 - b. Matplotlib y Seaborn para visualización de datos.

Módulo 4

2. Leemos el archivo "DatosFinalesFiltrados.csv" que es la base final ya limpia.

Análisis Estadístico

3. Desarrollamos una función para calcular los estadísticos principales de la base de datos. Con esta función se puede filtrar para un país específico o se puede calcular para toda la base.
4. Siguiendo con la misma lógica, hicimos una función que calcula la correlación entre las variables. También se le puede poner un país como parámetro para que filtre la correlación entre variables para ese país específico.
 - a. Si aparecen missing values es porque std = 0 para esa columna. Es decir, todos los valores son iguales.

Análisis Univariado - Objetivo: Entender la distribución de variables clave (composición, outliers, concentración)

1. Histogramas de las variables relevantes (casos, muertes, vacunación)
2. Boxplot de temperatura media por país
3. Distribución de la población por grupos de edad

Análisis Bivariado / Multivariado - Objetivo: Identificar relaciones y patrones entre variables para descubrir factores críticos

1. Matriz de correlación (filtrada por > 0.5)
2. Diagrama de dispersión: temperatura media vs. casos confirmados
3. Diagrama de dispersión: temperatura media vs. muertes confirmadas
4. Relación entre la Cobertura de Vacunación y la Cantidad de Casos Nuevos

Análisis Temporal - Objetivo: Observar tendencias a lo largo del tiempo y detectar momentos críticos

1. Evolución de dosis administradas por mes de cada país
2. Evolución del número de muertes diarias
3. Variación Mensual de Casos y Muertes

Análisis Geográfico / Comparativo - Objetivo: Evaluar la situación por país/región y priorizar territorios para expansión

1. Gráficos de barras de países contra variables clave

4° avance

En esta fase final del proyecto nos enfocamos en integrar y presentar los hallazgos analíticos y hacer reporting utilizando Power BI para la visualización. Nuestro objetivo fue sintetizar el análisis realizado en las fases anteriores en dashboards interactivos y reportes que faciliten la toma de decisiones estratégicas para la expansión de laboratorios y centros de vacunación. Estos insights ayudarán a la empresa farmacéutica a priorizar áreas de inversión basadas en la incidencia de COVID-19. Para lograrlo llevamos a cabo el siguiente proceso:

Carga

1. Importamos la base de datos final con el conector de Python.

Módulo 4

2. Chequeamos que los datos se hayan cargado con el formato correcto en Power Query. De hacer falta lo corregimos en este entorno.
3. Cerramos y aplicamos para seguir trabajando en Power BI.

Creación del Dashboard

Generamos un reporte con los siguientes dashboards como para poder navegar y extraer las conclusiones necesarias con respecto al objetivo de este proyecto, hacer una evaluación sobre la demanda de vacunas, la logística de su distribución y la infraestructura sanitaria ya existente.

1. Presentación - Esta página simplemente presenta el objetivo del reporte y está compuesta por el logo de la empresa y los botones para acceder a las otras páginas pertinentes.
2. Demanda COVID-19 - El objetivo de este dashboard es adentrarnos al nuevo mercado el cual quiere ingresar la empresa y poder analizar la oferta y la demanda de vacunación desde la perspectiva del COVID-19. La idea es poder tener una visión general de un caso de estudio real y no hacer simplemente estimaciones en base a estadísticas descriptivas de la población. A su vez, se encuentran segmentadores con el fin de poder detallar mejor las visualizaciones que queremos ver, así como poder especificar por año y por país.
3. Demanda - Acá nos centramos en la evaluación de Demanda Potencial de todo tipo de vacunas. La idea es identificar países con alta necesidad de vacunación por exposición, bajo acceso y riesgo poblacional. Es decir, enfocar la estrategia donde hay alto riesgo + baja cobertura.
4. Infraestructura - En este caso quisimos medir la infraestructura sanitaria existente y la logística y distribución potencial. Por un lado, evaluando cuán viable es llevar vacunas a distintos territorios, destacando los desafíos logísticos regionales. Y, por el otro, medir la capacidad del sistema de salud en cada país para gestionar campañas de vacunación y respuesta pandémica.

EDA e insights

Gráfico: Histogramas de las variables clave

Insight: Visualización de la distribución de un conjunto de variables numéricas clave a través de barras.

Gráfico: Distribución de Temperatura Media por País

Insight: El gráfico muestra que Brasil y Perú tienen las temperaturas medias más altas y consistentes. Argentina y Chile presentan mayor variabilidad, con valores extremos más bajos. Colombia y México tienen distribuciones más centradas, con menor dispersión. En general, los países tropicales tienden a temperaturas más elevadas.

Gráfico: Distribución de Población por Grupos de Edad

Insight: La mayor parte de la población de LAM se concentra entre los 0 y 39 años, mostrando un claro sesgo hacia edades jóvenes. A partir de los 50 años, la población disminuye progresivamente, con una fuerte caída a partir de los 60. Esta distribución sugiere que, aunque el riesgo sanitario es mayor en adultos mayores, la demanda

Módulo 4

futura de servicios de salud también debería anticiparse en las generaciones jóvenes que irán envejeciendo, justificando inversiones sostenidas en infraestructura médica.

Gráfico: Heatmap de la Matriz de Correlación de las Variables

Insight: Se identifican correlaciones muy fuertes entre los grupos etarios, reflejando consistencia en la estructura poblacional por país. También destacan correlaciones positivas entre variables demográficas y socioeconómicas como el índice de desarrollo humano, médicos/enfermeros por cada mil habitantes, y el PIB per cápita, indicando que los países con mejor infraestructura también tienden a tener mejores condiciones de salud. Llama la atención la correlación negativa entre la tasa de mortalidad infantil y la esperanza de vida, y entre nuevos casos y vacunación acumulada, lo cual sugiere efectividad en la inmunización.

Gráfico: Temperatura Media vs. Casos Confirmados (por país)

Insight: No se observa una relación lineal clara entre la temperatura media y la cantidad de casos confirmados de COVID-19. Si bien algunos países como Brasil y Argentina muestran una ligera concentración de casos en rangos de temperatura específicos (20-30 °C), en general los contagios se distribuyen ampliamente en distintos climas. Esto sugiere que la temperatura, por sí sola, no es un factor determinante en la propagación del virus, reforzando la necesidad de considerar otras variables como vacunación o densidad poblacional para tomar decisiones estratégicas.

Gráfico: Temperatura Media vs. Muertes Confirmadas (por país)

Insight: Los gráficos muestran que las muertes por COVID-19 no tienen una relación clara con la temperatura media, variando significativamente entre países. Brasil y México presentan mayor mortalidad en climas cálidos (20-30°C), mientras que Argentina y Colombia en rangos moderados (15-25°C). Esto sugiere que otros factores (como densidad poblacional o medidas sanitarias) influyen más que el clima. Estas zonas de alta mortalidad podrían ser prioritarias para la expansión de laboratorios.

Gráfico: Relación entre Acumulado de Dosis de Vacunas y Nuevos Casos Confirmados

Insight: A mayor acumulación de dosis de vacunas, se observa una tendencia a la disminución de nuevos casos confirmados de COVID-19. Esto respalda la eficacia de la vacunación como estrategia clave para reducir la incidencia del virus, reforzando la necesidad de priorizar zonas con baja cobertura vacunal en la expansión de laboratorios.

Gráfico: Evolución de las dosis acumuladas por mes de cada país

Insight: La evolución de dosis acumuladas muestra un crecimiento general en todos los países, con posibles variaciones bruscas atribuibles a correcciones de datos (outliers o valores imputados). Brasil y México lideran en volumen de vacunación, mientras que países como Perú presentan un avance más gradual, destacando disparidades en la capacidad de distribución que podrían orientar la expansión de laboratorios.

Gráfico: Evolución del Número de Muertes Diarias

Insight: La evolución de muertes diarias muestra fluctuaciones significativas, con picos que podrían asociarse a oleadas epidémicas o variantes del virus. La presencia de valores atípicos (outliers) sugiere la necesidad de validar la consistencia de los

Módulo 4

datos, especialmente en periodos con cambios bruscos. Esta tendencia resalta la importancia de reforzar la capacidad de respuesta en regiones con mayor mortalidad. De todas maneras, se ve una clara tendencia a la baja de las muertes diarias con el pasar del tiempo, posiblemente explicado por el aumento en la campaña vacunatoria.

Gráfico: Variación Porcentual Mensual: Casos Nuevos vs Muertes

Insight: El gráfico muestra que tanto los casos nuevos como las muertes presentan fluctuaciones mensuales moderadas. En general, ambas curvas siguen una tendencia similar, con aumentos y disminuciones relativamente sincronizados. A lo largo del periodo, no se observan variaciones extremas sostenidas, lo que sugiere un comportamiento relativamente estable de la pandemia, con repuntes y descensos graduales. Aunque se podría observar un pequeño lag en la variación de muertes con respecto a la de casos nuevos. Esto puede ser porque de una tanda alta de nuevos casos, hay más chances de que se muera gente, pero estas muertes no son instantáneas, sino que, pueden venir con un retraso de semanas.

Gráfico: Gráficos de barras de países contra el resto de variables

Insight: Serie de gráficos útiles para comparaciones categoricas. Es decir, podemos evaluar cada variable individualmente cómo se compara entre los distintos países de interés.

Análisis del dashboard

Demanda COVID-19

Análisis visualizaciones

1. Indicadores superiores (Casos Totales, Muertes, Recuperados, Vacunas Aplicadas)
 - a. Propósito: Ofrecer un resumen rápido del impacto del COVID-19 en términos generales.
 - b. Relevancia para el objetivo: Ayuda a dimensionar la magnitud de la pandemia en la región analizada, lo cual es esencial para justificar la necesidad de expansión de laboratorios. El alto número de casos y muertes sugiere una fuerte demanda histórica (y posiblemente futura) de vacunas.
2. Evolución Mensual del Total de Muertes por COVID-19
 - a. Propósito: Mostrar la tendencia temporal de fallecimientos.
 - b. Relevancia: Es clave para entender en qué momentos hubo mayor necesidad sanitaria. Una disminución en muertes puede reflejar efectividad de vacunación, pero también una reducción en la demanda. Esta información puede orientar decisiones sobre si el enfoque debe ser más preventivo o reactivo.
 - c. Observación clave: Las muertes por COVID-19 tuvieron picos entre febrero y abril de 2021, y entre diciembre y febrero de 2022, seguido de una tendencia descendente sostenida. Podríamos observar que durante los meses de verano/vacaciones de estos países aumentan las muertes y podría ser por un aumento en actividad social que lleva a mayores contagios y, por ende, mayores muertes.

Módulo 4

3. Vacunas de COVID-19 Administradas vs Población por País
 - a. Propósito: Comparar la cantidad de dosis administradas frente a la población total por país.
 - b. Relevancia: Muy útil para analizar la cobertura vacunal y detectar brechas de vacunación, lo que puede traducirse en oportunidades de mercado para BIOGENESYS. Por ejemplo, países con alta población pero baja proporción de vacunación pueden ser prioritarios.
 - c. Observación clave: Brasil lidera en dosis administradas, incluso superando su población total (alta cobertura), seguido por México. Esto los posiciona como mercados maduros o con alta penetración vacunal, aunque aún pueden necesitar refuerzos o nuevas campañas.
4. Evolución Mensual de Dosis de Vacunas Acumuladas de COVID-19 (por país)
 - a. Propósito: Mostrar el ritmo de vacunación acumulada por país.
 - b. Relevancia: Permite identificar qué países tuvieron un despliegue más rápido y cuáles fueron más lentos o estables. Esto es clave para evaluar la logística de distribución existente y la posible eficiencia en la adopción de vacunas nuevas.
 - c. Observación clave: Brasil y México tienen un crecimiento rápido y sostenido en vacunación. Colombia y Perú presentan un avance más tardío y moderado, indicando una potencial demanda insatisfecha o dificultad logística. A su vez, nos permite identificar en qué países va a ser mayor la competencia debido a su respuesta de vacunación.
5. Mapa de Casos Totales por País
 - a. Propósito: Representación geoespacial de la distribución de casos.
 - b. Relevancia: Muy útil para el análisis territorial. Permite vincular áreas de alta incidencia con necesidades sanitarias críticas, lo cual favorece la selección de ubicaciones óptimas para nuevos laboratorios o centros de distribución.
 - c. Observación clave: Argentina y Brasil tienen los mayores volúmenes de casos, lo que podría traducirse en una mayor conciencia pública y demanda sostenida de soluciones médicas o vacunas en esos países.

Conclusion

Este primer dashboard proporciona una base sólida para el análisis exploratorio del entorno de mercado. Ofrece:

- Un panorama epidemiológico claro (casos, muertes, recuperación).
- Información clave sobre la demanda histórica de vacunas.
- Datos relevantes para estimar la capacidad logística y eficiencia de distribución.
- Indicadores visuales que ayudan a identificar países con oportunidades de expansión debido a la cobertura vacunal insuficiente o alta incidencia.

Demanda General

Análisis visualizaciones

1. Distribución regional de la población mayor a 60 años (gráfico de dona)
 - a. Propósito: Muestra qué porcentaje de la población mayor de 60 años pertenece a cada país.
 - b. Relevancia: Altamente útil, ya que la población mayor a 60 años representa un grupo prioritario para vacunación por su vulnerabilidad.

Módulo 4

- c. Observación clave: Brasil concentra más del 50% de esta población, lo que representa una demanda sanitaria crítica y una oportunidad estratégica de expansión.
2. GDP Total y Población Total (indicadores superiores)
 - a. Propósito: Contextualizar el tamaño económico y poblacional del conjunto de países analizados.
 - b. Relevancia: El GDP es relevante para evaluar la capacidad de inversión en infraestructura sanitaria, y la población ayuda a estimar la demanda potencial total de vacunas.
 - c. Observación clave: Con 466M de habitantes y un GDP conjunto alto, existe una masa crítica suficiente para justificar la inversión farmacéutica.
3. Distribución de Población por Grupos de Edad
 - a. Propósito: Representar la pirámide poblacional.
 - b. Relevancia: Esencial para prever la demanda sostenida de vacunas y otros productos a largo plazo. Una población joven sugiere demanda futura constante.
 - c. Observación clave: Aunque hay un segmento joven dominante, también hay una proporción considerable de personas mayores, sobre todo de 60 a 79 años.
4. Esperanza de Vida por País
 - a. Propósito: Mostrar el promedio de años de vida por país.
 - b. Relevancia: Un indicador indirecto de la calidad del sistema sanitario y también de la longevidad de la demanda farmacéutica.
 - c. Observación clave: Chile lidera en esperanza de vida, lo que puede indicar una infraestructura sanitaria más madura que podría facilitar la instalación o colaboración con laboratorios.
5. Media de la Tasa de Mortalidad:
 - a. Propósito: Comparar las causas de mortalidad promedio por cada 1.000 personas en distintos factores de riesgo.
 - b. Relevancia: Identifica áreas de vulnerabilidad sanitaria que pueden incrementar la necesidad de intervención farmacéutica, como vacunas o tratamientos.
 - c. Observación clave: Las comorbilidades (154.27) y la contaminación (81.43) son los factores más críticos, sugiriendo regiones con alta fragilidad sanitaria y, por lo tanto, alta prioridad en campañas de vacunación y desarrollo de soluciones médicas.

Conclusion

Este dashboard es clave para evaluar la demanda potencial de vacunas y otros servicios farmacéuticos. Destaca por:

- Proporcionar una visión clara del segmento prioritario (mayores de 60 años).
- Ofrecer una mirada demográfica que permite estimar la sostenibilidad y evolución futura del mercado.
- Incluir variables económicas y sanitarias que permiten cruzar capacidad de respuesta con necesidades epidemiológicas.

Infraestructura

Análisis visualizaciones

1. Enfermeras y Médicos por cada 1000 personas por país

Módulo 4

- a. Propósito: Evaluar la disponibilidad de personal médico por país.
 - b. Relevancia: Es un indicador directo de la capacidad del sistema de salud para manejar campañas de vacunación o distribución farmacéutica.
 - c. Observación clave: Chile y Brasil sobresalen por su alta proporción de enfermeras, lo cual facilita el despliegue operativo de vacunas; mientras que Argentina lidera en médicos.
2. Porcentaje de Área Urbana por país
 - a. Propósito: Indicar qué proporción del territorio es urbano.
 - b. Relevancia: Las zonas urbanas tienden a tener mejor infraestructura logística y sanitaria, lo que puede facilitar la implementación de laboratorios o centros de distribución.
 - c. Observación clave: México y Colombia tienen la mayor proporción de área urbana, lo que sugiere mejor acceso logístico para distribución.
3. Lluvia mensual por país
 - a. Propósito: Mostrar la variabilidad climática que puede afectar la logística.
 - b. Relevancia: Alta humedad o lluvias intensas pueden afectar transporte, almacenaje y conservación de vacunas.
 - c. Observación clave: Perú y Colombia muestran picos de lluvias considerables, lo que podría ser un desafío logístico durante ciertos meses.
4. Diagrama de dispersión: Temperatura media vs Casos confirmados por país
 - a. Propósito: Analizar la posible relación entre temperatura y casos de COVID-19.
 - b. Relevancia: Ayuda a entender si hay patrones climáticos asociados a brotes, lo cual puede informar estrategias de respuesta estacional.
 - c. Observación clave: México muestra una alta densidad de casos en temperaturas altas (25-35°C), lo que puede influir en la planificación de campañas según la estacionalidad.

Conclusion

Este dashboard aporta información esencial para tomar decisiones sobre capacidad instalada, logística y adecuación climática. Sus aportes clave son:

- Identificación de países con mejor dotación de personal sanitario (Chile, Brasil, Argentina).
- Visión clara de facilidades urbanas que podrían facilitar la logística (México, Colombia).
- Análisis climático que permite prever riesgos de distribución estacionales.

Conclusiones y Recomendaciones

A partir del análisis integral de los tres dashboards ("Impacto COVID-19", "Demanda General" e "Infraestructura"), se pueden destacar los siguientes países como candidatos óptimos, clasificados según prioridad estratégica:

1. Brasil – Alta prioridad

- a. Demanda: Mayor número de casos y muertes totales, así como el mayor volumen de vacunación acumulada. Esto refleja un mercado altamente movilizado y consciente de la importancia sanitaria.

Módulo 4

- b. Infraestructura: Alta cantidad de enfermeras por cada 1000 personas y una gran población mayor de 60 años (50% de la región), lo que sugiere alta capacidad de respuesta y demanda sostenida.
- c. Logística: Aunque su área urbana es baja, el tamaño de su sistema sanitario compensa esta desventaja.
- d. Conclusión: Brasil combina una demanda alta, infraestructura sólida y experiencia previa en campañas de vacunación masiva. Ideal para establecer centros de producción y distribución regionales.

2. México – Alta prioridad

- a. Demanda: Altos casos confirmados y segundo en vacunación. Aún con buena cobertura, los datos climáticos indican nuevos brotes posibles en zonas cálidas, lo que sostiene la demanda.
- b. Infraestructura: Tiene el mayor porcentaje de área urbana, lo que facilita enormemente la distribución logística. Moderada dotación de personal sanitario.
- c. Conclusión: Su ubicación geográfica, urbanización y volumen poblacional lo posicionan como un nodo estratégico para distribución y logística farmacéutica.

3. Colombia – Potencial de crecimiento

- a. Demanda: Casos moderados y crecimiento de vacunación lento, lo que indica espacio para expansión del mercado.
- b. Infraestructura: Buen porcentaje de urbanización y distribución poblacional diversa, pero baja dotación médica.
- c. Conclusión: Ideal para expansión progresiva o establecimiento de centros de investigación y vacunación en zonas urbanas.

4. Perú – Oportunidad emergente

- a. Demanda: Baja cobertura médica y alta mortalidad por comorbilidades. Esto señala una población vulnerable con barreras de acceso.
- b. Logística: Área urbana limitada, pero buen crecimiento en vacunación y patrón estacional claro en lluvias.
- c. Conclusión: Potencial interesante como foco de intervención social y médica, aunque dependerá de mejoras en infraestructura logística.

5. Chile y Argentina – Alta infraestructura, menor demanda latente

- a. Infraestructura: Altos estándares sanitarios y esperanza de vida. Argentina tiene mayor dotación médica; Chile, la más alta esperanza de vida.
- b. Demanda: Menor incidencia de casos o crecimiento moderado en vacunación, lo que sugiere mercados más estabilizados.
- c. Conclusión: Útiles como centros de soporte técnico, investigación o respaldo, más que como primeros puntos de expansión comercial.

Recomendación final

BIOGENESYS debería priorizar la expansión de laboratorios farmacéuticos en Brasil y México, por su combinación ideal de alta demanda, infraestructura médica sólida y capacidad logística. Colombia y Perú pueden ser considerados en una segunda fase

Módulo 4

como regiones de crecimiento estratégico o atención prioritaria por vulnerabilidad sanitaria.