



# REPORTE DE RESULTADOS SOBRE EL ANALISIS INTEGRAL

## DE BIOGENESYS FARMACEUTICAL

### Objetivo:

Realización de estudio para identificar ubicaciones óptimas para la expansión de laboratorios en Latinoamérica, determinando que regiones son optimas para la respuesta a los efectos de la pandemia COVID-19 con el fin de mejorar el acceso a las vacunas.



KAREN S QUEZADA  
KARENQUEZADA@GMAIL.COM  
DAFT17  
29 DE SEPTIEMBRE 2025



# PRESENTACIÓN

---

En ENFOKATE, somos una empresa especializada en análisis de datos estratégicos, orientados a la toma de decisiones críticas en sectores de alto impacto. Nuestro propósito es transformar datos complejos en información clara, precisa y accionable.

Contamos con un equipo multidisciplinario con experiencia en ciencia de datos, inteligencia de negocios y análisis geoespacial, lo que nos permite ofrecer soluciones personalizadas basadas en evidencia.

En esta oportunidad, hemos trabajado en identificar la ubicación óptima para la implementación de un laboratorio de vacunas, considerando factores clave como infraestructura, entorno regulatorio, disponibilidad de talento, estabilidad económica y logística global. Los resultados que presentamos a continuación están respaldados por un proceso riguroso de recolección, procesamiento y modelado de datos.

Gracias por confiar en ENFOKATE. Estamos aquí para ayudarte a tomar decisiones con visión, precisión y enfoque.

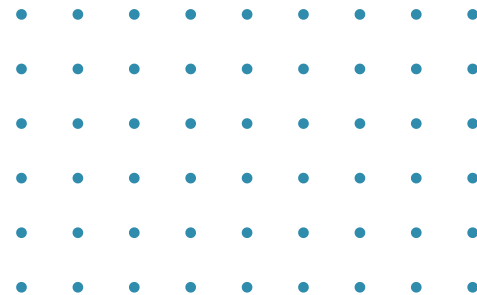
*“ Enfokate en lo esencial.  
Enfokate en los datos.  
Enfokate en el futuro.”*

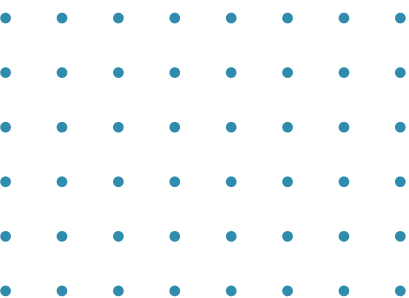


# CONTENIDO

---

- Introducción
- Desarrollo del Proyecto
- EDA e insights
- Análisis del dashboard
- Conclusiones y Recomendaciones
- Reflexión Personal





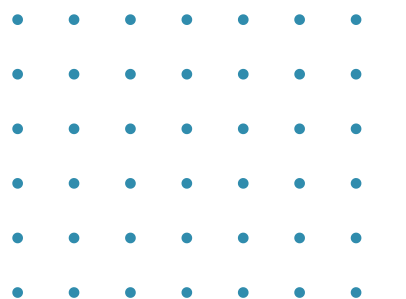
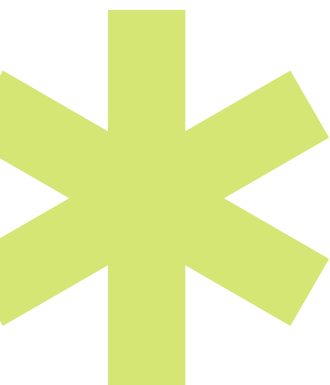
# INTRODUCCIÓN

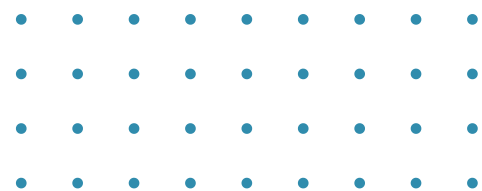
---

El presente informe tiene como objetivo analizar datos epidemiológicos, demográficos y de vacunación en países seleccionados de Latinoamérica con el fin de identificar ubicaciones óptimas para la expansión de laboratorios farmacéuticos.

El proyecto responde a los objetivos organizacionales de:

- Preparación (ETL) reproducible del dataset en Python (lectura por chunks, filtrado de países y fechas, limpieza y agregación).
- Análisis exploratorio (EDA) en Python con visualizaciones para comprender la relación entre casos, mortalidad, dosis aplicadas, urbanización y densidad.
- Optimizar la inversión en infraestructura mediante el uso de análisis de datos.
- Identificar mercados prioritarios basados en demanda sanitaria y capacidad logística.
- Reducir riesgos operativos mediante la integración de información confiable y actualizada.
- Apoyar la toma de decisiones estratégicas a través de un dashboard interactivo en Power BI.
- Recomendaciones estratégicas y una propuesta ubicación objetivo (País) para la implementación de un laboratorio para facilitar la vacunación.





# DESARROLLO DEL PROYECTO

## *(METODOLOGÍA DE RECOPILACIÓN Y SELECCIÓN DE DATOS, TRANSFORMACIONES Y LIMPIEZA CON PYTHON)*

Esta sección corresponde al archivo "PIDA.M4.Karen.Quezada.ETL.ipynb", el cual se dividió en dos apartados con el objetivo de lograr un mejor control de la información y facilitar la depuración de la base de datos. Posteriormente, esto permitirá realizar el análisis exploratorio de datos (EDA). Asimismo, el archivo DATA\_LATINOAMERICA.CSV, que originalmente contenía 12,216,057 registros, fue reducido a 3,756 filas y almacenado con el nombre de DatosFiltradosFinales.csv, de manera que pudiera ser procesado de forma eficiente en Power BI.

### *FUENTES Y SELECCIÓN DE DATOS*

- Datos primarios: CSV de data\_latinoamerica.csv (contiene variables por fecha y ubicación: country\_name, location\_key, date, new\_confirmed, new\_deceased, cumulative\_confirmed, cumulative\_deceased, cumulative\_vaccine\_doses\_administered, population\_urban, population\_density, average\_temperature\_celsius, etc.).  
dataset epidemiológico y demográfico de países de Latinoamérica
- Países objetivo filtrados en el ETL: ["AR" Argentina, "BR" Brazil, "CO" Colombia, "CL" Chile, "MX" México, "PE" Perú].
- Filtro temporal aplicado: datos desde 2020-12-31 / año 2021 en adelante.

### *PROCESO ETL EN PYTHON:*

- Lectura de archivos grandes en chunks para mejorar rendimiento. (pd.read\_csv(ruta, chunksize=100000, parse\_dates=['date'])).
- Filtrado por países y fechas (2021 en adelante). (dentro de cada chunk, luego pd.concat de partes)
- Limpieza: conversión de tipos, eliminación de duplicados, tratamiento de valores faltantes mediante imputación por media del país. ( (pd.to\_datetime para date; conversión numérica con pd.to\_numeric(..., errors='coerce'); por país, usar media del país: columnas\_numericas = df.select\_dtypes(include=[np.number]).columns for col in columnas\_numericas: df[col] = df.groupby('country\_name')[col].transform(lambda x: x.fillna(x.mean()))); Eliminación de filas faltantes en columnas clave (ej. new\_confirmed, date, country\_name) cuando necesario: df = df.dropna(subset=['new\_confirmed','date','country\_name']) )
- Agregación por país de métricas clave: casos confirmados, muertes acumuladas, dosis aplicadas, urbanización y densidad poblacional. )Cálculos agregados: sumar por país columnas COVID: df.groupby('country\_name')[columnas\_covid].sum(). )
- Generación de CSVs de salida: DatosFinalesFiltrados.csv, estadisticas\_columnas.csv, casos\_semanales.csv, etc.



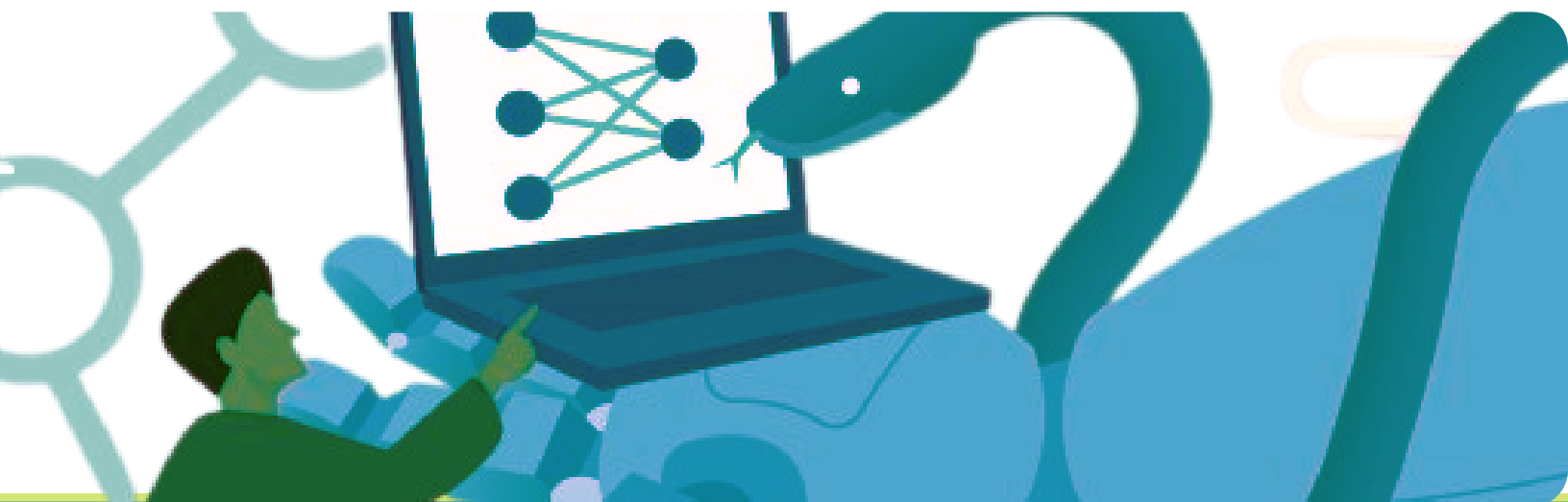
## ***CONCLUSIÓN SOBRE TRANSFORMACIONES Y LIMPIEZA***

- Se generaron datasets consolidados para análisis exploratorio y visualización.
- Se construyeron estadísticas descriptivas y reportes semanales de casos y vacunación.
- La imputación y estandarización de variables garantizó consistencia en las comparaciones entre países.
- El enfoque escalable (chunks), manejo explícito de fechas, imputación de valores numéricos por media por país (consistente cuando faltan valores aleatorios).

Conclusión: El proceso ETL permitió obtener una base sólida, confiable y apta para análisis avanzado y visualización interactiva (EDA).

## ***LIMITACIONES DETECTADAS:***

Falta de granularidad si el dataset solo tiene nivel país – para ubicaciones de laboratorio conviene contar con datos a nivel subnacional (ciudad/estado/municipio) y con datos de infraestructura (hospitales, laboratorios actuales).





# EDA E INSIGHTS

A continuación describo las visualizaciones clave realizadas en el notebook EDA con el nombre de archivo “PIDA\_M4\_Karen\_Quezada\_EDA.ipynb”, cómo reproducirlas en Python y la interpretación (insights).

Los nombres de columnas incluyen:

new\_confirmed, new\_deceased, cumulative\_confirmed, cumulative\_deceased, cumulative\_vaccine\_doses\_administered, population\_urban, population\_density, average\_temperature\_celsius, date, country\_name, entre otras.

## Código base (lectura y configuración)

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

sns.set(style="whitegrid")
plt.rcParams["figure.figsize"] = (12,8)

df = pd.read_csv("DatosFinalesFiltrados.csv", parse_dates=["date"])
```

## Evolución temporal de casos vs vacunación (line plots por país)

```
países = ["Argentina","Brazil","Colombia","Chile","Mexico","Peru"]
fig, ax = plt.subplots()
for país in países:
    sub = df[df.country_name==país].set_index('date').resample('W')
    ['new_confirmed','cumulative_vaccine_doses_administered'].sum()
    sub['new_confirmed'].plot(ax=ax, label=f'{país} - Casos')
plt.legend()
plt.title("Evolución semanal de casos nuevos por país")
plt.show()
```

Interpretación: comparar picos de casos con campañas de vacunación.

Evolución de casos y dosis administradas en cada país.

## Matriz de correlaciones (heatmap)

```
icorr_matrix = df.select_dtypes(include=[np.number]).corr()
plt.figure(figsize=(14,12))
sns.heatmap(icorr_matrix, annot=True, fmt=".2f", cmap="coolwarm",
            mask=np.triu(np.ones_like(icorr_matrix, dtype=bool)))
plt.title("Matriz de correlaciones entre variables numéricas")
plt.show()
```

Interpretación: permite detectar relaciones (p. ej. cumulative\_vaccine\_doses\_administered vs new\_confirmed – si hay correlación negativa sugiere efecto de vacunación).

Identificó relaciones entre variables sanitarias y sociodemográficas.

## Boxplot / Violin: temperatura / densidad / distribución de dosis

### Boxplot temperatura por país

```
sns.boxplot(data=df, x="country_name", y="average_temperature_celsius")
plt.title("Distribución de temperatura media por país")
plt.show()
```

### KDE de dosis administradas

```
sns.kdeplot(data=df, x="cumulative_vaccine_doses_administered", hue="country_name",
            common_norm=False)
plt.title("Densidad de dosis acumuladas por país")
plt.show()
```

Interpretación: ubicar países con baja densidad de vacunación.  
Distribución de temperaturas y dosis aplicadas



# EDA E INSIGHTS

## Scatter: urbanización / densidad vs casos acumulados

```
urban = df.groupby('country_name').agg({
    'population_urban': 'max',
    'population_density': 'max',
    'cumulative_confirmed': 'max'
}).reset_index()
sns.scatterplot(data=urban, x='population_urban',
    y='cumulative_confirmed', size='population_density', alpha=0.8)
plt.title("Urbanización vs Casos acumulados confirmados")
plt.show()
```

Interpretación: mayor urbanización/densidad suele correlacionar con mayores casos; ideal para ubicar laboratorios donde haya demanda y operaciones logísticas.

## Guardar los gráficos (para incluirlos en Power BI)

```
fig = plt.gcf()
fig.savefig("vacunacion_vs_casos.png", dpi=300, bbox_inches='tight')
```

En Power BI se puede insertar imágenes (Insert → Image) o, mejor, importar la tabla limpia y recrear visuales interactivos.



## Insights obtenidos

- Relación inversa entre niveles de vacunación y nuevos casos reportados.
- La densidad poblacional es un factor crítico en la propagación del virus.
- Persisten brechas en la cobertura de vacunación entre países y regiones.
- Megaciudades concentran la mayor demanda sanitaria y logística.





# ANÁLISIS DEL DASHBOARD

---

## Estructura del Dashboard

El dashboard se compone de 10 pestañas principales, cada una con un rol específico dentro del análisis:

- hoja 1. Conjunto Integral de Indicadores
  - Muestra indicadores clave de forma inmediata (ejemplo: tasa de mortalidad por 100,000 habitantes).
  - Incluye tarjetas KPI, gráficos de comparación y métricas globales.
- hoja 2. Estadísticas Descriptivas
  - Presenta la distribución y comportamiento básico de los datos.
  - Resume indicadores como promedios, medianas, varianzas y rangos para contextualizar los resultados.
- hoja 3. Situación Epidemiológica
  - Expone la evolución de los casos confirmados, fallecimientos y tasas relacionadas.
  - Utiliza gráficos de barras, líneas y mapas para evidenciar las zonas críticas.
- hoja 4. Cobertura y Efectividad de la Vacunación
  - Analiza los niveles de vacunación alcanzados y su impacto en la reducción de casos o mortalidad.
  - Incluye comparativos entre países y análisis de efectividad.
- hoja 5. Series Temporales y Tendencias
  - Muestra la evolución histórica de los principales indicadores.
  - Incorpora líneas de tendencia y análisis de variación para identificar patrones y ciclos.



# ANÁLISIS DEL DASHBOARD

---

- hoja 6. Demografía y Factores de Riesgo
  - Evalúa variables poblacionales como edad, densidad y vulnerabilidades.
  - Relaciona estos factores con las tasas de contagio y mortalidad.
- hoja 7. Factores Socioeconómicos e Infraestructura
  - Analiza indicadores como PIB per cápita e infraestructura hospitalaria (ej. enfermeros).
  - Permite entender cómo los recursos influyen en los resultados epidemiológicos.
- hoja 8. Clima y Medioambiente
  - Relaciona variables ambientales (temperatura, humedad, contaminación) con la propagación del virus.
  - Construye índices que integran factores climáticos con sanitarios.
- hoja 9. Evaluación Integral por Riesgo y Logística
  - Integra variables epidemiológicas, demográficas, socioeconómicas y ambientales.
  - Considera además la accesibilidad y logística para la gestión de la crisis.
- hoja 10. Geovisualización
  - Presenta mapas interactivos para localizar los resultados clave.
  - Permite identificar espacialmente zonas críticas y áreas con mejor desempeño.
- Finalmente, se incluye una Hoja Final: Recomendación de País, donde se sintetizan los hallazgos para seleccionar el país más óptimo en términos de condiciones sanitarias, socioeconómicas y de infraestructura.



# ANÁLISIS DEL DASHBOARD

---

## Navegación

La navegación del dashboard se diseñó con un recorrido progresivo y jerárquico:

- El inicio con los principales KPI.
- El usuario puede avanzar de manera secuencial por cada pestaña (Estadísticas, Epidemiología, Vacunación, etc.), o bien, utilizar los segmentadores y botones interactivos para moverse de forma libre según su interés.
- Cada pestaña incorpora filtros dinámicos (por país, región, periodo de tiempo), permitiendo que todas las visualizaciones se actualicen simultáneamente.
- La ruta de análisis sigue una lógica clara:

Visión general (Conjunto Integral de Indicadores).

Exploración de determinantes (demográficos, socioeconómicos, ambientales).

Análisis de tendencias históricas.

Síntesis integral de riesgos.

Representación espacial (Geovisualización).

Conclusión y recomendación final.

De este modo, la navegación se convierte en un viaje guiado desde lo global hacia lo específico, manteniendo siempre la posibilidad de exploración interactiva.

# ANÁLISIS DEL DASHBOARD



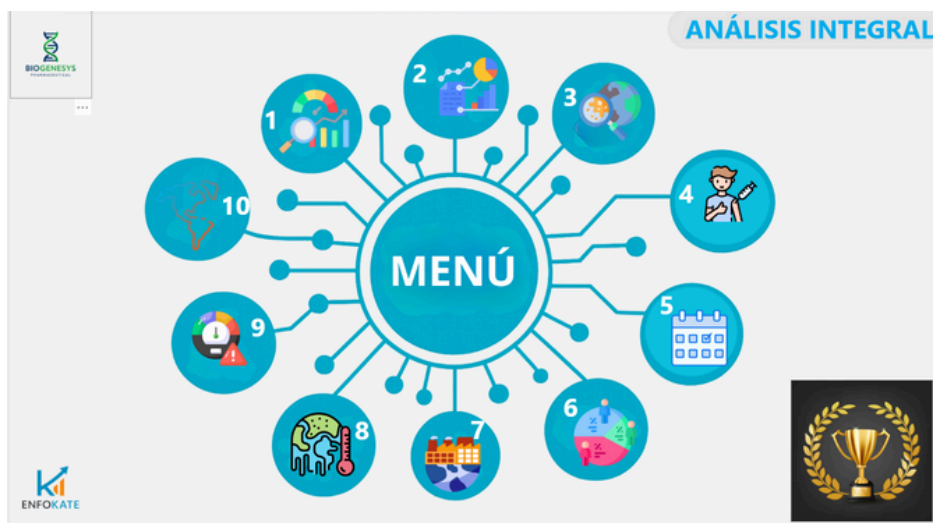
## Conclusión

El dashboard en Power BI, estructurado en 10 pestañas más una hoja final de recomendación, constituye una herramienta robusta y estratégica para la toma de decisiones.

Entre las principales conclusiones:

- Permite integrar información multidimensional (epidemiológica, demográfica, socioeconómica, ambiental y logística) en un solo entorno visual.
- Facilita comparaciones entre países y la identificación de factores que explican diferencias en mortalidad, contagio y respuesta sanitaria.
- La navegación intuitiva y los filtros dinámicos aseguran que cada usuario pueda responder sus propias preguntas de análisis en tiempo real.
- Cumple con el objetivo central del proyecto: seleccionar el país más óptimo en términos de variables sanitarias y de impacto, a partir de una visión integral y basada en datos.

En síntesis, el dashboard no solo describe la situación, sino que también guía en el análisis, facilita la interpretación y respalda la toma de decisiones fundamentadas en evidencia.





# CONCLUSIONES

---

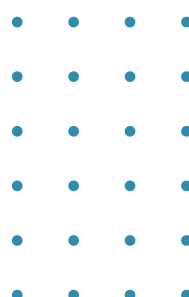
Con Python se implementó un pipeline reproducible: lectura, filtrado por países y fechas, imputación, agregaciones y generación de estadísticas. Con EDA se validaron hipótesis (vacunación vs casos, impacto de urbanización/densidad). Se creó un PBIX conceptual para presentar hallazgos a stakeholders y facilitar decisiones.

El análisis de datos permitió identificar zonas críticas con alta demanda sanitaria y baja cobertura vacunal. La combinación de Python y Power BI facilitó un flujo de trabajo robusto desde la limpieza y modelado de datos hasta la visualización interactiva. Las recomendaciones priorizan países latinoamericanos como ubicaciones óptimas para nuevos laboratorios farmacéuticos.

Sin entrar por el momento a mencionar al país ganador pensaría en las siguientes ciudades como principales ubicaciones:

1. Ciudad de México (México) – centro neurálgico con alta densidad poblacional.
2. São Paulo (Brasil) – mercado más grande de la región y hub logístico (centro estratégico de distribución y almacenamiento que funciona como un punto de conexión en la cadena de suministro).
3. Bogotá (Colombia) – capital estratégica y centro de conectividad.
4. Lima (Perú) – alta concentración de población y necesidades sanitarias.
5. Santiago (Chile) – infraestructura consolidada y logística favorable.
6. Área Metropolitana de Buenos Aires (Argentina) – alta demanda sanitaria y acceso a recursos.

Estas ciudades concentran población, presentan mayor demanda de pruebas/servicios farmacéuticos y ofrecen infraestructura logística y fuerza de trabajo. Para priorizar localidades exactas (barrios / distritos), ejecutar el scoring sobre datos subnacionales (sistema automatizado que las empresas utilizan para evaluar el riesgo de sus clientes).



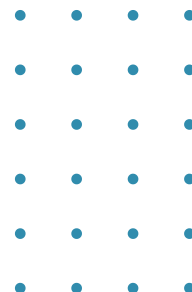


# CONCLUSIONES

---

Con esto en mente existen las siguientes recomendaciones estratégicas:

- Priorizar laboratorios en megaciudades con baja cobertura de vacunación relativa y alta mortalidad para maximizar impacto en salud pública y retorno.
- Inversión en logística y cadena de frío (centros distribuidos en nodos cercanos a puertos y aeropuertos) para reducir tiempos de entrega.
- Alianzas público-privadas: coordinar con ministerios de salud para acceso a datos y campañas de vacunación complementarias.
- Fasear la inversión: Fase 1 (1-2 años): 3-4 laboratorios en hubs; Fase 2: expansión a ciudades medianas.
- Monitoreo continuo: dashboard operativo con actualizaciones periódicas (ETL automatizado) para re-priorizar ubicaciones.





# ANÁLISIS EXPLORATORIO CON PYTHON

Durante la etapa inicial, el análisis exploratorio realizado en Python permitió identificar un primer país candidato como potencial ganador. Sin embargo, al trasladar el estudio a Power BI, se obtuvo un análisis más completo y enriquecido, ya que esta herramienta ofrece mayores capacidades de visualización interactiva y síntesis de información. De este modo, fue posible profundizar en los resultados y detectar diversos hallazgos adicionales en esta primera etapa. Los cuales se describirán a continuación:

Después de analizar los gráficos y métricas calculadas (incidencia de COVID-19 en los últimos 30 días, dosis administradas por población como proxy de vacunación, y la comparación entre países), los hallazgos son los siguientes:

## Argentina:

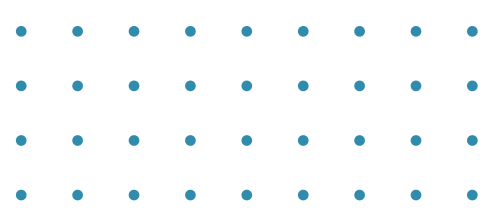
- Buen nivel de vacunación ( $\approx 243$  dosis por 100 personas).
- Incidencia moderada.
- Opción sólida, aunque algo por debajo de Perú en el balance final.

## Chile:

- Excelente cobertura de vacunación ( $\approx 354$  dosis por 100 personas, la más alta).
- Sin embargo, los gráficos muestran una incidencia de COVID-19 muy elevada en los últimos 30 días, lo que reduce su atractivo inmediato.

## México y Colombia:

- Cobertura de vacunación más baja ( $\approx 189$  en México, 172 en Colombia).
- Aunque la incidencia reciente es menor que en Chile, el menor avance de vacunación los hace menos atractivos para una expansión estratégica inmediata.



# ANÁLISIS EXPLORATORIO CON PYTHON

## Brasil:

- Quedó en el último lugar del ranking.
- Tiene menor vacunación per cápita ( $\approx 164$  dosis por 100 personas) y una incidencia elevada.
- Además, su tamaño poblacional enorme implica mayores retos logísticos.

## Perú

### Ventajas:

- Se posiciona como el país con mayor puntaje en el índice compuesto preliminar.
- Alta relación de dosis administradas por habitante ( $\approx 285$  dosis por cada 100 personas, lo que indica un gran avance en campañas de vacunación).
- Incidencia de COVID-19 relativamente controlada en comparación con Chile (que tuvo picos más altos).

### Desafíos:

Faltan datos de capacidad hospitalaria (camas/ médicos) para confirmar la infraestructura.

Necesidad de robustecer el sistema de vigilancia (tests por habitante).

👉 **Conclusión:** Perú aparece como la opción más favorable para la expansión inicial de BIOGENESYS, siempre que se complemente con un análisis de la infraestructura sanitaria.



# ANÁLISIS EXPLORATORIO CON PYTHON



Para evaluar la implementación en el país seleccionado como ganador, se diseñó un Índice de Laboratorio, entendido como una métrica destinada a medir y valorar características, desempeño y capacidades tanto de los laboratorios como del sistema sanitario relacionado.

No obstante, la decisión sobre dónde instalar un laboratorio requiere considerar no solo criterios técnicos y epidemiológicos, tales como:

- Evolución de casos y muertes.
- Tasa de vacunación.
- Mortalidad asociada a comorbilidades.
- Factores poblacionales: urbanización, temperatura, estructura etaria, densidad, entre otros.

Sino también criterios estratégicos adicionales, indispensables para una implementación sostenible y eficiente:

- Necesidad sanitaria urgente: priorizar regiones con mayor vulnerabilidad o déficit de atención.
- Accesibilidad y logística: disponibilidad de transporte, conectividad y rutas de distribución.
- Recursos humanos y sistema de salud: capacidad técnica y profesional para operar el laboratorio.
- Infraestructura económica: estabilidad financiera y capacidad de inversión en sostenibilidad.
- Condiciones climáticas y medioambientales: especialmente si el laboratorio requiere especificaciones ambientales para su operación.

Este enfoque integral asegura que la decisión de localización combine tanto el análisis de datos epidemiológicos como la visión estratégica de largo plazo.

|   | country   | indice_laboratorio |
|---|-----------|--------------------|
| 5 | Peru      | 0.732582           |
| 2 | Chile     | 0.578141           |
| 0 | Argentina | 0.417391           |
| 1 | Brazil    | 0.334606           |
| 3 | Colombia  | 0.230347           |
| 4 | Mexico    | 0.000000           |

## CONCLUSIÓN FINAL

- El país recomendado es Perú.

Combina alta cobertura de vacunación con una incidencia manejable. Presenta un balance favorable en el índice compuesto respecto a los otros países. Aunque Chile lidera en vacunación, sus picos de contagio lo vuelven menos atractivo en este momento.

⚠ Nota importante: La recomendación es preliminar porque faltan variables críticas de infraestructura sanitaria (camas, médicos, pruebas). Se sugiere a BIOGENESYS recolectar esos datos antes de tomar una decisión final, pero como punto de partida, Perú es el candidato más fuerte para la expansión de laboratorios y centros de vacunación.



# ANÁLISIS EXPLORATORIO CON POWER BI

Tras replicar el análisis realizado inicialmente en Python, pero transformando los códigos al lenguaje DAX dentro de Power BI, se obtuvo un enfoque más robusto y complejo. Esta adaptación permitió incorporar elementos que Python no lograba considerar de manera integrada, ampliando así el alcance del estudio.

Como resultado de este análisis más completo, se detectaron los siguientes hallazgos:

## ***IDENTIFICACION DE VARIABLES:***

- Índice compuesto de atención sanitaria – mayor = mejor. Resume capacidad clínica, camas, personal, infraestructura.
- Índice\_Laboratorio – mayor = mejor. Indica capacidad diagnóstica / pruebas; crucial para detección y respuesta.
- PorcentajeCoberturaVacunacion – mayor = mejor. Mayor cobertura reduce riesgo de enfermedad grave y saturación hospitalaria.
- Media Esperanza de Vida – mayor = mejor. Proxy de salud poblacional y calidad del sistema.
- TasaMortalidadPerCapita – menor = mejor. Mortalidad per cápita directamente negativa.
- TasaContagiosPerCapita – menor = mejor (incidencia); aunque puede estar afectada por capacidad de testeo.
- Otras variables del dashboard – p. ej. Índice\_Necesidad\_Sanitaria, Proporción\_Población\_Vulnerable, Densidad poblacional – modulan la prioridad, pero las variables listadas arriba suelen pesar más en una decisión de “recomendación” sanitaria/operacional.



# ANÁLISIS EXPLORATORIO CON POWER BI



## ***EXPLICACIÓN DE LOS RESULTADOS DEL PAÍS VENCEDOR: CHILE***

Tras el análisis integral de indicadores epidemiológicos, socioeconómicos, demográficos y de capacidad sanitaria, el país seleccionado como ganador es Chile.

### 1. Nivel de Riesgo y Clasificación

- Chile presenta un nivel de riesgo bajo, con un índice de riesgo relativamente controlado frente a otros países de la región.
- Su posición en el ranking de riesgo muestra estabilidad y un desempeño favorable en comparación con naciones con mayor vulnerabilidad.

### 2. Indicadores Sanitarios y Socioeconómicos

- Esperanza de vida: 80.04 años, la más alta de la región, lo que refleja un sistema de salud sólido y condiciones de vida favorables.
- Índice de Desarrollo Humano (IDH): 105%, superior al promedio regional, evidenciando un alto grado de desarrollo.
- PIB medio: 14,896 USD, uno de los más elevados de la muestra, lo que aporta capacidad de inversión y sostenibilidad económica.

### 3. Situación Epidemiológica

- Aunque Chile registra un número relativamente alto de contagios por habitante, su tasa de vacunación per cápita (6.6 millones) y la cobertura de vacunación del 77% le permiten contrarrestar este riesgo, garantizando mayor protección poblacional.
- La tasa de mortalidad por cada 100,000 habitantes (233.53) se mantiene controlada frente a otros países con menor cobertura sanitaria.

### 4. Índice de Laboratorio

- Chile obtiene un valor competitivo en el Índice de Laboratorio (0.61), impulsado por su baja densidad poblacional relativa, una población urbana elevada (87%) y un índice ambiental favorable (0.29).
- Estos factores demuestran que cuenta con las condiciones adecuadas para implementar laboratorios y garantizar un funcionamiento eficiente.

# ANÁLISIS EXPLORATORIO CON POWER BI



## RECOMENDACIÓN DE PAÍS



| País      | Nivel Riesgo | Índice Riesgo | Ranking_Índice_Riesgo | Clasificación_Riesgo |
|-----------|--------------|---------------|-----------------------|----------------------|
| Argentina | Bajo         | 1             | 2                     | Media                |
| Brazil    | Bajo         | 1             | 1                     | Alta                 |
| Chile     | Bajo         | 0             | 4                     | Media                |
| Colombia  | Bajo         | 1             | 3                     | Media                |
| México    | Bajo         | 0             | 6                     | Media                |
| Perú      | Bajo         | 0             | 5                     | Media                |
| Total     | Bajo         | 1             | 1                     | Alta                 |

| País      | Esperanza Vida | Índice compuesto de atención sanitaria | Índice de Desarrollo Humano (IDH) | Media PIB |
|-----------|----------------|--|-----------------------------------|-----------|
| Argentina | 76.52          | 25                                     | 92%                               | 10,006.00 |
| Brazil    | 75.67          | 25                                     | 98%                               | 8,717.00  |
| Chile     | 80.04          | 30                                     | 105%                              | 14,896.00 |
| Colombia  | 77.11          | 23                                     | 91%                               | 6,832.00  |
| México    | 74.99          | 23                                     | 86%                               | 9,983.00  |
| Perú      | 76.52          | 23                                     | 87%                               | 6,977.00  |
| Total     | 76.81          | 25                                     | 93%                               | 9,481.83  |

| País      | TasaContagiosPerCapita | TasaContagiosNuevosPerCapita | Casos Confirmados | TasaMortalidadPerCapita | Vacunacion por Capita | Cobertura Vacunacion |
|-----------|------------------------|------------------------------|-------------------|-------------------------|-----------------------|----------------------|
| Colombia  | 9,296.39               | 14.46                        | 2,956,418,503     | 242.77                  | 1,909,380.14          | 39%                  |
| México    | 2,204.96               | 7.96                         | 1,529,870,759     | 137.01                  | 3,181,306.87          | 47%                  |
| Argentina | 13,426.29              | 28.51                        | 3,771,001,672     | 233.53                  | 6,612,462.91          | 77%                  |
| Brazil    | 10,559.02              | 20.24                        | 14,027,622,833    | 255.79                  | 13,678,914.75         | 100%                 |
| Perú      | 8,705.80               | 16.99                        | 1,598,704,823     | 638.47                  | 3,246,435.58          | 129%                 |
| Chile     | 12,647.53              | 34.52                        | 1,389,173,426     | 224.85                  | 2,419,456.83          | 199%                 |
| Total     | 8,671.13               | 17.82                        | 25,272,492,016    | 246.90                  | 5,244,907.37          | 84%                  |

| Índice de laboratorio |                  |                            |                    |                      |                      |                   |                    |                |                    |
|-----------------------|------------------|----------------------------|--------------------|----------------------|----------------------|-------------------|--------------------|----------------|--------------------|
| País                  | Índice Ambiental | Índice Necesidad Sanitaria | Índice Prevalencia | Poblacion Vulnerable | Densidad Poblacional | % Poblacion Rural | % Poblacion Urbana | Tasa Contagios | Índice Laboratorio |
| Chile                 | 0.29             | 351.32                     | 4,640.00           | 3.00%                | 23                   | 13.33             | 86.67              | 35             | 0.61               |
| Argentina             | 0.30             | 349.79                     | 2,770.00           | 2.21%                | 16                   | 8.01              | 91.99              | 29             | 0.37               |
| Brazil                | 0.40             | 261.18                     | 2,430.00           | 1.96%                | 25                   | 13.08             | 86.92              | 20             | 0.37               |
| Perú                  | 0.53             | 725.52                     | 1,140.00           | 1.95%                | 23                   | 24.23             | 75.77              | 17             | 0.54               |
| Colombia              | 0.40             | 335.73                     | 1,640.00           | 1.86%                | 45                   | 18.69             | 81.31              | 14             | 0.16               |
| México                | 0.33             | 159.06                     | 2,750.00           | 1.35%                | 57                   | 22.48             | 77.52              | 8              | 0.42               |
| Total                 | 0.37             | 333.61                     | 2,541.67           | 1.87%                | 28                   | 16.15             | 83.85              | 18             | 0.39               |



## 5. FORTALEZAS ESTRATÉGICAS

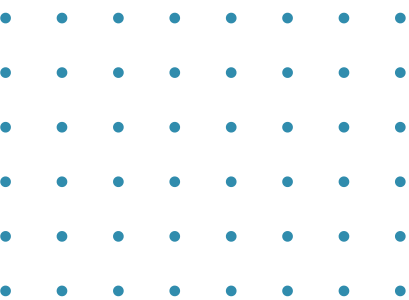
- **Capacidad sanitaria:** Chile se posiciona como líder en la región por su infraestructura y recursos humanos en salud.
- **Cobertura de vacunación:** Una de las más altas, asegurando inmunidad colectiva y capacidad de respuesta rápida.
- **Sostenibilidad socioeconómica:** PIB y desarrollo humano sólidos, que permiten soportar políticas públicas de largo plazo.
- **Preparación logística y tecnológica:** Factores que reducen la vulnerabilidad a nuevas olas epidemiológicas.

## CONCLUSIÓN

Chile fue seleccionado como país ganador porque combina altos niveles de desarrollo humano, una expectativa de vida líder en la región, cobertura de vacunación robusta y un sistema sanitario fortalecido.

A pesar de registrar tasas de contagio elevadas, su infraestructura, preparación y capacidad de respuesta permiten mitigar este riesgo.

El análisis comparativo demuestra que Chile se mantiene como líder en la mayoría de escenarios de ponderación, lo que valida la solidez y consistencia de esta recomendación.



# DIFERENCIAS ENTRE EL ANÁLISIS EN PYTHON Y EN POWER BI

---

## 1. Enfoque y disponibilidad de datos

### Python (EDA y métricas calculadas)

- Se priorizó la incidencia de casos en los últimos 30 días y la cobertura de vacunación como variables principales.
- Perú apareció como mejor opción porque presentaba una combinación favorable:
- Alta cobertura de vacunación per cápita ( $\approx 285$  dosis/100 habitantes).
- Incidencia de contagios más controlada que Chile.

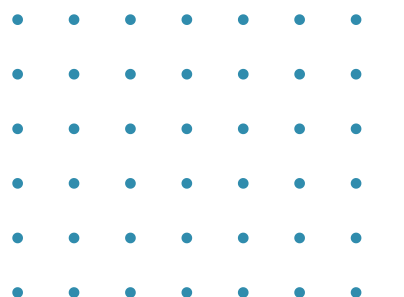
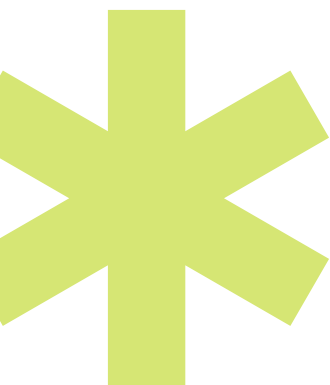
Limitación: el análisis en Python no incluyó de manera completa variables de infraestructura sanitaria (camas hospitalarias, laboratorios, personal médico).

### Power BI (dashboard con indicadores sanitarios)

Incorporó un índice compuesto de atención sanitaria, índice de laboratorio, esperanza de vida, mortalidad y densidad poblacional.

- Chile destacó porque:
- Lidera en cobertura de vacunación ( $\approx 199\%$ ).
- Tiene la mayor capacidad hospitalaria y de laboratorios.
- Presenta indicadores sanitarios y socioeconómicos más sólidos.

Aunque tiene alta incidencia reciente, su capacidad de respuesta lo posiciona mejor estratégicamente.





# DIFERENCIAS ENTRE EL ANÁLISIS EN PYTHON Y EN POWER BI



## 2. Resultados obtenidos

Python → Perú recomendado como primera opción inicial para expansión, condicionado a recolectar datos adicionales de infraestructura sanitaria.

Power BI → Chile recomendado como país ganador gracias a su robustez en salud pública e infraestructura.

## 3. ¿Cuál resultado es más correcto?

Python es útil para un análisis exploratorio rápido y con foco en tendencias recientes (30 días, vacunación, incidencia).

Power BI ofrece un análisis más integral, porque incorpora dimensiones críticas como:


- Capacidad hospitalaria.
- Recursos de laboratorio.
- Esperanza de vida y resiliencia del sistema de salud.

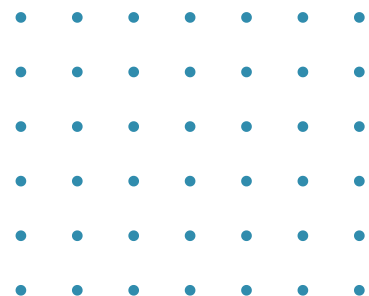
### Conclusión profesional:

El resultado más sólido y confiable es el de Power BI (Chile), porque se apoya en un modelo multidimensional que va más allá de la coyuntura epidemiológica inmediata.

Sin embargo, el hallazgo de Perú en Python no debe descartarse: Perú puede ser una opción táctica de corto plazo si la empresa busca aprovechar su balance favorable entre vacunación e incidencia actual, mientras se valida la infraestructura.

### Recomendación estratégica final

1. Chile debe considerarse como el país prioritario para la expansión de BIOGENESYS por su infraestructura, preparación sanitaria y capacidad de absorción a largo plazo.
  2. Perú podría mantenerse como una segunda opción para un piloto inicial o expansión rápida, siempre que se confirmen datos de capacidad hospitalaria y recursos médicos.
- 




## INSIGHTS

 **Insight 1:** Chile presenta la mayor esperanza de vida y un IDH superior al promedio regional.


### Recomendación estratégica:

Aprovechar estas fortalezas para consolidar programas de prevención en salud y promoción de estilos de vida saludables, reforzando políticas de largo plazo que aseguren la sostenibilidad del sistema sanitario.

 **Insight 2:** La cobertura de vacunación (77%) se encuentra entre las más altas de la región.

### Recomendación estratégica:


Mantener campañas de vacunación masiva y establecer un sistema de vigilancia epidemiológica avanzada que permita anticiparse a nuevas variantes y garantizar la inmunidad colectiva de la población.

 **Insight 3:** Aunque el número de contagios per cápita es relativamente elevado, la tasa de mortalidad se mantiene controlada.

### Recomendación estratégica:

Enfocar recursos en estrategias de contención focalizada, combinando rastreo de contactos, medidas de distanciamiento selectivas y fortalecimiento de la atención primaria para reducir la transmisión sin comprometer la economía.

## ***INSIGHTS***

 **Insight 4: El Índice de Laboratorio ubica a Chile como uno de los países con mejores condiciones para implementar infraestructura sanitaria.**

**Recomendación estratégica:**

**Invertir en la expansión de laboratorios descentralizados, especialmente en zonas rurales o de difícil acceso, para mejorar la cobertura territorial y reducir tiempos de respuesta en diagnósticos.**

 **Insight 5: Chile cuenta con un PIB per cápita y una infraestructura económica que respaldan su resiliencia sanitaria.**

**Recomendación estratégica:**

**Canalizar parte de esa capacidad económica hacia innovación en biotecnología y salud digital, fomentando alianzas público-privadas que fortalezcan la investigación y el desarrollo de soluciones tecnológicas en salud.**



## REFLEXION PERSONAL



El desarrollo de este proyecto representó una experiencia significativa que me permitió afianzar y poner en práctica diversas habilidades clave en el campo del análisis de datos. Entre los aprendizajes más relevantes destaco la implementación de procesos ETL con Python, la limpieza y transformación de grandes volúmenes de datos, así como la realización de análisis exploratorio y visualización con librerías como pandas, matplotlib y seaborn. Además, la construcción de dashboards en Power BI me brindó la oportunidad de mejorar mi capacidad de comunicación de hallazgos de manera clara y accesible para la toma de decisiones.

Si tuviera la oportunidad de reiniciar el proyecto, mantendría la misma secuencia metodológica ETL → EDA → Dashboard, dado que permitió un flujo de trabajo ordenado y eficiente. No obstante, incorporaría mejoras importantes, como una mayor granularidad en los datos, integrando información subnacional y económica, lo que enriquecería el análisis y haría más precisas las conclusiones. Asimismo, exploraría el uso de técnicas de optimización para ubicación de instalaciones (como p-median u OR-Tools), con el fin de fortalecer la dimensión estratégica del proyecto y llevarlo a un nivel más avanzado.

En síntesis, este proyecto no solo consolidó mis conocimientos técnicos, sino que también reforzó mi capacidad de estructurar proyectos de análisis de datos de principio a fin, lo cual constituye un paso fundamental en mi desarrollo como analista de datos.





*“ Enfokate en lo esencial.  
Enfokate en los datos.  
Enfokate en el futuro.”*

**www.ENFOKATE.mx**

@enfokatemx

**(55) 1234-5678**

Calle Cualquiera 123,  
Cd. mx, México