

Expansión Estratégica de Biogenesys con Python

Nombre del autor: Lorena Maza

Email: lorenavmaza@yahoo.es

Cohorte: DA-PT08

Fecha de entrega: 25/08/25

Institución: Empresa farmacéutica **BIOGENESYS**

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.



Desarrollo del proyecto

Carga y transformación de datos

1. Lectura del archivo data_latinoamerica.csv con código Python en tu Visual Studio Code:
2. Importación de librerías Pandas, Numpy, Seaborn, Matplotlib.pyplot
3. Filtrado de los datos
 - a. Selecciona los países donde se expandirán: Colombia, Argentina, Chile, México, Perú y Brasi
 - b. Se observa una dependencia jerárquica en los códigos
 - c. Con la dependencia demostrada, decido filtrar por location key
 - d. Sobre un total de 12.216.057 filas, quedan 5.946 luego del filtrado por país
4. Filtrado los datos en fechas mayores a 2021-01-01
 - a. Utilizando un bucle For, completo la lista de columnas vacías (por location_key):

Módulo 4

- b. con un promedio de los valores para las variables climáticas: 'rainfall_mm', 'relative_humidity', 'average_temperature_celsius', 'minimum_temperature_celsius', 'maximum_temperature_celsius'
 - c. con el valor anterior para los acumulados: 'cumulative_confirmed', 'cumulative_deceased', 'cumulative_recovered', 'cumulative_vaccine_doses_administered'
 - d. con cero: 'new_confirmed', 'new_deceased', 'new_recovered'
5. Exportación de los datos limpios y filtrados
 6. Importación de DataSet nuevo (1,47 mb)
 7. Revisión de las Estadísticas por distintos métodos
 8. Indexación de columna Date para series temporales

EDA e insights

[Deberás detallar los insights obtenidos en el Análisis Exploratorio de Datos y visualizaciones]

1. Generación de Mapa de Correlación en donde se observa que la administración de vacunas tiene una alta correlación (0.94) con el acumulado de recuperados
2. Generación de Diagramas de barras por país

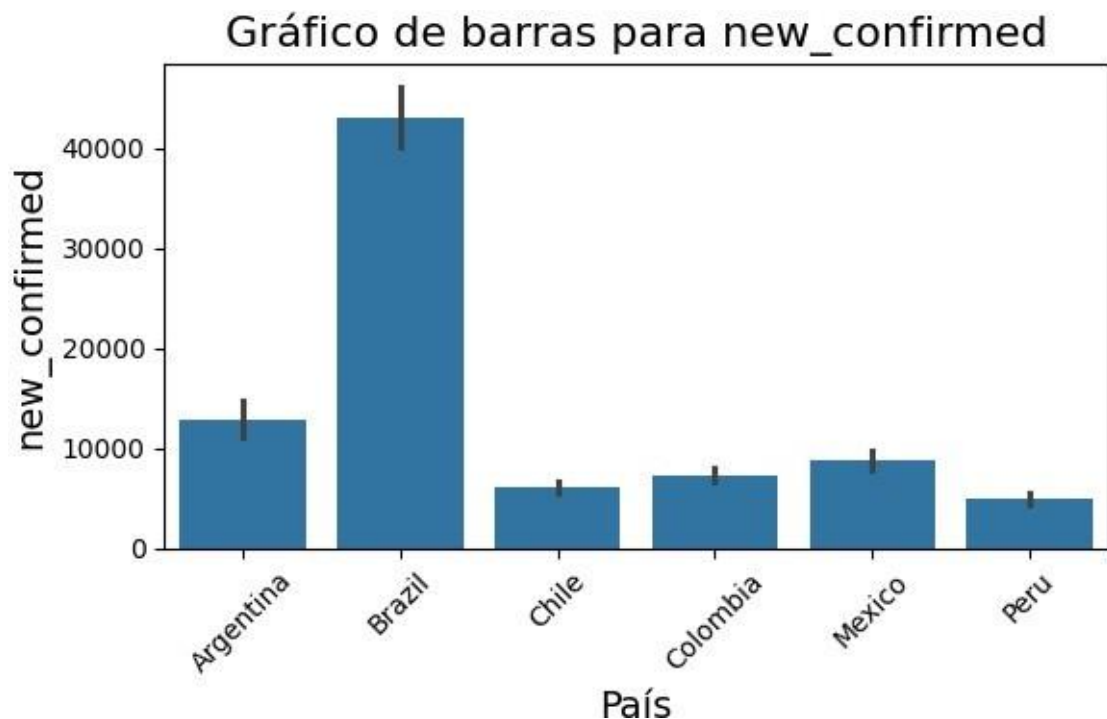


Gráfico de barras para new_deceased

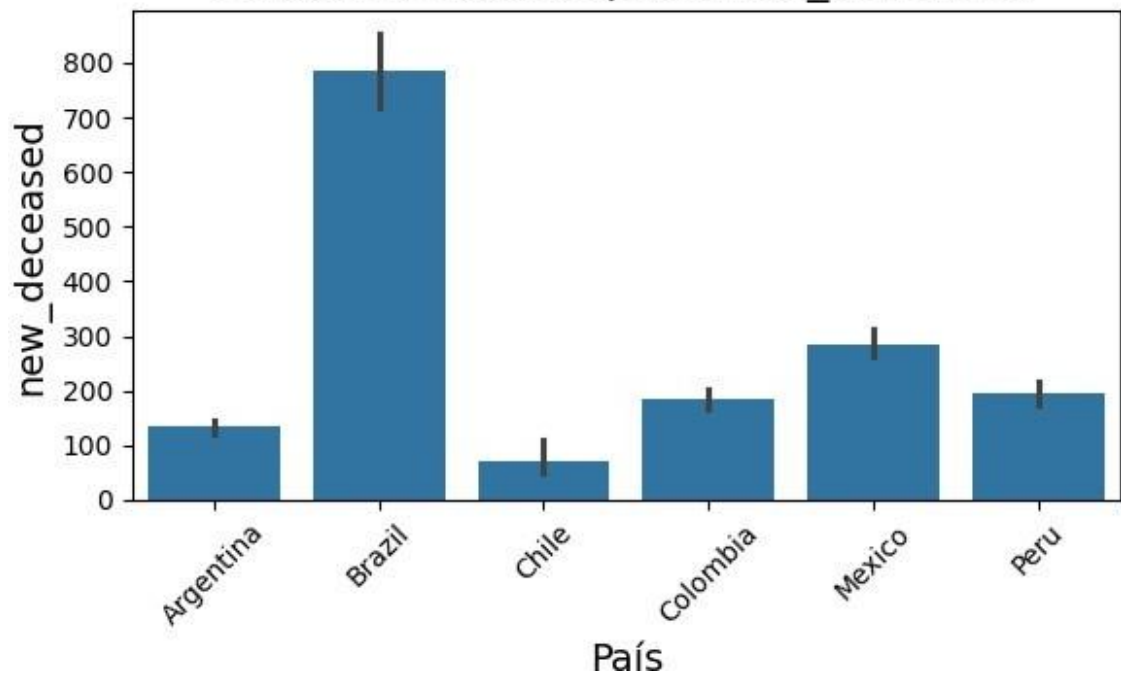
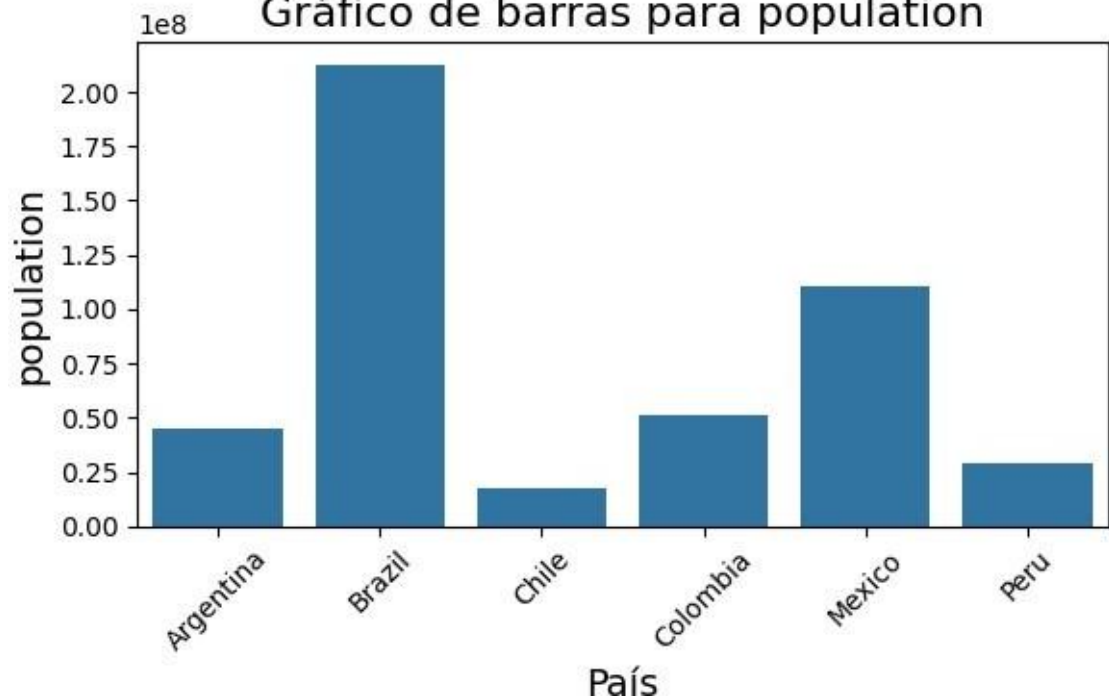
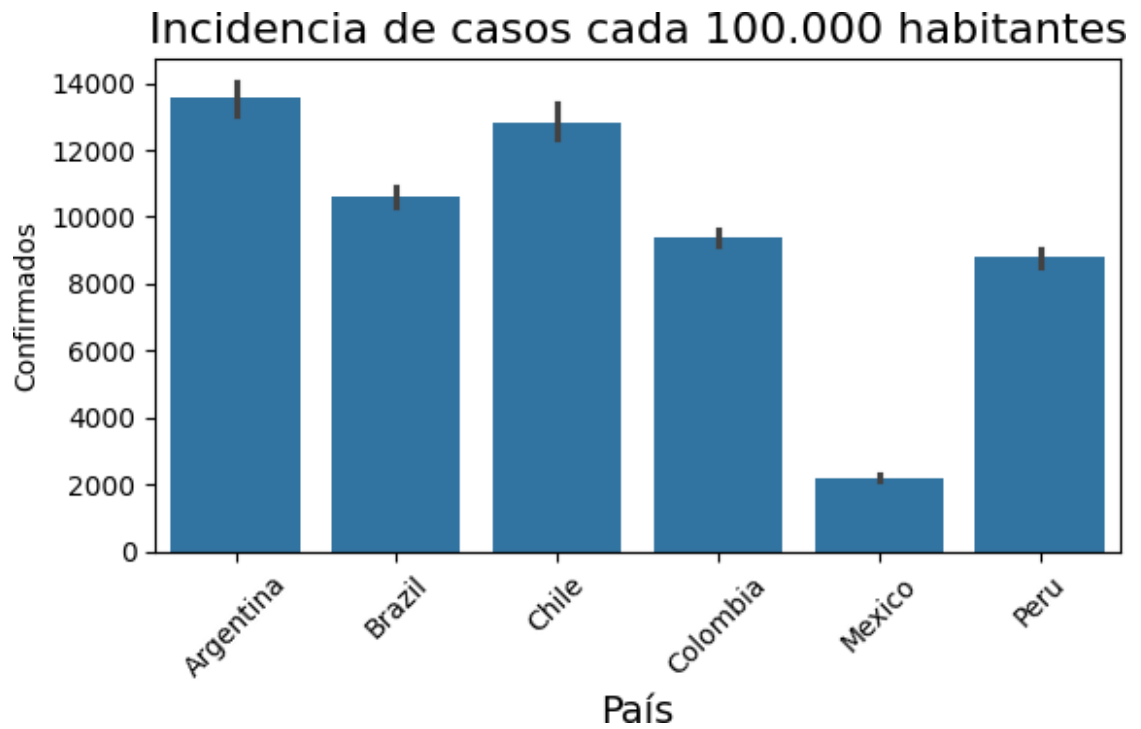


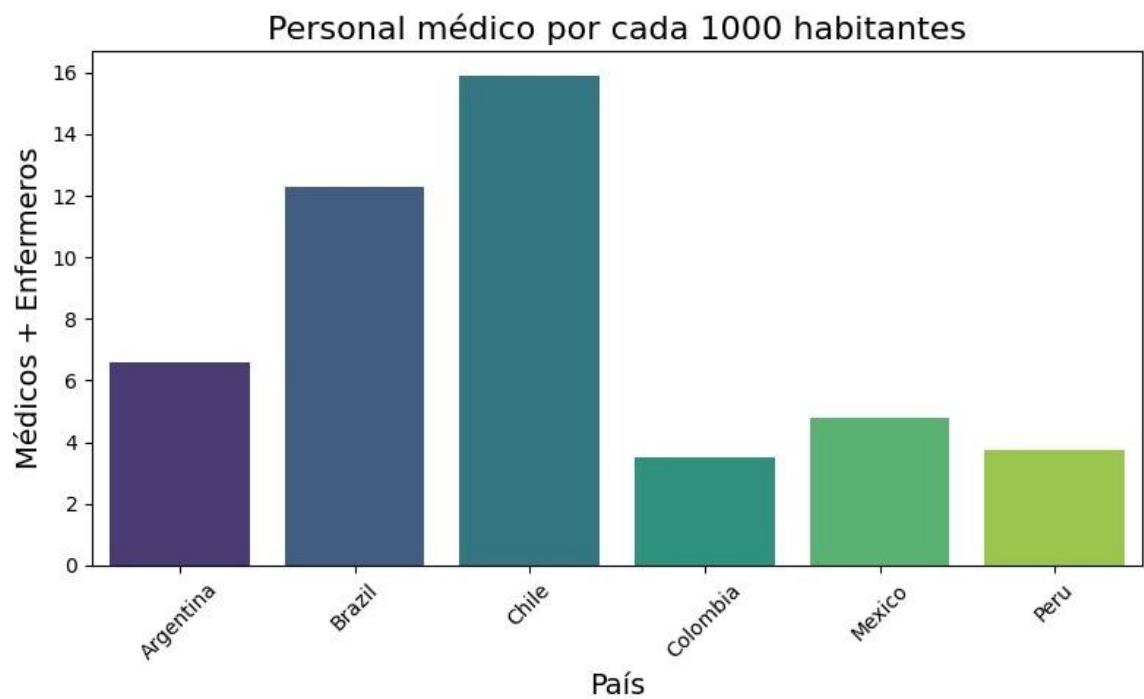
Gráfico de barras para population



Dada la diferencia considerable en la población entre países, no se puede establecer a partir de los gráficos, una comparativa entre los mismos. Decido continuar el análisis calculando tasas según la población de cada país.

Módulo 4

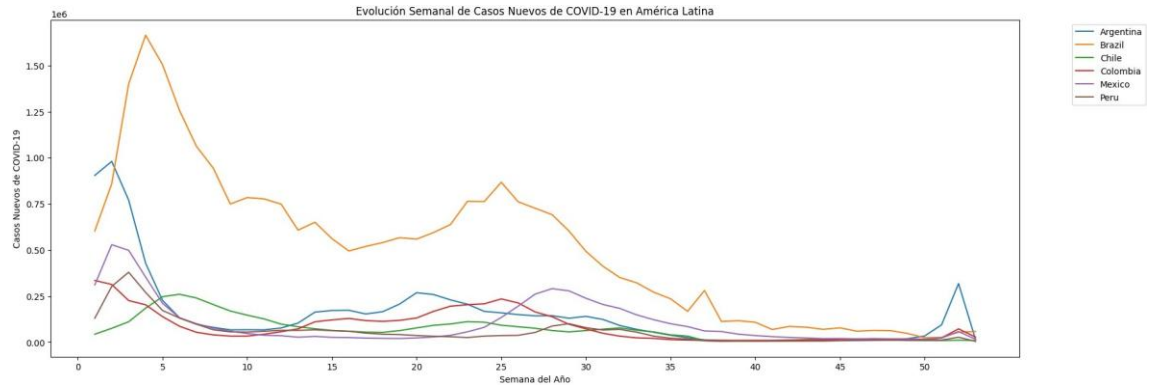
3. Cálculo de cuerpo médico total por países



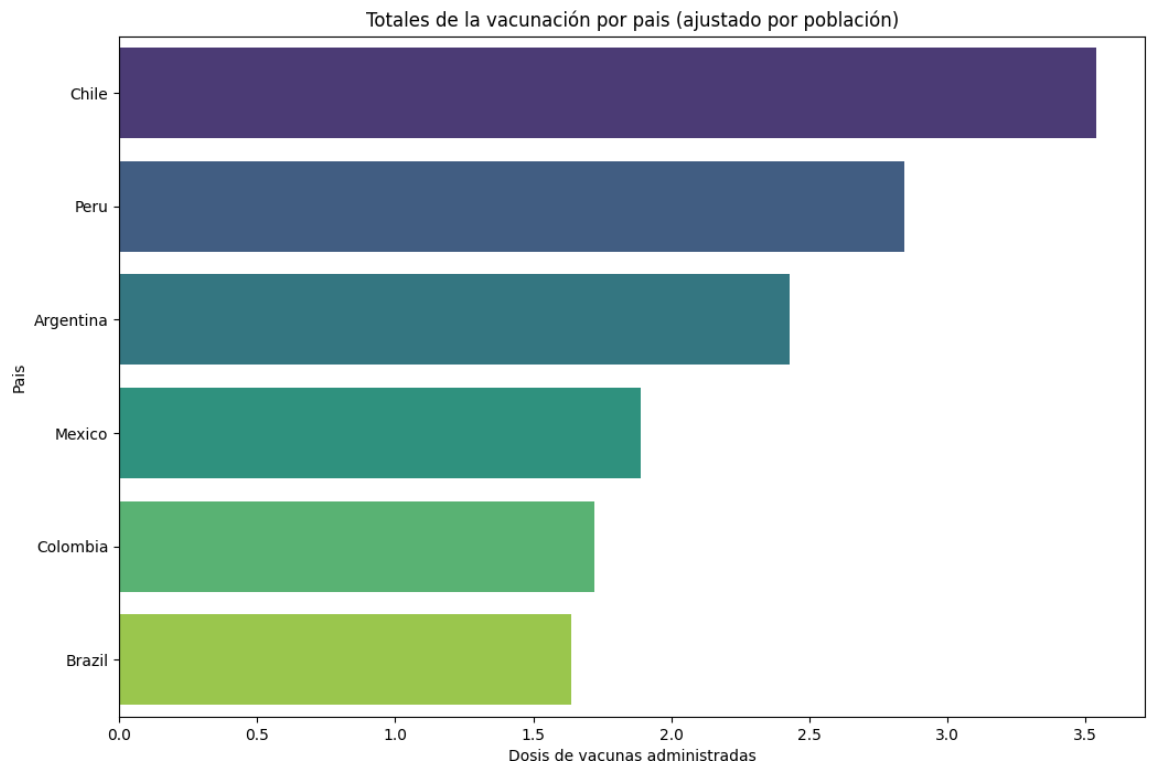
4. La exploración de variables ambientales no reporta información relevante.

Análisis temporal de casos y vacunas

1. Análisis de la evolución de casos por semana

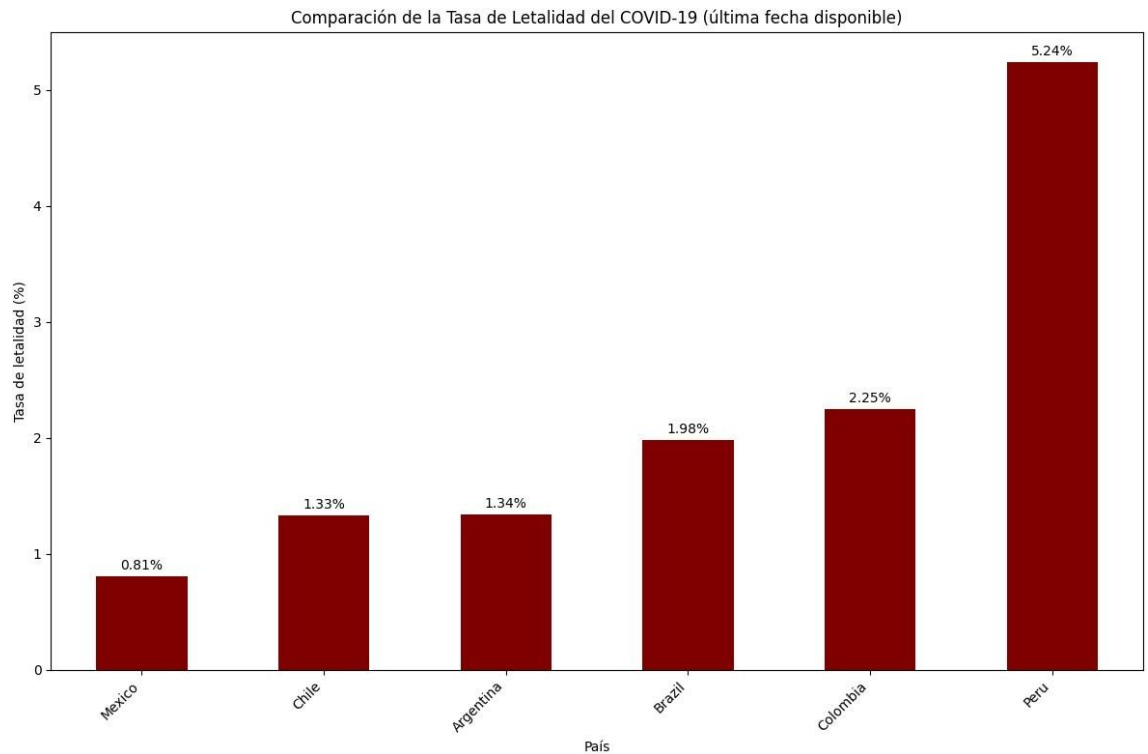
Módulo 4

2. Mientras a mayoría de los países muestran olas hacia abajo, Brasil obtuvo un pico alrededor de la semana 5. Argentina presenta un rebrote luego de la semana 50.
3. Cálculo de totales de vacunación per cápita



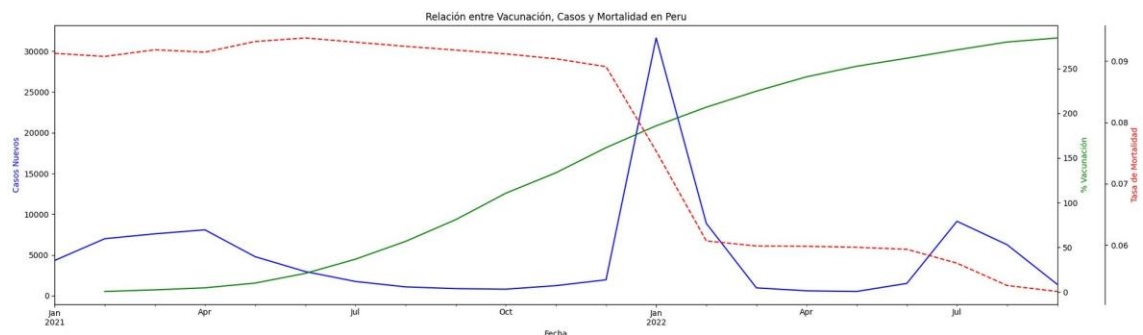
Chile y Perú dominan en la cantidad de vacunados según su población, mientras que Brasil queda en inferioridad.

4. Análisis de tasa de letalidad

Módulo 4

Pese a que Perú tiene una alta dosis de vacunas suministradas, la tasa de letalidad en este país es muy alta. Aun así la cantidad de casos fue disminuyendo a lo largo de las semanas. Según los gráficos desarrollados anteriormente, Perú no muestra una alta densidad poblacional. Sin embargo su cobertura total del cuerpo médico es de las más bajas. Aun así la cantidad de casos fue disminuyendo a lo largo de las semanas.

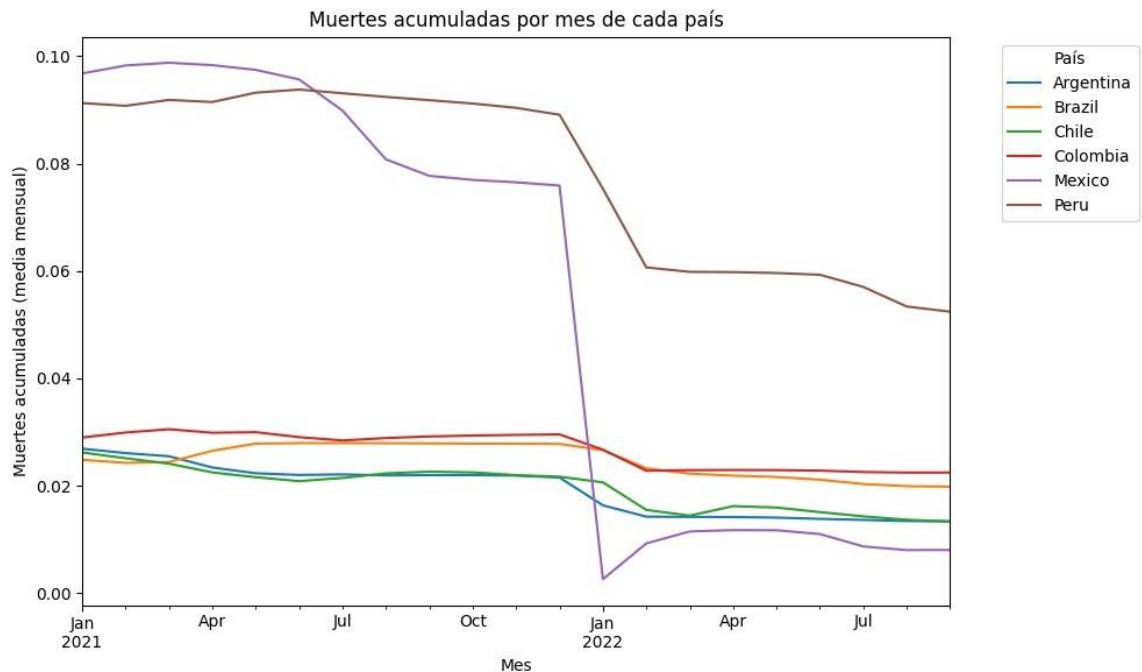
5. Búsqueda de una relación entre las vacunas aplicadas y los resultados en cuanto a disminución de casos y muertes.



A Perú le costó controlar los casos de muerte en toda la fase inicial. Parece haber mayor número de casos también frente a los otros países.

6. Ratio de mortalidad por país cada mes

Módulo 4



7. El análisis de enfermedades pre existentes tampoco arrojó resultados significativos.

Análisis del dashboard

Al momento de la entrega del informe, el dashboard se encuentra en construcción.

En principio cuenta con una página de presentación y una navegación de 2 páginas con información global y de vacunación.

Conclusiones y Recomendaciones

Recomendación de expansión (priorizada)

1) CHILE — Prioridad alta

- Muy alta vacunación per cápita; lidera el grupo con Perú.
- Alta incidencia acumulada por 100k (junto con Argentina) → mercado con historial de demanda y monitoreo/diagnóstico sostenido.
- Mayor dotación total de personal de salud por 1.000 hab. (suma médicos+enfermeros) → buena capacidad de adopción/implementación.

2) ARGENTINA — Prioridad alta

- Incidencia por 100k entre las más altas → necesidad histórica + sensibilidad del sistema a soluciones de vigilancia y terapias.
- Fuerte dotación de médicos (aunque menos enfermería que BR/CL) → capacidades clínicas para productos especializados.
- Rebotes tardíos (semana 50) → interés por soluciones de preparación/prevencción y testeo estacional.

3) BRASIL — Prioridad media-alta

- Gran mercado y curva de muertes más alta en el grupo (históricamente) → oportunidad para portafolios hospitalarios/UTI, antiinfecciosos, soporte crítico.

Módulo 4

- Fuerte cobertura de enfermería; total de personal supera a AR (aunque por debajo de CL) → personal operativo para despliegue.
- Vacunación per cápita por debajo de CL/PE → espacio para prevención complementaria/diagnóstico.

4) MÉXICO — Prioridad media (potencial grande, revisar data)

- Mercado grande; pero con data dudosa en junio 2022 (posible error de registro) → antes de invertir fuerte, conviene depurar/validar series con fuentes locales.
- Aun así, por escala poblacional y necesidades, es candidato para comercial tras validación de datos.

5) PERÚ — Prioridad selectiva (alto impacto en salud + alianzas)

- Vacunación per cápita alta, pero **tasa de letalidad muy alta** → foco en productos/servicios que reduzcan mortalidad (diagnóstico temprano, soporte crítico, manejo de comorbilidades).
- Dotación total de personal de salud entre las más bajas del grupo → entrar vía programas de acceso/educación y soluciones fáciles de escalar.
- Mayoría de la población es urbana → logística de despliegue más sencilla que si fuera predominantemente rural.

6) COLOMBIA — Prioridad observación

- Menos señales diferenciales en el análisis actual.

Justificación

- **Vacunación vs. resultados:** fuertísima correlación entre vacunas administradas y recuperados ($r \approx 0,94$) → los mercados con mejor cobertura y reporting tienden a sostener inversión en prevención/diagnóstico, y a adoptar innovaciones.
- **Incidencia histórica por 100k:** Argentina y Chile arriba → mayor conciencia/experiencia del sistema con COVID y síndromes respiratorios; buen terreno para soluciones de vigilancia/diagnóstico.
- **Capacidad del sistema:** Chile lidera en personal total por 1.000 hab.; Brasil y Chile fuertes en enfermería; Argentina más médica → define el mix de productos/soporte requerido por país.
- **Letalidad:** Perú destaca por letalidad alta pese a buena vacunación per cápita → oportunidad de impacto sanitario con soluciones que ataquen “última milla” (detección oportuna, manejo de agudos, comorbilidades).
- **Curvas de muertes:** Brasil con curva elevada; México con anomalía puntual → priorizar BR en portafolio hospitalario y auditar MX antes de escalar.

Reflexión personal

Fue un proyecto muy complejo. Con poco tiempo para asimilar todos los conceptos.

En cuanto al rol del analista fue interesante posicionarse en un rol más asociado a la investigación y mejoras a nivel social que dentro de una empresa de retail.

Carrera: Data Analytics

Módulo 4

EXTRA CREDIT



Por falta de tiempo no pude incorporar ExtraCredits.