



**RIDUNAJ**  
Repositorio Institucional  
Digital UNAJ



Universidad Nacional  
**ARTURO JAURETCHE**

## Práctica Profesional Supervisada

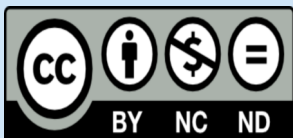
Nuzzo, Julián

# Aplicación web interactiva para la consulta y visualización de datos de salud

*Instituto de Ingeniería y Agronomía*

2025

*Carrera: Ingeniería en Informática*



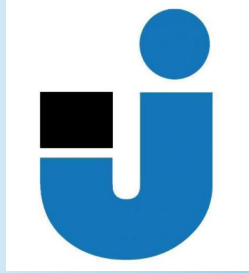
Esta obra está bajo una Licencia Creative Commons.  
Atribución – No comercial – Sin obra derivada 4.0  
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Documento descargado de RID - UNAJ Repositorio Institucional Digital de la Universidad Nacional Arturo Jauretche

Cita recomendada:

Nuzzo, J. (2025). *Aplicación web interactiva para la consulta y visualización de datos de salud* [Práctica Profesional Supervisada, Universidad Nacional Arturo Jauretche].

<https://rid.unaj.edu.ar/handle/123456789/3613>



**PRÁCTICA PROFESIONAL SUPERVISADA**  
**Informe final**

*Aplicación web interactiva para la consulta y visualización  
de datos de salud*

**Julián Nuzzo**

**Florencio Varela, mayo 2025**

Universidad Nacional Arturo Jauretche

Instituto de Ingeniería y Agronomía

Ingeniería en Informática



TRABAJO FINAL DE LA PRÁCTICA  
PROFESIONAL SUPERVISADA

*Aplicación web interactiva para la consulta y  
visualización de datos de salud*

Estudiante:

Nuzzo Julián

Tutores:

Sebastián Martin Gonzalez

Bein Paula

Martín Morales

**Florencio Varela, 2025**

## DATOS DEL ESTUDIANTE

Nombres y Apellidos	Nuzzo Julián
DNI	41.209.868
Correo electrónico	julinuzzo19@gmail.com
Carrera	Ingeniería en Informática

PRÁCTICA PROFESIONAL SUPERVISADA enmarcada en el artículo 4, punto 2 inciso “a” de la Resolución (CS) 123/24

## DOCENTE SUPERVISOR

Nombres y Apellidos	Sebastián Martin Gonzalez
Correo electrónico	sgonzalez@unaj.edu.ar

## DATOS DE LA ORGANIZACIÓN DONDE SE REALIZA EL PIP

Nombre o Razón Social	Universidad Nacional Arturo Jauretche
CUIT	30-71134016-1
Dirección	Av. Calchaquí 6200, B1888 Florencio Varela, Provincia de Buenos Aires
Teléfono	11 4275-6100
Sector	Programa Tecnologías de la Información y la Comunicación (TIC) en Aplicaciones de Interés Social

## TUTOR ORGANIZACIONAL

Nombres y Apellidos	Martín Morales
Correo electrónico	martin.morales@unaj.edu.ar

## Resumen

En la era de la información, los datos se han convertido en uno de los recursos más valiosos para la toma de decisiones, más aún en áreas tan sensibles como la salud pública, donde los indicadores en este sector son herramientas fundamentales que permiten identificar problemas al analizar tendencias a lo largo del tiempo y comparar la situación entre diferentes regiones o países. Su importancia radica en que proporcionan información crítica para la toma de decisiones en políticas de salud y la asignación de recursos.

El proyecto refuerza el rol crítico del desarrollador, este no solo realiza la tarea de codificación de la herramienta siguiendo procedimientos claros, sino que es un traductor de datos complejos en experiencias útiles para el usuario final. Las decisiones tomadas a lo largo del desarrollo van desde la selección del stack tecnológico, el procesamiento de los datos mediante técnicas de normalización y predicciones, hasta la elección de visualizaciones capaces de adaptarse a distintos escenarios y el diseño del dashboard. Estas decisiones no fueron sólo técnicas, sino estratégicas: orientadas a construir una experiencia coherente, clara y centrada en el usuario.

## Abstract

In the information age, data has become one of the most valuable resources for decision-making—especially in sensitive areas such as public health, where indicators serve as fundamental tools to identify problems by analyzing trends over time and comparing situations across different regions or countries. Their importance lies in the critical information they provide for shaping health policies and allocating resources.

This project reinforces the critical role of the developer—not merely as a coder who follows clear procedures, but as a translator of complex data into meaningful, user-friendly experiences. The decisions made throughout the development process ranged from selecting the technology stack, processing data through normalization and prediction techniques, to choosing visualizations that can adapt to various scenarios and designing the dashboard. These decisions were not only technical, but also strategic—aimed at building a coherent, clear, and user-centered experience.

# Índice

<b>Resumen</b>	<b>4</b>
<b>Abstract</b>	<b>5</b>
<b>Índice</b>	<b>6</b>
<b>Introducción</b>	<b>8</b>
Tecnologías Utilizadas en el Desarrollo	10
Visualización de Mapas	11
Gráficos con D3.js	11
Estilos del dashboard con Tailwind	12
Integración y Uso de la API del Banco Mundial en la Aplicación	12
Consulta de todos los indicadores	13
Obtención de Regiones y Países	14
Consulta de indicador seleccionado	14
Metadata del indicador	15
Obtención del Rango Temporal Disponible	16
Procesamiento de datos	18
Normalización	18
Métodos de Normalización utilizados	18
Criterios para la obtención de valores representativos de Indicadores en intervalos de tiempo	19
Técnicas de predicción	20
Técnicas de regresión utilizadas	22
Patrones de Crecimiento Epidemiológico: De la Exponencial a la Logística (Caso COVID-19)	25
Lógica de selección de técnica	26
Descripción y uso del dashboard	29
1. Selección del Indicador de Salud	31
Componentes	31
Selector de tópicos (opcional)	31
Buscador	31
Metadata del indicador seleccionado	32
2. Filtros de búsqueda adicionales	33
Componentes	34
Selector de años ("Desde - Hasta")	34
Selector de países	34
Funciones de agregación	35
3. Selección de Vista de Visualización	35
Mapa mundial	36
Gráfico de barras horizontales	37
Gráfico de barras animado	38
Flujo completo de la herramienta	40

Explicación de la Estructura y Organización del Dashboard	41
1. Filtros por Tópicos	41
2. Selector de Indicador	41
3. Filtros de Tiempo/Países + Función	42
4. Selector de Vistas	42
Estructura del código de la herramienta	43
Propuestas de ampliación del trabajo	44
Mejora en las técnicas predictivas y su elección	44
Ampliar áreas temáticas más allá de la salud	45
Ampliación en la variedad de visualizaciones	46
<b>Conclusión</b>	<b>47</b>
<b>Fuentes</b>	<b>48</b>

## Introducción

En la era de la información, los datos se han convertido en uno de los recursos más valiosos para la toma de decisiones, más aún en áreas tan sensibles como la salud pública, donde los indicadores en este sector son herramientas fundamentales que permiten identificar problemas al analizar tendencias a lo largo del tiempo y comparar la situación entre diferentes regiones o países. Su importancia radica en que proporcionan información crítica para la toma de decisiones en políticas de salud y la asignación de recursos.

Sin embargo, el acceso a la información no siempre está al alcance de todos ya que ésta puede no ser precisa o comprensible. Si bien el Banco Mundial ya ofrece un dashboard oficial para la consulta de estos indicadores, dicha herramienta presenta limitaciones importantes que restringen su potencial analítico. Entre sus debilidades se destacan:

- **Limitaciones en tipos de visualización:** Aunque el dashboard oficial ofrece algunas vistas con posibilidades de personalización, está severamente limitado en cuanto a los tipos de representaciones disponibles. Esta restricción afecta significativamente las posibilidades de análisis.
- **Estancamiento en innovación visual:** El dashboard no muestra señales de evolución ni incorporación de nuevas formas de visualización, lo que limita la evolución de la herramienta frente a las crecientes necesidades de análisis de datos modernos.
- **Navegación poco intuitiva:** Para utilizar de forma efectiva el dashboard del Banco Mundial, es necesario tener conocimientos previos sobre la organización de sus bases de datos, los códigos de los indicadores. Además, si bien ofrece algunas funciones estadísticas, estas no están presentadas de manera clara ni accesible en la interfaz.
- **Valores faltantes:** Cuando hay años sin datos, el dashboard simplemente los omite, generando visualizaciones incompletas. Esto dificulta el análisis comparativo y puede llevar a conclusiones erróneas o imprecisas.

Este proyecto busca cerrar esa brecha mediante el desarrollo de una aplicación web interactiva que permita consultar y visualizar datos de salud global de manera sencilla. Esta

herramienta está diseñada para ser accesible tanto para profesionales del área, investigadores, responsables de políticas públicas, hasta personas interesadas en explorar y entender los datos de salud global, sin necesidad de conocimientos previos.

A través de una interfaz intuitiva, la aplicación permitirá a los usuarios explorar indicadores de salud, compararlos entre países y períodos, así como visualizar la información mediante gráficos y mapas interactivos. Además, se incorporarán técnicas de predicción para garantizar la consistencia de los datos en caso de que la fuente original no cuente con toda la información disponible. También permitirá aplicar métodos estadísticos, como agregaciones por promedio, tasas de cambio o valores extremos, de manera simple para obtener análisis representativos. Todo esto se presenta en un entorno visual accesible, que evita sobrecargar al usuario con interfaces complicadas. A su vez, la visualización incluye animaciones dinámicas que permiten analizar la evolución temporal de los datos de manera intuitiva, facilitando la detección de tendencias, cambios y patrones a lo largo del tiempo. Esto contribuirá a una mejor experiencia de usuario y asegurará que los datos presentados sean útiles para la toma de decisiones.

En términos técnicos, el desarrollo de esta aplicación implica la integración de una API (Application Programming Interface), que es un conjunto de protocolos y herramientas que permiten la comunicación entre diferentes sistemas de software. En este caso, la API del Banco Mundial será la fuente principal de información, utilizando específicamente la base de datos **World Development Indicators (WDI)**, que proporciona información detallada sobre indicadores de salud global y otros aspectos del desarrollo.

Sin embargo, la información no solamente es recopilada por el Banco Mundial<sup>1</sup>, también obtiene información de otras fuentes internacionales confiables como la OMS (Organización Mundial de la Salud), OCDE (Organización para la Cooperación y el Desarrollo Económico), y demás oficinas nacionales de estadística.

---

<sup>1</sup> El Banco Mundial es una de las instituciones más importantes a nivel global en el ámbito del desarrollo económico y social. Fundado en 1944, su objetivo es “poner fin a la pobreza extrema y promover la prosperidad compartida en un planeta habitable”. Para lograrlo, proporciona financiamiento, asistencia técnica, además de recopilar y difundir una gran cantidad de datos y estadísticas sobre indicadores en áreas de salud, educación e infraestructura.

Además, se utilizarán herramientas de visualización de datos para crear gráficos y mapas que faciliten la interpretación de la información. Estas herramientas serán fundamentales para entender cómo funciona la aplicación y cómo se logra su objetivo principal de transformar datos complejos en información accesible y útil.

## Tecnologías Utilizadas en el Desarrollo

Para el desarrollo del Dashboard, se eligió “**React**” como framework de JavaScript. Una de las principales razones fue la rapidez de desarrollo que ofrece, gracias a su enfoque basado en componentes reutilizables que permite estructurar la interfaz del dashboard de manera modular. Además, cuenta con un gran ecosistema respaldado por una comunidad activa, lo que garantiza acceso a una amplia variedad de librerías y herramientas que aceleran el proceso de desarrollo. Otra razón clave fue que existe conocimiento previo de la tecnología, es decir, cierta familiaridad con sus conceptos fundamentales, como hooks, estado y props. Todo esto permitió reducir drásticamente la curva de aprendizaje y centrar los esfuerzos en la implementación de las funcionalidades requeridas.

Otra tecnología importante de mencionar - que fue utilizada como complemento a “**React**”- fue “**TypeScript**”. Proporciona un sistema de tipado estático que resultó valioso durante el desarrollo del dashboard. Al definir interfaces y tipos para los datos que fluyen a través de los componentes, se pudo detectar errores potenciales durante la fase de desarrollo en lugar de encontrarlos en tiempo de ejecución. Esto fue particularmente útil al manejar estructuras de datos provenientes de la API que alimenta las vistas. Además la legibilidad y mantenibilidad del código mejoraron debido a que el tipado actúa como documentación, permitiendo a cualquier desarrollador entender rápidamente qué datos espera y devuelve cada componente o función.

## Visualización de Mapas

Para la representación geográfica de los indicadores de salud, se implementó la librería “**react-simple-maps**”, que simplifica la creación de mapas interactivos. La herramienta permite personalizar distintos aspectos del mapa, como resaltar los países donde esté el puntero del mouse, modificar el color de relleno de los países según el valor del indicador seleccionado, ofreciendo una visión clara para la comparación entre países.

En cuanto a la escala de colores, se definió utilizando la librería “**chroma-js**”, que es ideal para la manipulación cromática en la visualización de datos en el mapa. La paleta de colores utilizada fue de tonalidades anaranjadas a rojas, por ser intuitiva para representar magnitudes (valores más altos es igual a un rojo intenso).

## Gráficos con D3.js

Para la generación de gráficos (barras horizontales y gráfico de barras animado en este caso), se optó por la librería “**D3.js**”

Para una mejor visualización de los datos, se implementó una funcionalidad para adaptar el ancho de las barras según la cantidad de países a mostrar. Permitiendo expandir verticalmente los gráficos manteniendo proporciones legibles de los nombres de los países y la separación entre barras.

En cuanto a las ventajas del uso de la librería para la generación de las vistas:

- Permite personalizar cada aspecto del gráfico (ejes, transiciones, tooltips, textos) sin limitaciones y con un control total.
- Se integra perfectamente con React para actualizaciones dinámicas.

## Estilos del dashboard con Tailwind

La elección de Tailwind como herramienta principal para estilizar el dashboard fue resultado de una evaluación de las necesidades del proyecto y las características que ofrecen en comparación a otras librerías como Bootstrap. Las ventajas que llevaron a decantar por ella son:

- **Enfoque utility-first:** Tailwind permite aplicar estilos directamente en el markup mediante clases utilitarias. Esto elimina la necesidad de cambiar constantemente entre archivos de componentes y hojas de estilo separadas, acelerando el desarrollo del dashboard y manteniendo todo el código relacionado con un componente en un solo lugar.
- **Mayor control sobre el diseño:** A diferencia de “**Bootstrap**”, donde los componentes vienen con estilos predefinidos y rígidos, Tailwind permite una personalización más profunda sin estar atado a patrones visuales preestablecidos.
- **Conocimiento de la herramienta y velocidad de desarrollo:** Otra razón fundamental fue la familiaridad previa con la herramienta, lo que contribuyó significativamente a acelerar el proceso de desarrollo. Conocer en profundidad su sistema de clases utilitarias permitió implementar estilos de manera rápida, sin necesidad de dedicar tiempo al aprendizaje de un framework nuevo.

## Integración y Uso de la API del Banco Mundial en la Aplicación

La aplicación utiliza la API del Banco Mundial como fuente principal de datos para los indicadores de salud global, en específico la base de datos **World Development Indicators (WDI)**, que proporciona información detallada sobre indicadores de salud global y otros aspectos del desarrollo.

En las consultas se le indica a la API, que la fuente a utilizar debe ser la **WDI**, agregando el siguiente parámetro de consulta al endpoint “**source=2**” (donde 2 es el código de esta fuente de datos).

Para la demostración de la integración realizada se utilizará un indicador de ejemplo para facilitar la explicación.

- Indicador: **Población, hombres (SP.POP.TOTL.MA.IN)**
- Países seleccionados: **TODOS**
- Año seleccionado: **2010**

A continuación, se detalla cómo se realiza la integración y las funcionalidades clave implementadas.

### Consulta de todos los indicadores

Para optimizar el rendimiento, la aplicación precarga todos los indicadores de salud disponibles mediante una única consulta al endpoint correspondiente. Esta información se almacena localmente, permitiendo búsquedas instantáneas sin necesidad de consultas adicionales a la API.

Endpoint para precarga de todos los indicadores de salud:  
<https://api.worldbank.org/v2/es/topic/8/indicator?format=json&source=2>

### Ventajas de la precarga inicial

- Permite evitar rate limits (límites de solicitudes a la API).
- Poder filtrar indicadores por tópicos instantáneamente.
- Poder navegar por todos los indicadores disponibles sin necesidad de realizar una búsqueda más específica.

## Obtención de Regiones y Países

Se implementan consultas de regiones y países a la API, para ser utilizados en el selector de países, permitiendo agrupar países por regiones geográficas y filtrar entidades no relevantes para las vistas como agregados regionales.

Endpoint para regiones:

<https://api.worldbank.org/v2/es/regions?format=json>

Endpoint para países:

<https://api.worldbank.org/v2/es/countries?format=json>

## Consulta de indicador seleccionado

Se realiza una petición que incluye los países seleccionados (o todos si no hay filtro), el indicador elegido y el rango de años especificado por el usuario.

Endpoint de consulta del indicador:

[https://api.worldbank.org/v2/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per\\_page=10000&date=2000:2020&source=2](https://api.worldbank.org/v2/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per_page=10000&date=2000:2020&source=2)

### 1. Manejo de países:

La parte de la consulta “/country/ALL”, esto quiere decir que se están consultando los registros de todos los países. En caso de seleccionar solamente países específicos se debe modificar esta parte de la consulta, reemplazando “ALL” por los códigos ISO oficiales de los países separados por punto y coma.

## 2. Ampliación del rango temporal:

Para soportar el sistema de predicción de datos faltantes, la aplicación consulta automáticamente un rango de años más amplio que el seleccionado por el usuario (es de  $\pm 10$  años). Estos datos adicionales:

- Se utilizan internamente para mejorar la precisión de las interpolaciones
- No se muestran en las visualizaciones principales
- Permiten detectar tendencias que ayudan a predecir valores faltantes

## Metadata del indicador

Para proporcionar información detallada sobre el indicador seleccionado, la aplicación implementa un sistema de consulta de metadatos.

Contenido de los metadatos obtenidos:

- Nombre completo y código del indicador
- Definición técnica
- Fuente original de los datos
- Notas sobre estimaciones o interpolaciones realizadas.
- Tópicos a los que pertenece el indicador

Se implementa una lógica de consulta de metadatos en dos idiomas (español e inglés) con el siguiente flujo:

### 1. Primera consulta en español

Se realiza una petición inicial al endpoint en español para obtener la descripción, definición y notas metodológicas del indicador. Esto garantiza que los usuarios hispanohablantes reciban la información en su idioma nativo cuando esté disponible.

Endpoint para metadatos en español:

**<https://api.worldbank.org/v2/es/indicator/SP.POP.TOTL.MA.IN?format=json&source=2>**

## 2. Consulta en inglés

Si la respuesta en español no contiene suficiente información o devuelve campos vacíos, se ejecuta automáticamente una segunda consulta al mismo endpoint pero en inglés, que suele tener metadatos más completos.

Endpoint para metadatos en inglés:

**<https://api.worldbank.org/v2/en/indicator/SP.POP.TOTL.MA.IN?format=json&source=2>**

Esto permite una consistencia en los datos obtenidos asegurando que los usuarios siempre reciban descripciones detalladas.

## Obtención del Rango Temporal Disponible

Para determinar los años de cobertura del indicador seleccionado, se realizan dos consultas estratégicas:

Último año disponible:

**[https://api.worldbank.org/v2/es/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per\\_page=1&page=1&source=2](https://api.worldbank.org/v2/es/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per_page=1&page=1&source=2)**

Para obtener el registro más reciente se agrega a la url “per\_page=1” y “page=1”, que permite obtener un solo resultado que es el más reciente, permitiendo conocer el año máximo disponible. También la respuesta trae información relacionada a la cantidad de registros disponibles. Este dato se usará en la siguiente consulta para obtener el dato más antiguo y así determinar el mínimo año de datos disponibles.

Primer año disponible:

[https://api.worldbank.org/v2/es/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per\\_page=1&page=ULTIMA\\_PAGINA&source=2](https://api.worldbank.org/v2/es/country/ALL/indicator/SP.POP.TOTL.MA.IN?format=json&per_page=1&page=ULTIMA_PAGINA&source=2)

Con esta última solicitud se obtiene el registro más antiguo existente para determinar el año mínimo disponible.

Los beneficios de realizar estas consultas es establecer límites temporales reales del indicador, y a su vez evitar búsquedas en años sin datos disponibles.

## Procesamiento de datos

### Normalización

Para determinar el color del valor del país en la vista del mapa se debe transformar datos numéricos de diferentes escalas y magnitudes a un rango común (típicamente [0, 1]), lo que permite una representación visual coherente, especialmente en mapas coropléticos. A continuación, se detalla la lógica y los métodos utilizados:

Cuando se trabaja con datos que abarcan múltiples órdenes de magnitud (ejemplo: tasas de mortalidad infantil [0.1, 50] vs. población [1,000, 1,000,000,000]), surgen dos problemas:

1. Dominancia de valores extremos: Un valor muy alto (ejemplo: 1,000,000) puede opacar las diferencias entre valores pequeños (ejemplo: 10 y 30).
2. Distorsión en escalas de color: Sin normalización, los colores no representarán proporciones reales.

La normalización resuelve esto ajustando los datos para que:

- Todos los valores estén dentro de un rango estándar (ejemplo: 0 a 1).
- Las diferencias relativas se preserven (ejemplo: 100 es el doble de 50, y eso se refleja en el color).

### Métodos de Normalización utilizados

1. Normalización lineal: Transforma los datos al rango [0, 1], es útil para valores en rango similar. Su fórmula es la siguiente:

$$\text{Valor normalizado} = \frac{\text{valor} - \text{min}}{\text{max} - \text{min}}$$

2. Normalización logarítmica: Para datos que abarcan múltiples órdenes de magnitud (ejemplo: [1, 10, 100, 1000]). Reduce el impacto de valores extremos y preserva diferencias relativas. Su fórmula es la siguiente:

$$\text{Valor normalizado} = \frac{\log_{10}(\text{valor}) - \log_{10}(\text{min})}{\log_{10}(\text{max}) - \log_{10}(\text{min})}$$

3. Normalización con raíz cuadrada: Para datos con muchos valores cercanos a 0 (ejemplo: [0.1, 0.5, 0.9]). Útil cuando los valores son pequeños pero no logarítmicos (ejemplo: tasas de crecimiento). Su fórmula es la siguiente:

$$\text{Valor normalizado} = \frac{\sqrt{\text{valor}} - \sqrt{\text{min}}}{\sqrt{\text{max}} - \sqrt{\text{min}}}$$

Criterios para la obtención de valores representativos de Indicadores en intervalos de tiempo

Al visualizar datos en gráficos o mapas como es el caso, es común trabajar con series temporales que abarcan rangos de años. Sin embargo, no es posible mostrar todos los valores individuales de los indicadores de cada año para comparar varios países. Debido a esto, se implementaron diferentes criterios de agregación para extraer valores representativos de un rango de años, cada uno con un propósito específico según lo que se quiera analizar:

- **Máximo y Mínimo**

Estos valores permiten identificar picos o valles, que son útiles para detectar situaciones extremas. Por ejemplo, analizar cambios drásticos en la esperanza de vida de una población en un período de tiempo específico o crisis sanitarias.

- **Promedio**

Proporciona una medida de tendencia que suaviza fluctuaciones y da una visión general del comportamiento de un indicador específico en un periodo seleccionado.

- **Tasa de cambio**

Se calcula como el porcentaje de cambio entre el valor inicial (el primer año del período seleccionado) y el final (último año del período seleccionado). Permite evaluar si un indicador mejora, empeora o se mantiene estable en el intervalo. Por ejemplo, una tasa de cambio positiva en la esperanza de vida sugiere mejoras en políticas sanitarias.

- **Más reciente y Más antiguo**

Estos valores permiten contrastar la relación inicial y final del período, facilitando un análisis de la evolución del indicador. El dato más antiguo sirve como punto de referencia para medir el progreso, mientras que el más reciente indica si tuvo un comportamiento mejor, peor o estable en el tiempo.

La elección del criterio depende del objetivo del análisis, si se busca identificar tendencias, los promedios o la tasa de cambio son útiles, mientras que si se buscan extremos o cambios abruptos, los valores máximos y mínimos son más relevantes. Estos métodos se pueden aplicar dinámicamente sobre las vistas permitiendo una mayor flexibilidad a la hora de analizar los gráficos.

## Técnicas de predicción

En el análisis de datos es común encontrar valores faltantes de ciertos años o países debido a limitaciones o problemas en la recolección de información. Para garantizar que la

aplicación muestre datos consistentes y útiles, se implementó un sistema de predicción basado en técnicas de regresión. Para determinar el método más adecuado se tuvieron en cuenta tres variables claves que describen el comportamiento histórico del indicador:

1. Correlación de Pearson: Este coeficiente mide la fuerza y la dirección de la relación lineal entre la variable temporal (por ejemplo, los años) y los valores históricos del indicador que se está analizando. Su valor oscila entre -1 y 1, y su interpretación es la siguiente:

- Un valor cercano a 1: Sugiere una fuerte relación lineal positiva. Esto significa que a medida que el tiempo avanza, los valores del indicador tienden a aumentar de manera consistente a lo largo de una línea recta.
- Un valor cercano a 0: Sugiere que no existe una relación lineal significativa entre el tiempo y los valores del indicador. Los cambios en el tiempo no se asocian con cambios predecibles y lineales en el indicador.
- Un valor cercano a -1: Sugiere una fuerte relación lineal negativa. Esto significa que a medida que el tiempo avanza, los valores del indicador tienden a disminuir de manera consistente a lo largo de una línea recta.
- Valores intermedios: Cuanto más cerca esté el valor a 1 o -1, más fuerte será la relación lineal. Cuanto más cerca esté a 0, más débil será la relación lineal.

2. Promedio de crecimiento: Este valor proporciona una indicación de la tendencia general de los valores del indicador a lo largo del tiempo. Esto dice si, en promedio, los valores están aumentando, disminuyendo o permaneciendo relativamente estables de un período a otro.

- Un promedio de crecimiento mayor que 1: Indica una tendencia general al crecimiento. Los valores del indicador tienden a aumentar con el tiempo. Cuanto mayor sea el valor, más rápido será el crecimiento promedio.
- Un promedio de crecimiento cercano a 1: Sugiere una tendencia relativamente estable. Los valores del indicador no muestran un aumento o disminución significativa en promedio.

- Un promedio de crecimiento menor que 1: Indica una tendencia general al decrecimiento. Los valores del indicador tienden a disminuir con el tiempo. Cuanto menor sea el valor, más rápida será la disminución promedio.
- Un promedio de crecimiento igual a 1: Significa que, en promedio, no hay cambio en los valores del indicador a lo largo del tiempo.

3. Estabilización: Esta variable evalúa si la tasa de cambio del indicador (ya sea crecimiento o disminución) está mostrando signos de ralentización con el tiempo. Es decir, si la magnitud del cambio entre períodos sucesivos se está haciendo cada vez menor.

Al analizar estas variables en conjunto, el sistema de predicción puede seleccionar la técnica de regresión que mejor se adapte al patrón histórico de los datos y, por lo tanto, genere predicciones más confiables para los valores faltantes.

#### Técnicas de regresión utilizadas

Las regresiones lineal, exponencial y logística son métodos genéricos de fácil implementación e interpretación, adecuados cuando las series temporales son cortas y no se dispone de abundantes datos históricos. Estas técnicas permiten generar predicciones rápidas sin requerir infraestructura compleja ni requieren de largos historiales o complejos procesos de análisis, por lo que funcionan con series de pocos puntos y sin mucha estacionalidad, lo cual encaja con el alcance acotado del proyecto.

Las ventajas del uso de estas técnicas:

- **Consistencia visual:** Evita huecos en las vistas que dificulten la interpretación de los datos sin complejidad excesiva.
- **Mejor comparabilidad:** Permite analizar países con datos incompletos en el mismo rango temporal.
- **Toma de decisiones informada:** Proporciona estimaciones razonables cuando no hay datos oficiales.

## 1. Regresión lineal

Modelo estadístico que establece una relación lineal entre una variable dependiente (indicador) y una o más variables independientes (tiempo, en este caso).

La fórmula es:

$$y = m x + b$$

donde **y** es el valor predicho, **x** es el año, **m** es la pendiente (tasa de crecimiento) y **b** es la ordenada al origen.

Para analizar indicadores que cambian de forma gradual y constante a lo largo del tiempo, como la esperanza de vida o la tasa de natalidad, se utiliza esta técnica. Es ideal cuando no hay patrones que se repitan en ciertas épocas (como picos anuales de enfermedades) ni cambios bruscos (por ejemplo, por una crisis o políticas públicas). En esos casos, el método permite predecir fácilmente la tendencia general del indicador sin problemas. Según Ratner (2017), este enfoque funciona bien cuando la relación entre el tiempo y el indicador se mantiene estable. Sin embargo, si hay ciclos repetitivos o saltos abruptos, se necesitan métodos más avanzados para capturar esos detalles. La regresión lineal, aunque simple, podría no ser suficiente en escenarios complejos o dinámicos.

## 2. Regresión logística

En su forma más común es una técnica que se usa para predecir la probabilidad de que un evento ocurra. La fórmula es:

$$y = \frac{L}{1+e^{-x}}$$

donde **L** es el límite superior de la probabilidad, y **x** es el logaritmo de odds, que traduce las probabilidades a una escala lineal.

Además de su uso tradicional en modelos binarios (probabilidades entre 0 y 1), la regresión logística puede adaptarse para modelar procesos de crecimiento continuo con saturación, como en el caso del modelo logístico de crecimiento. Este enfoque es especialmente útil para representar fenómenos donde, tras una etapa inicial de expansión acelerada, se observa una desaceleración progresiva que tiende a estabilizarse cerca de un límite superior (por ejemplo, la cobertura de servicios o la adopción tecnológica). En estos casos, el parámetro  $L$  deja de ser 1 — como en la regresión logística binaria — y pasa a representar el valor real de saturación del indicador (como el 100 % de cobertura o la capacidad máxima de un recurso).

### 3. Regresión exponencial

Técnica que permite modelar datos que siguen una tendencia exponencial. La fórmula es:

$$y = a \cdot e^{kx}$$

donde  $a$  es el valor inicial,  $k$  es la tasa de crecimiento exponencial y  $e$  es la base del logaritmo natural.

En contextos donde las series temporales de un indicador de salud no muestran señales claras de saturación — es decir, no hay indicios de que se aproxime a un límite superior — la regresión exponencial resulta más adecuada que el modelo logístico por las siguientes razones:

1. Fase inicial idéntica: En las primeras etapas de crecimiento, sea una epidemia o la adopción de un servicio de salud, el modelo logístico y el exponencial lucen igual: ambos crecen casi exactamente con la misma forma porque, mientras estamos lejos de cualquier límite, no actúa el freno que introduce la logística. Esta observación coincide con lo señalado por Chen et al. (2020) al analizar la evolución temprana de la pandemia por COVID-19 en Estados Unidos, quienes indicaron que, en las fases iniciales, el crecimiento era

esencialmente exponencial, y el modelo logístico se comportaba de manera similar al no manifestarse aún el límite de crecimiento.

2. Menos parámetros, más sencillez: La logística requiere de un parámetro extra (la “capacidad máxima”) denominado  $L$ . Si la serie es corta y no hay datos cercanos a ese supuesto límite, no hay información real para fijar  $L$ , y la logística puede hacer predicciones erráticas. En cambio, la exponencial basa su fortaleza precisamente en su simplicidad y en la suposición de que cada periodo el cambio es siempre un porcentaje fijo de lo que ya hay.

Por estas razones, en ausencia de evidencia empírica de desaceleración o techo en la serie, la regresión exponencial ofrece un balance óptimo entre **precisión** y simplicidad, justificando su uso antes que el modelo logístico.

Patrones de Crecimiento Epidemiológico: De la Exponencial a la Logística (Caso COVID-19)

Como se detalló previamente, la regresión logística y la exponencial ofrecen perspectivas distintas sobre el crecimiento de fenómenos. En particular, se ha observado en numerosos procesos epidemiológicos, un patrón que comienza con una fase de crecimiento rápido, modelado eficazmente por la regresión exponencial en sus etapas iniciales. Sin embargo, a medida que el proceso evoluciona y se acerca a límites impuestos por factores diversos, el crecimiento tiende a desacelerarse y estabilizarse, adoptando la forma de una curva en S característica del modelo logístico. El caso de la propagación del COVID-19 ilustra claramente esta dinámica. Inicialmente, la expansión del virus fue de naturaleza exponencial, pero estudios como el de Debecker y Modis (2021) evidenciaron que en al menos 25 países, la trayectoria de la pandemia se ajustó posteriormente a un patrón logístico, demostrando la capacidad de este modelo para capturar la saturación gradual propia de muchos procesos del mundo real. Esta observación resalta la importancia de considerar el modelo logístico cuando se anticipa una eventual limitación del crecimiento, en contraste con la suposición de expansión ilimitada inherente a la regresión exponencial.

## Lógica de selección de técnica

El sistema implementa un mecanismo para elegir la técnica de predicción más adecuada según las características de los datos. Aquí se explica el proceso paso a paso:

### 1. Filtrado Inicial de Datos (Preprocesamiento)

Antes de analizar los datos, se aplica un filtrado riguroso para garantizar que solo se consideran países reales. Esto elimina:

- Agregados regionales: Como la Unión Europea (EU) o la OCDE (OE), que no representan países individuales.
- Códigos especiales: Identificadores que comienzan con X (excepto Kosovo XK) o Z (excepto Sudáfrica ZA, Zambia ZM, y Zimbabue ZW), típicamente usados para grupos o categorías no geográficas.

Permiten asegurar que los datos predichos correspondan únicamente a naciones, evitando el procesamiento de datos indeseados.

### 2. Selección de la Técnica de Predicción

Una vez filtrados los datos, el sistema decide automáticamente qué técnica de regresión aplicar mediante un proceso en dos etapas:

#### a) Análisis por País Individual

Para cada país con datos históricos suficientes (más de dos valores registrados), se evalúa:

Condición	Técnica seleccionada	Motivo
Correlación $\leq 0.7$ y se estabiliza	LOGÍSTICA	Comportamiento no lineal con tendencia al equilibrio; la logística modela bien ese límite.
Correlación $\leq 0.7$ y no se estabiliza	EXPONENCIAL	Indicador con bajo ajuste y sin estabilización; caída o crecimiento continuo.
Correlación $> 0.7$ y crecimiento muy estable ( $0.9 \leq$ promedio crecimiento $\leq 1.1$ )	LINEAL	El crecimiento promedio es mínimo; la recta representa adecuadamente la tendencia
Correlación $> 0.7$ y promedio crecimiento $< 0.9$ y se estabiliza	LOGÍSTICA	El indicador cae pero tiende a estabilizarse, típico de una curva con forma de S.
Correlación $> 0.7$ y promedio crecimiento $< 0.9$ y <b>no</b> se estabiliza	LINEAL	Descenso constante sin evidencia de estabilización; la caída lineal es más realista
Correlación $> 0.7$ y promedio crecimiento $> 1.1$ y se estabiliza	LOGÍSTICA	Crecimiento con desaceleración progresiva; proyecta un límite.
Correlación $> 0.7$ y $1.1 <$ promedio crecimiento $\leq 1.3$ y datos $\geq 5$ y no se estabiliza	EXPONENCIAL	Crecimiento moderado sostenido con respaldo de datos; se ajusta bien a la curva exponencial.
Correlación $> 0.7$ y promedio crecimiento muy alto o pocos datos	LOGÍSTICA	Evita proyecciones exageradas o sobreajuste por escasez de datos.

#### b) Votación Global

- Se cuenta cuántos países se ajustan mejor a cada técnica (lineal, exponencial, logística).
- La técnica con más votos se aplica uniformemente a todos los países.
- Regla de respaldo: Si no hay consenso claro, se usa regresión lineal por defecto (la más estable).

### 3. Aplicación de la Predicción

Con la técnica seleccionada, el sistema permite:

1. Calcular valores faltantes: Para años donde no hay datos, usando la tendencia histórica extendida del país.
2. Registrar el método usado: Cada valor predicho incluye un campo que documenta mediante qué técnica se obtuvo.

Las ventajas de este enfoque son:

- Consistencia: Todos los países se analizan con el mismo método en un período dado.
- Transparencia: Los usuarios pueden ver qué técnica se usó para cada predicción.
- Adaptabilidad: El sistema elige automáticamente entre modelos simples (lineal) o complejos (logística, exponencial) según los datos.

Es importante aclarar que los valores predichos mediante regresión en este trabajo no se utilizan para realizar nuevas predicciones, sino para completar series temporales en las que hay datos ausentes. Como advierte Pyle (1999):

*(...) los reemplazos de valores faltantes no deberían aportar información al modelo. Lo que hacen es permitir que la información presente [...] sea utilizada por la herramienta de modelado, añadiendo la menor distorsión posible al conjunto de datos [...] dado que el valor de reemplazo se deriva exclusivamente de información ya presente en el conjunto, solo puede volver a expresar dicha información (traducción propia, p. 160).*

Esta perspectiva enfatiza que el objetivo de completar valores ausentes no es enriquecer el modelo predictivo con información nueva, sino preservar la coherencia interna del conjunto de datos. En ese sentido, el uso de regresiones en este proyecto cumple una función operativa más que analítica.

## Descripción y uso del dashboard

La aplicación está diseñada para permitir a los usuarios explorar y visualizar datos de salud global de manera intuitiva. A continuación, se detalla el proceso paso a paso, incluyendo la selección del indicador, la funcionalidad de cada filtro y cómo interactúan entre sí para generar visualizaciones útiles.

Para simplificar el uso de la herramienta, se muestra en el lugar de la vista, un modal con el estado actual de los filtros necesarios para realizar una búsqueda precisa. Es necesario que los tres pasos estén completos para poder obtener los resultados.

### Seleccione los filtros correspondientes para realizar la búsqueda

- Indicador seleccionado
- Países seleccionados
- Intervalo de tiempo seleccionado

Aplicar búsqueda

Figura 1 - Guía visual de filtros a seleccionar para obtener resultados.

Fuente: Elaboración propia, basada en la práctica.

Como se observa en el modal, indica qué filtros ya se encuentran seleccionados y los restantes a completar. No es necesario seleccionarlos en el orden establecido pero hay que tener en cuenta que el filtro de intervalo de tiempo no estará disponible hasta que no se seleccione un indicador, debido a que es necesario saber sobre cuál se va a trabajar para conocer los años en los que hay datos disponibles. Una vez completado los pasos se debe clicar el botón “Aplicar búsqueda” para realizar la consulta de datos.



Figura 2 - Interfaz de la herramienta que muestra el proceso de selección de indicadores, aplicación de filtros y elección de vista.

Fuente: Elaboración propia, basada en la práctica.

## 1. Selección del Indicador de Salud

Este permite al usuario elegir el indicador de salud que desea analizar.

Componentes

Selector de tópicos (opcional)

Tiene como función filtrar los indicadores por categorías. Facilita al usuario la búsqueda de estos a pesar de desconocer el nombre exacto y navegar por indicadores pertenecientes al tópico sin necesidad de completar el buscador.

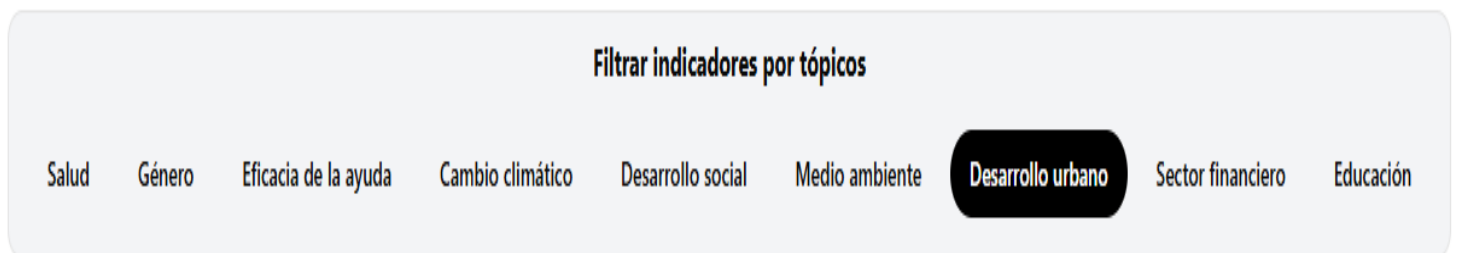


Figura 3 - Selector de tópicos, permite a los usuarios categorizar y filtrar los indicadores, facilitando la búsqueda de datos específicos.

Fuente: Elaboración propia, basada en la práctica.

Buscador

Es el campo de texto que filtra dinámicamente los indicadores disponibles a medida que el usuario escribe. Si anteriormente selecciona un tópico, las sugerencias que aparezcan, pertenecerán solamente al tópico elegido, facilitando la búsqueda.

Otra de las funcionalidades de este es la de permitir la búsqueda exacta de un indicador por su código, facilitando y agilizando la selección del mismo.

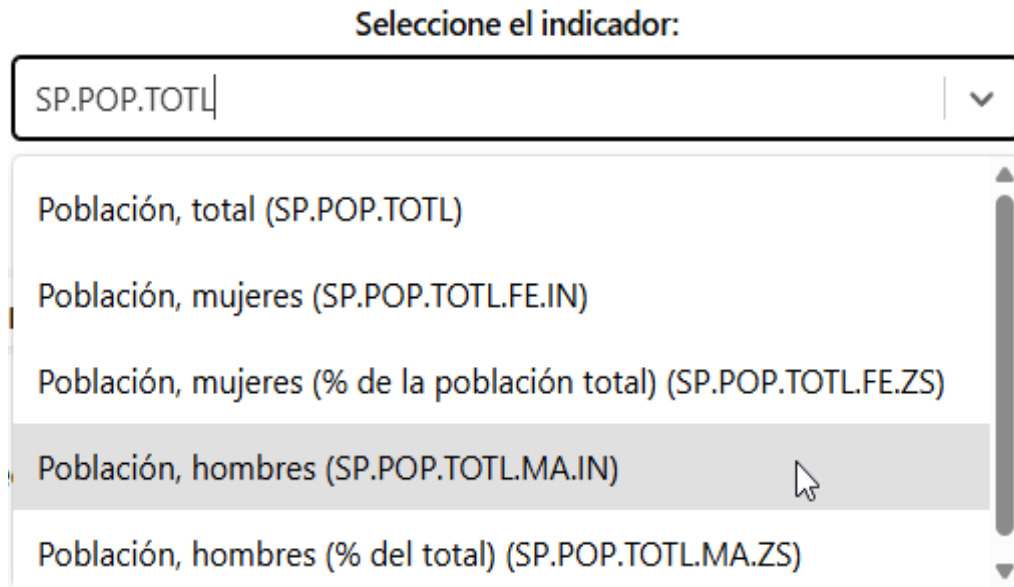


Figura 4 - Buscador que permite ir filtrando los indicadores a medida que se escribe.

Fuente: Elaboración propia, basada en la práctica.

Metadata del indicador seleccionado

Una vez seleccionado el indicador requerido, se puede observar la selección de manera visual con el nombre y su código asociado.

**Indicador seleccionado**  
**Población, total (SP.POP.TOTL) ⓘ**

Figura 5 - Detalle del nombre y código del indicador seleccionado.

Fuente: Elaboración propia, basada en la práctica.

Se puede obtener mayor información de este al clickear en el icono que aparece a su derecha, abrirá un modal con la metadata, es decir, información clave sobre el indicador seleccionado, brindando contexto y transparencia sobre los datos mostrados en las visualizaciones. Debido a limitaciones de la API, al consultar información preferentemente en el idioma español, algunas veces no trae esta información adicional, la cual - en su mayoría - se encuentra en el idioma inglés. Es decir, el sistema preferentemente muestra los datos en español y, en caso de no encontrarlo, se realiza la consulta de datos en el idioma inglés, como se ve en la imagen siguiente.

## Metadata del indicador ×

**Id del indicador:** SP.POP.TOTL.MA.IN

**Nombre del indicador:** Población, hombres

**Fuente:** Indicadores del desarrollo mundial

**Definición del indicador:** Male population is based on the de facto definition of population, which counts all male residents regardless of legal status or citizenship.

**Nota del indicador:** World Bank staff estimates using the World Bank's total population and age/sex distributions of the United Nations Population Division's World Population Prospects: 2022 Revision.

**Fuente:** Indicadores del desarrollo mundial

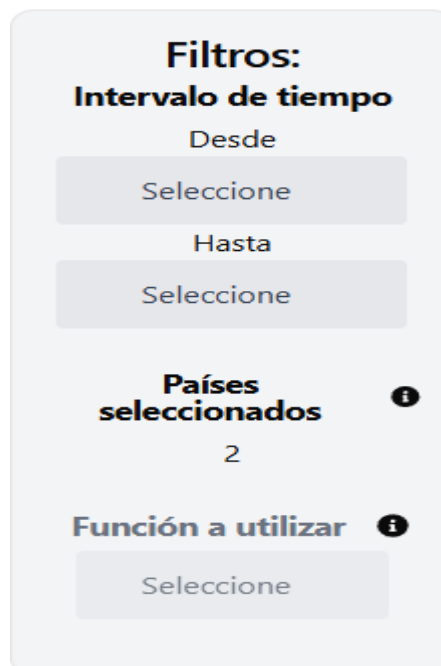
**Tópicos:** Salud

Figura 6 - Proporciona información detallada sobre el indicador de salud seleccionado, incluyendo su definición y fuente de consulta.

Fuente: Elaboración propia, basada en la práctica.

## 2. Filtros de búsqueda adicionales

Tienen como objetivo refinar el análisis seleccionando rangos temporales, países y funciones de agregación.



**Filtros:**

**Intervalo de tiempo**

Desde  
Seleccione

Hasta  
Seleccione

**Países seleccionados** ⓘ  
2

**Función a utilizar** ⓘ  
Seleccione

Figura 7 - Sección de filtros permite a los usuarios refinar el análisis de datos seleccionando intervalos de tiempo, países y funciones de agregación específicas.

Fuente: Elaboración propia, basada en la práctica.

### Componentes

Selector de años ("Desde - Hasta")

Permite definir el intervalo de tiempo de los datos a visualizar. Al no seleccionar un rango de años, por defecto se consulta el dato del último año disponible del indicador.

Selector de países

Esta función permite seleccionar determinados países, permitiendo filtrar solo los datos útiles para el análisis. Es posible seleccionar un país, regiones (clickeando el nombre de la región) o todos los países.

- Filtrar datos por nación: Visualizar información específica de uno o varios países.
- Comparar indicadores: Analizar diferencias y similitudes entre regiones o grupos de países.
- Personalizar visualizaciones: Adaptar gráficos y mapas a las necesidades del usuario, ya sea para un estudio detallado o una comparación general.

**Seleccionar países**

<p><b>Asia oriental y el Pacífico</b></p> <input type="checkbox"/> Samoa Americana <input type="checkbox"/> Australia <input type="checkbox"/> Brunei Darussalam <input type="checkbox"/> China <input type="checkbox"/> Fiji <input type="checkbox"/> Micronesia (Estados Federados de) <input type="checkbox"/> Guam <input type="checkbox"/> Hong Kong, Región Administrativa Especial <input type="checkbox"/> Indonesia <input type="checkbox"/> Japón <input type="checkbox"/> Camboya <input type="checkbox"/> Kiribati <input type="checkbox"/> Corea, República de <input type="checkbox"/> República Democrática	<p><b>Europa y Asia central</b></p> <input type="checkbox"/> Albania <input type="checkbox"/> Andorra <input type="checkbox"/> Armenia <input type="checkbox"/> Austria <input type="checkbox"/> Azerbaiyán <input type="checkbox"/> Bélgica <input type="checkbox"/> Bulgaria <input type="checkbox"/> Bosnia y Herzegovina <input type="checkbox"/> Belarús <input type="checkbox"/> Suiza <input type="checkbox"/> Islas del Canal <input type="checkbox"/> Chipre <input type="checkbox"/> República Checa <input type="checkbox"/> Alemania <input type="checkbox"/> Dinamarca <input type="checkbox"/> España	<p><b>América Latina y el Caribe</b></p> <input checked="" type="checkbox"/> Aruba <input checked="" type="checkbox"/> Argentina <input checked="" type="checkbox"/> Antigua y Barbuda <input checked="" type="checkbox"/> Bahamas <input checked="" type="checkbox"/> Belice <input checked="" type="checkbox"/> Bolivia <input checked="" type="checkbox"/> Brasil <input checked="" type="checkbox"/> Barbados <input checked="" type="checkbox"/> Chile <input checked="" type="checkbox"/> Colombia <input checked="" type="checkbox"/> Costa Rica <input checked="" type="checkbox"/> Cuba <input checked="" type="checkbox"/> Curacao <input checked="" type="checkbox"/> Islas Caimán <input checked="" type="checkbox"/> Dominica <input checked="" type="checkbox"/> República Dominicana	<p><b>Oriente Medio y Norte de África</b></p> <input type="checkbox"/> Emiratos Árabes Unidos <input type="checkbox"/> Bahrein <input type="checkbox"/> Djibouti <input type="checkbox"/> Argelia <input type="checkbox"/> Egipto, República Árabe de <input type="checkbox"/> Irán, República Islámica del <input type="checkbox"/> Iraq <input type="checkbox"/> Israel <input type="checkbox"/> Jordania <input type="checkbox"/> Kuwait <input type="checkbox"/> Líbano <input type="checkbox"/> Libia <input type="checkbox"/> Marruecos <input type="checkbox"/> Malta <input type="checkbox"/> Omán <input type="checkbox"/> Ribera Occidental y Gaza
<p><b>América del Norte</b></p> <input type="checkbox"/> Bermudas <input type="checkbox"/> Canadá <input type="checkbox"/> Estados Unidos	<p><b>Asia meridional</b></p> <input type="checkbox"/> Afganistán <input type="checkbox"/> Bangladesh <input type="checkbox"/> Bhután <input type="checkbox"/> India <input type="checkbox"/> Sri Lanka <input type="checkbox"/> Maldivas <input type="checkbox"/> Nepal <input type="checkbox"/> Pakistán	<p><b>África al sur del Sahara</b></p> <input type="checkbox"/> Angola <input type="checkbox"/> Burundi <input type="checkbox"/> Benin <input type="checkbox"/> Burkina Faso <input type="checkbox"/> Botswana <input type="checkbox"/> República Centroafricana <input type="checkbox"/> Côte d'Ivoire <input type="checkbox"/> Camerún	

Figura 8 -Modal que permite seleccionar uno o varios países para enfocar el análisis en áreas geográficas específicas.

Fuente: Elaboración propia, basada en la práctica.

### Funciones de agregación

Cuando un indicador tiene múltiples valores en un intervalo de tiempo, esta función extrae los que son representativos de cada país seleccionado para poder visualizar los datos en las vistas de manera correcta.

- **Máximo/Mínimo:** Para identificar picos o valles.
- **Promedio:** Para tendencias generales.
- **Tasa de cambio:** Para medir crecimiento/reducción porcentual en el período.
- **Más reciente/más antiguo:** Para comparar situaciones iniciales vs. actuales.

Debe elegir una función a utilizar en el indicador debido al tener valores de diferentes años del país ×

**Máximo**    Mínimo    Promedio    Tasa de cambio    Más reciente    Más antiguo

**Máximo**

**Descripción:**  
Encuentra el valor más alto registrado dentro del rango de años.

**Sentido de uso:**  
Resalta el punto de mayor desempeño o el pico del indicador.

**Fórmula:**  
Máximo: Compara todos los valores y selecciona el más alto.

Aceptar

Figura 9 - Modal que permite elegir cómo agregar los datos durante un período de tiempo.

Fuente: Elaboración propia, basada en la práctica.

### 3. Selección de Vista de Visualización

El dashboard ofrece tres tipos de visualizaciones interactivas diseñadas para explorar datos de salud global desde diferentes perspectivas. Cada vista está diseñada para responder preguntas específicas, facilitando el análisis tanto para expertos como para usuarios ocasionales.

## Mapa mundial

El mapa proporciona una visión geográfica de los datos, coloreando cada país según el valor del indicador seleccionado. Con escalas cromáticas intuitivas (ej.: rojo para valores altos, an para bajos), esta vista ayuda a detectar patrones regionales, como concentraciones de enfermedades. Al combinar información visual con contexto geográfico, responde a preguntas como: ¿Dónde es más crítica la problemática? o ¿Existe correlación entre las características de una región y la problemática?. Además, al interactuar con el mapa, los usuarios pueden acceder a datos específicos de cada país, como por ejemplo, el año del dato; el continente; si el dato es real o predicho.

El mapa es una poderosa herramienta de visualización que proporciona una perspectiva geográfica clara y contextualizada de los datos, coloreando cada país en función del valor del indicador seleccionado. A través de escalas cromáticas intuitivas, como el uso de tonalidades rojas para valores altos y anaranjadas para valores bajos, esta representación facilita la identificación rápida de patrones regionales y tendencias globales. Por ejemplo, permite detectar concentraciones geográficas de problemas como enfermedades o pobreza, lo que resulta valioso para comprender la distribución espacial de una problemática específica. Al combinar información visual con contexto geográfico, este tipo de mapa responde a preguntas fundamentales: ¿Dónde es más crítica la situación? o ¿Existe correlación entre las características de una región y la problemática? Además, al interactuar con el mapa, los usuarios pueden acceder a datos específicos de cada país, como el año del dato, el continente al que pertenece, si el valor es real o predicho.

En resumen, el mapa mundial no sólo comunica información de manera visual, sino que también potencia el análisis al conectar datos con su contexto geográfico, facilitando decisiones informadas y estrategias basadas en evidencia.

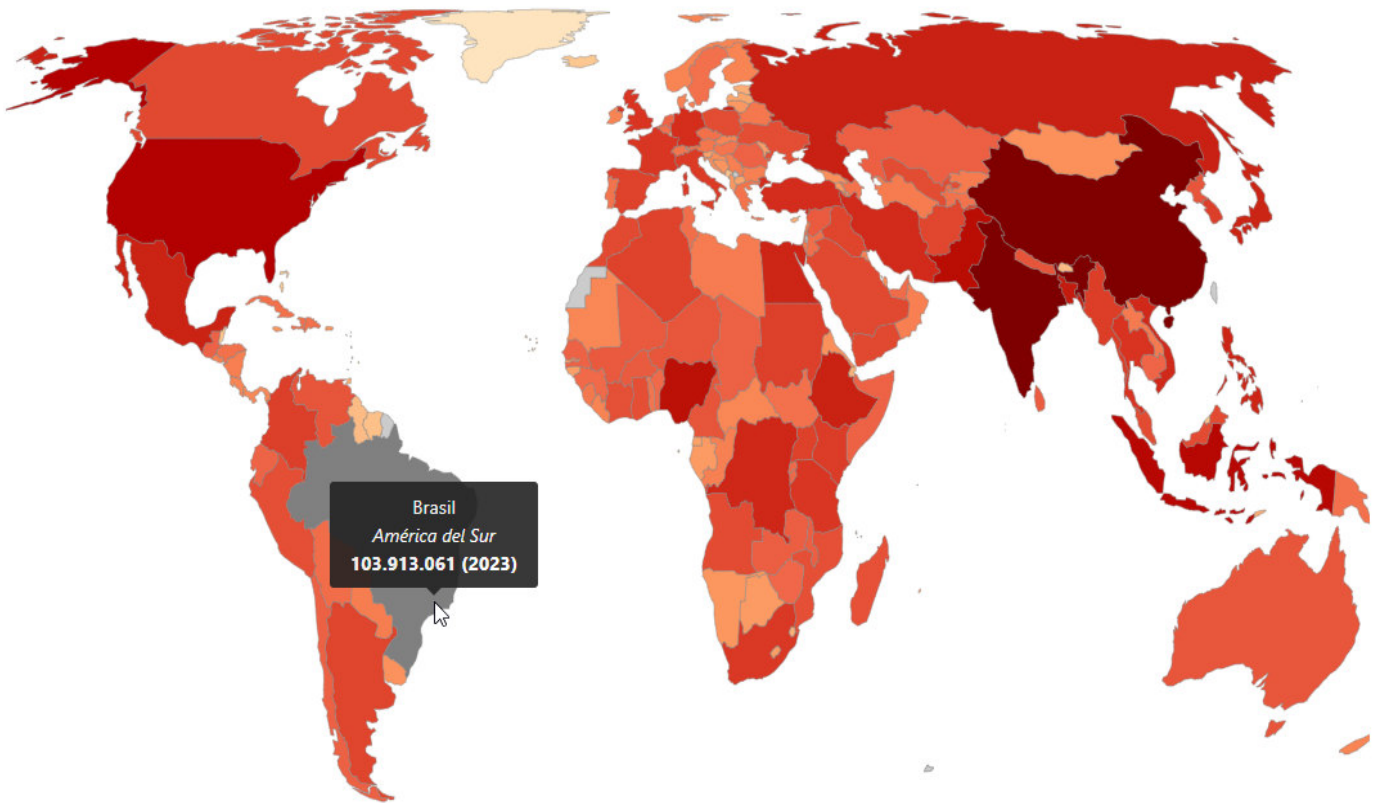


Figura 10 - El mapa coroplético visualiza los datos del indicador de salud geográficamente, utilizando gradientes de color para representar los valores entre los países.

Fuente: Elaboración propia, basada en la práctica.

#### Gráfico de barras horizontales

El gráfico de barras horizontales es una herramienta visual especialmente efectiva para comparar, de manera clara y ordenada, los valores de un indicador entre distintos países en un período determinado. A diferencia de las barras verticales, la disposición horizontal facilita la lectura de los nombres de los países o etiquetas largas, evitando superposiciones o recortes de texto que podrían dificultar la interpretación de los datos. Este diseño también permite organizar los elementos de forma jerárquica, ordenándolos de mayor a menor valor, lo que resalta automáticamente qué territorios ocupan las primeras posiciones en el ranking y cuáles se encuentran en los últimos lugares. Su principal ventaja radica en su capacidad para mostrar rankings de manera intuitiva, destacando diferencias significativas entre los valores más altos

y más bajos de forma inmediata. Además, este tipo de gráfico es útil cuando el objetivo es establecer comparaciones entre un gran número de países, ya que la orientación horizontal maximiza el espacio disponible y mejora la legibilidad incluso con conjuntos de datos extensos.

También cuenta con un tooltip al interactuar con las barras, los datos mostrados son:

- Nombre del país
- Año del dato
- Técnicas de predicción utilizada (si aplica)

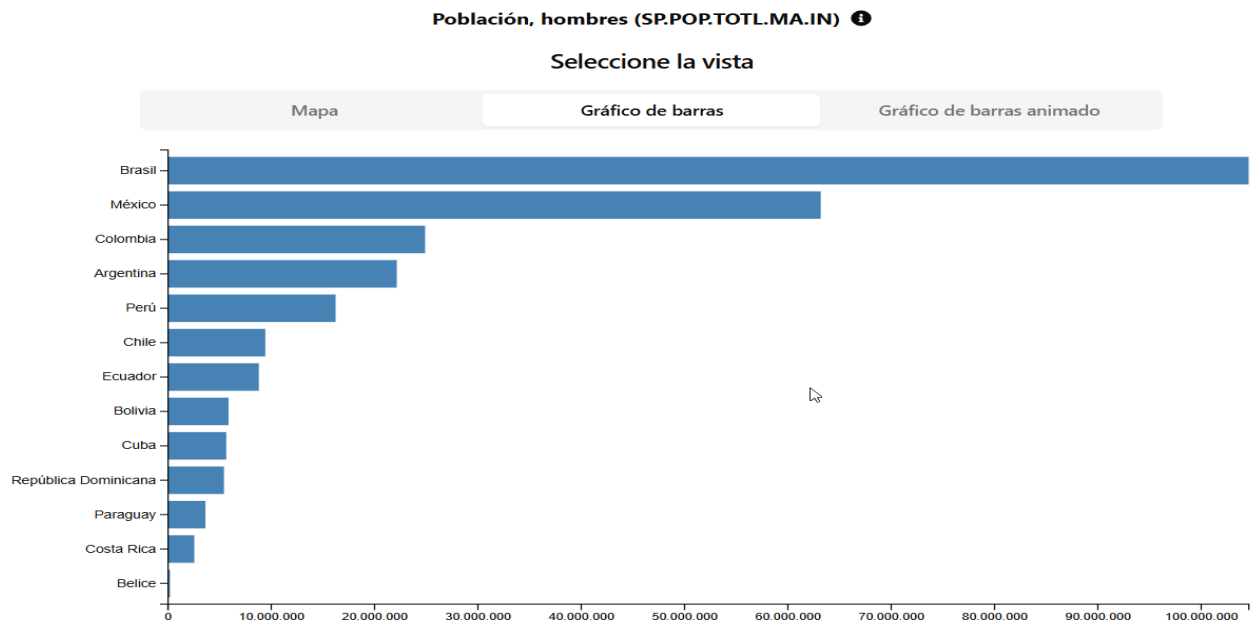


Figura 11 - Gráfico de barras horizontales compara los valores de un indicador entre diferentes países, ordenados de mayor a menor.

Fuente: Elaboración propia, basada en la práctica.

### Gráfico de barras animado

Este tipo de visualización dinámica se destaca por su capacidad para mostrar cómo cambia la posición o el valor absoluto de un indicador a lo largo del tiempo, ofreciendo una representación clara y atractiva de los datos. Las barras "compiten" animadamente entre sí, desplazándose vertical u horizontalmente para reflejar avances, retrocesos o incluso cambios

drásticos en el ranking. Esta animación facilita la comprensión de tendencias y patrones que podrían ser difíciles de interpretar en un formato estático. Por ejemplo, este tipo de gráfico es especialmente útil para identificar tendencias a largo plazo, como el progreso en la cobertura de vacunación en diferentes países, el aumento de la esperanza de vida o la evolución de indicadores económicos y sociales. Al transformar datos complejos en un proceso visual intuitivo, permite responder preguntas clave de manera efectiva: ¿Qué país redujo más la pobreza en la última década?

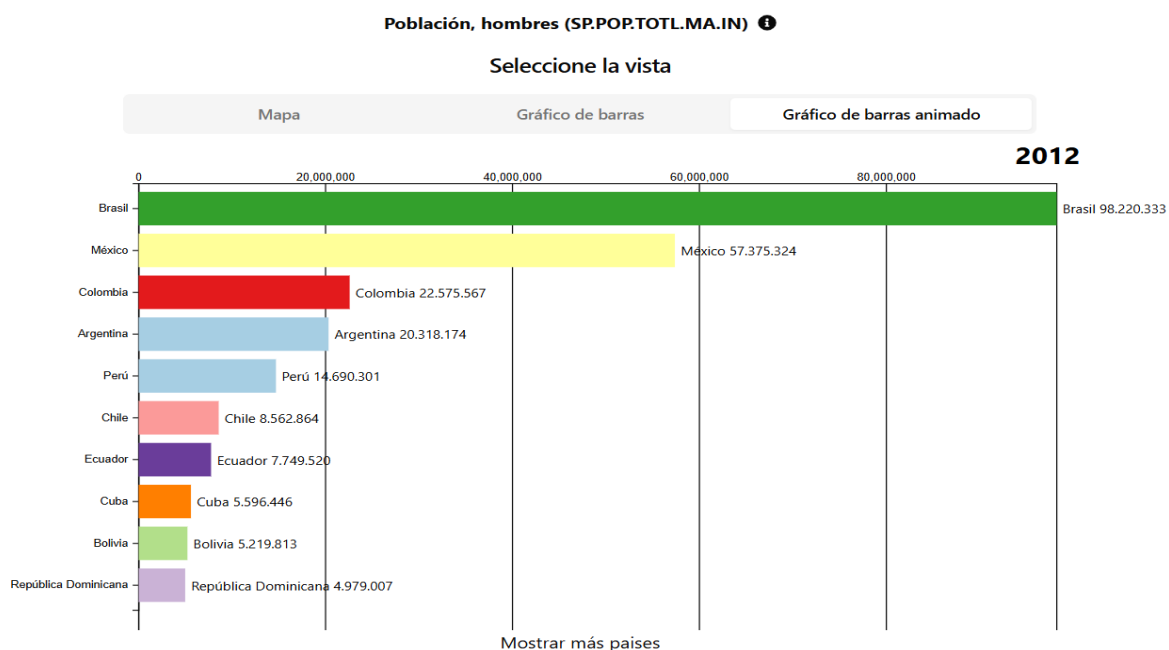


Figura 12 - Esta vista de barras horizontales animada muestra la evolución del ranking de los países a lo largo del tiempo, destacando cómo cambian sus posiciones en el indicador seleccionado.

Fuente: Elaboración propia, basada en la práctica.

Los gráficos de barras horizontales y animados cuentan con la función de “Mostrar más países”, que permiten al usuario aumentar dinámicamente el número de países mostrados, aprovechando la ventaja de estas vistas de analizar múltiples países a la vez.

## Flujo completo de la herramienta

1. La aplicación obtiene datos de todos los indicadores y países disponibles.
2. El usuario selecciona el indicador a analizar.
3. La aplicación consulta el rango de años disponibles del indicador.
4. La aplicación consulta los metadatos del indicador (español/inglés).
5. El usuario selecciona el rango de tiempo y países a consultar.
6. El usuario selecciona la vista a visualizar.
7. Se ejecuta la consulta de datos con años extendidos.
8. Obtenemos los datos del rango seleccionado.
9. Se verifican que los valores existan, y en caso de no existir, se realizan las predicciones necesarias utilizando los resultados de datos en años extendidos.
10. Se muestran los datos en las visualizaciones.

Ejemplo:

Si un usuario selecciona:

- Indicador: Población, hombres (SP.POP.TOTL.MA.IN).
- Países: Argentina y Brasil.
- Años: Del 2005 al 2010.

La aplicación:

1. Se consulta el rango de años disponibles del indicador en los países seleccionados.
2. Se consultan los metadatos del indicador en español/inglés.
3. Se obtienen datos reales para el intervalo de tiempo de 1995-2020 (años extendidos).
4. En caso de falta de valores, se utilizan los datos de los intervalos 1995-2004 y 2011-2020 para realizar las predicciones necesarias.
5. Mostrará solo datos del intervalo 2005-2010 en las vistas.

## Explicación de la Estructura y Organización del Dashboard

El dashboard está diseñado para facilitar la exploración intuitiva de indicadores de salud global, siguiendo principios de jerarquía visual y flujo lógico de análisis. A continuación, se detalla la razón detrás de cada sección:

### 1. Filtros por Tópicos

- Posición: Parte superior central
- Propósito:
  - Enfoque temático: Permite al usuario acotar primero el área de interés antes de seleccionar indicadores específicos.
  - Reducción de opciones: Evita agobiar al usuario con un listado de cientos de indicadores sin correlación.
  - Agrupación lógica: Los indicadores se agrupan según el tópico seleccionado.

### 2. Selector de Indicador

- Posición: Parte superior central debajo de los tópicos
- Propósito:
  - El usuario elige el indicador exacto después de filtrar por tópico.
  - Información técnica: Muestra códigos oficiales de los indicadores para referencia experta.
  - Jerarquía visual: Ubicado debajo de los tópicos porque permite agregar contexto previo a las opciones.

### 3. Filtros de Tiempo/Países + Función

- Posición: Sección central izquierda
- Propósito:
  - Al ubicarse luego de la selección del indicador, aseguran que el usuario ya tenga el contexto necesario para aplicar filtros más precisos.
  - Al estar ubicados en una zona visible del dashboard, siempre se puede verificar rápidamente qué filtros están activos.
  - Se alinean con el orden lógico del flujo de la aplicación: *qué ver* → *cuándo* → *dónde* → *cómo*.

### 4. Selector de Vistas

- Posición: Sección central
- Propósito:
  - Visualización adaptativa: Diferentes formatos para distintas necesidades de análisis.
  - Ubicación estratégica: Es el último paso porque depende de todos los filtros anteriores.
  - Jerarquía visual: Se muestra después de definir todos los parámetros de datos.

## Estructura del código de la herramienta

La estructura se encuentra organizada para favorecer la escalabilidad, legibilidad y mantenibilidad del código, siguiendo buenas prácticas de desarrollo. A continuación, se describe la organización del código:

### **Raíz del Proyecto**

En la raíz se ubican archivos esenciales para la configuración y ejecución de la aplicación, desde el punto de entrada de la aplicación, hasta la configuración de los estilos.

### **/src - Carpeta principal del código fuente**

Contiene todo el código de la aplicación. Su estructura se organiza en submódulos que cumplen funciones bien delimitadas:

- **/components:** Agrupa todos los componentes reutilizables de la interfaz de usuario.
- **/modals:** Incluye ventanas utilizadas para mostrar información adicional sin abandonar la vista actual.
- **/ui:** Contiene componentes visuales que forman la base de la interfaz de usuario de la herramienta
- **/views:** Contiene componentes que definen las visualizaciones que permite la herramienta.
- **/hooks:** Incluye hooks personalizados de React que encapsulan lógica reutilizable.
- **/interfaces:** Define las interfaces y tipos de TypeScript utilizados en todo el proyecto. Esto permite mantener un tipado fuerte, detectar errores en tiempo de desarrollo y mejorar la legibilidad del código.
- **/utils:** Almacena funciones utilitarias puras utilizadas en el procesamiento de datos.
- **/store:** Contiene la lógica relacionada con el estado global de la aplicación

El código fuente completo se encuentra disponible en el siguiente repositorio público de GitHub: <https://github.com/julinuzzo19/interactive-dashboard-data>

## Propuestas de ampliación del trabajo

### Mejora en las técnicas predictivas y su elección

Si bien la herramienta actual implementa predicciones utilizando técnicas de regresión simples (lineal, exponencial y logística), estas resultan efectivas solo en escenarios con comportamientos relativamente estables. Sin embargo, muchos indicadores presentan dinámicas más complejas: algunos sufren cambios abruptos debido a eventos inesperados (como pandemias, crisis económicas o intervenciones de políticas públicas), mientras que otros responden a patrones cíclicos o estacionales. En estos casos, aplicar técnicas genéricas que suponen una progresión continua y predecible puede llevar a estimaciones erróneas, disminuyendo la utilidad de los resultados para la toma de decisiones.

Actualmente, la lógica de selección de técnica se basa en un análisis combinado entre países, considerando variables como la correlación de Pearson y la tasa de crecimiento. Esto permite aplicar una única técnica al conjunto completo de datos de un indicador. Si bien esta aproximación mantiene una coherencia visual general, no contempla las particularidades estadísticas de cada indicador de forma individual.

Una mejora clave consiste en avanzar hacia un enfoque más especializado: realizar un análisis exhaustivo del comportamiento histórico de cada indicador antes de aplicar cualquier técnica de predicción. Este análisis debería incluir la detección de estacionalidades, puntos de inflexión, valores atípicos y tendencias generales. A partir de esa caracterización, se podría determinar la técnica más adecuada (lineal, exponencial, logística u otra) a aplicar por defecto, adaptando así el modelo a las particularidades del fenómeno observado en lugar de forzar una técnica única. Los pasos para implementarlo incluyen:

- Un análisis estadístico automatizado del indicador (volatilidad, tendencia, estacionalidad).
- La clasificación del tipo de curva más probable.
- La aplicación automática de la técnica óptima al conjunto de países asociados a ese indicador.

En definitiva, este cambio elevaría la calidad predictiva de la herramienta y permitiría brindar estimaciones más realistas, confiables y ajustadas a la naturaleza de cada serie temporal analizada.

### Ampliar áreas temáticas más allá de la salud

Actualmente, el sistema se centra exclusivamente en indicadores de salud. Sin embargo, la API del Banco Mundial ofrece una enorme cantidad de indicadores agrupados en diversas áreas del desarrollo: educación, economía, medio ambiente, infraestructura, pobreza, empleo, entre otros. Expandir la aplicación para incluir indicadores de estas otras áreas abriría un abanico de posibilidades de análisis mucho más amplio, permitiendo a usuarios de distintas disciplinas beneficiarse de la herramienta. Las ventajas de esta mejora serían:

- Mayor público objetivo: La aplicación dejaría de estar dirigida únicamente a especialistas en salud pública, y podría ser utilizada por economistas, sociólogos, analistas de políticas públicas y organismos internacionales.
- Análisis multidimensional: La toma de decisiones requiere de análisis integrales.

La inclusión de nuevas temáticas implicarán modificaciones y mejoras en funcionalidades de la herramienta:

- Actualmente, el selector de tópicos está pensado para un conjunto limitado de categorías relacionadas exclusivamente con salud (por ejemplo: género, eficacia de ayuda, etc.). Al incorporar indicadores de áreas como economía, educación, medio ambiente, el número de categorías aumentaría considerablemente, y la estructura del selector debería adaptarse para mantener la usabilidad.

- Revisión de las funciones de agregación: las métricas como "promedio", "máximo" o "tasa de cambio" pueden no tener el mismo significado o aplicabilidad en todos los indicadores. Se necesitaría permitir funciones de agregación específicas según el tipo de dato.

### Ampliación en la variedad de visualizaciones

Una mejora sustancial para el futuro desarrollo de la herramienta consiste en ampliar la variedad de visualizaciones disponibles. Actualmente, el dashboard se apoya en representaciones gráficas que si bien son funcionales, pueden no ser suficientes para todos los tipos de indicadores ni para todos los perfiles de usuarios. Por ello, brindar múltiples alternativas de visualización no solo mejora la experiencia de uso, sino que también favorece a una comprensión más precisa y ajustada a cada necesidad.

Por ejemplo, un gráfico de líneas es eficaz para mostrar tendencias a lo largo del tiempo, permitiendo identificar de forma clara aumentos, descensos o variaciones abruptas en las variables observadas. Su capacidad para representar series cronológicas lo convierte en una herramienta clave para analizar el comportamiento histórico de un país o comparar la evolución de un mismo indicador entre diferentes regiones.

En definitiva, aumentar la variedad de visualizaciones disponibles no representa un cambio superficial, sino que habilita un análisis mucho más abarcativo y ajustado a distintas perspectivas, permitiendo explorar los datos desde múltiples enfoques y adaptarse mejor a la complejidad de cada indicador o necesidad del usuario.

## Conclusión

La conclusión de este proyecto evidencia que, más allá de cumplir con el objetivo central de ofrecer un dashboard accesible de indicadores de salud, quedan claras oportunidades de evolución. Entre las futuras líneas de desarrollo aparece la expansión hacia otros dominios (educación, economía, medio ambiente), que no solo aumentaría el volumen de usuarios, sino que, además, permitiría comparativas multidimensionales.

Respecto a las técnicas de predicción implementadas, el uso exclusivo de regresiones lineal, exponencial y logística resultó adecuado para series cortas y tendencias sencillas. Sin embargo, estas técnicas presentan limitaciones frente a series con estacionalidad o rupturas abruptas provocadas por eventos extraordinarios. Avanzar hacia implementaciones de técnicas más avanzadas permitiría adaptar mejor las estimaciones al comportamiento real de cada indicador, aportando predicciones más sólidas y confiables.

El proyecto refuerza el rol crítico del desarrollador, este no solo realiza la tarea de codificación de la herramienta siguiendo procedimientos claros, sino que es un traductor de datos complejos en experiencias útiles para el usuario final. Las decisiones tomadas a lo largo del desarrollo van desde la selección del stack tecnológico, el procesamiento de los datos mediante técnicas de normalización y predicciones, hasta la elección de visualizaciones capaces de adaptarse a distintos escenarios y el diseño del dashboard. Estas decisiones no fueron sólo técnicas, sino estratégicas: orientadas a construir una experiencia coherente, clara y centrada en el usuario.

Más allá de cumplir con los requisitos funcionales, el verdadero valor de este proyecto no reside sólo en lo que hace, sino en cómo fue concebido: partió de la identificación de un problema real y avanzó hacia el diseño de una solución con propósito, basada en decisiones técnicas y de experiencia de usuario orientadas a necesidades concretas. El desarrollo implicó adaptarse a limitaciones, superar obstáculos con criterio y, sobre todo, mantener el foco no en simplemente “hacer que funcione”, sino en “hacer que sea útil”.

## Fuentes

Chen, D.-G., Chen, X., & Chen, J. (2020). *Reconstructing and forecasting the COVID-19 epidemic in the United States using a 5-parameter logistic growth model*. *Global Health Research and Policy*. Recuperado el 16 de mayo de 2025 de <https://doi.org/10.1186/s41256-020-00152-5>

“Chroma.js”. Disponible en <https://gka.github.io/chroma.js/>

“D3.js”. Disponible en <https://d3js.org/>

DATAtab. (s.f.). *Correlación de Pearson*. Recuperado el 2 de mayo de 2025 de <https://datatab.es/tutorial/pearson-correlation>

Debecker, A., & Modis, T. (2021). Poorly known aspects of flattening the curve of COVID-19. *Technological Forecasting and Social Change*. Recuperado el 16 de mayo de 2025 de <https://pmc.ncbi.nlm.nih.gov/articles/PMC7603980/>

GeeksforGeeks. (s.f.). *Step-by-Step Guide to Modeling Time Series Data Using Linear Regression*. Recuperado el 2 de mayo de 2025 de <https://www.geeksforgeeks.org/step-by-step-guide-to-modeling-time-series-data-using-linear-regression/?ref=asr11>

“Node.js”. Disponible en <https://nodejs.org/>

Number Analytics. (s.f.). *5 key stats: linear regression in healthcare & pharmaceuticals*. Recuperado el 2 de mayo de 2025 de <https://www.numberanalytics.com/blog/5-key-stats-linear-regression-healthcare-pharmaceuticals>

Pyle, D. (1999). Replacing Missing and Empty Values. *Data preparation for data mining*. Morgan Kaufmann Publishers. United States

Ratner, B. (2011). *Statistical and machine-learning data mining: Techniques for better predictive modeling and analysis of big data* (2.<sup>a</sup> ed.). CRC Press. United States

“React”. Disponible en <https://react.dev/>

“React Simple Maps”. Disponible en <https://www.react-simple-maps.io/>

Statistics Easily. (s.f.). *¿Qué es la regresión exponencial?*. Recuperado el 2 de mayo de 2025 de <https://es.statisticseasily.com/glosario/%C2%BFQu%C3%A9-es-la-regresi%C3%B3n-exponencial-explicada-en-detalle%3F/>

“Tailwindcss”. Disponible en <https://tailwindcss.com/>

The World Bank Group. (s.f.). *Developer Information*. Recuperado el 2 de mayo de 2025 de <https://datahelpdesk.worldbank.org/knowledgebase/topics/125589-developer-information>

The World Bank Group. (s.f.). *World Development Indicators Database*. Recuperado el 2 de mayo de 2025 de <https://databank.worldbank.org/source/world-development-indicators>

Villalobos-Arias, M. A. (2020). *Using generalized logistics regression to forecast population infected by Covid-19*. Universidad de Costa Rica. Recuperado el 2 de mayo de 2025 de <https://www.kerwa.ucr.ac.cr/bitstreams/9a7e23c5-d627-4eb5-8324-478ff7828517/download>