

Junta de Castilla y León

---

PandemIA CyL: Monitorización de Hospitales por  
COVID-19 en Castilla y León

---



VIII Concurso de Datos Abiertos

**Autores**

Álvaro García Barragán  
Alberto González Calatayud  
Laura Masa Martínez  
Enrique Solera Navarro

Septiembre, 2024

---

## Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Fuentes de datos</b>	<b>1</b>
<b>3. Desarrollo del Proyecto</b>	<b>1</b>
3.1. Análisis descriptivo . . . . .	1
3.2. Desarrollo de la herramienta . . . . .	2
3.3. Futuro trabajo . . . . .	4
<b>4. Impacto y valor del proyecto</b>	<b>4</b>
<b>5. Accesibilidad</b>	<b>5</b>

---

## 1. Introducción

La pandemia de COVID-19 nos ha mostrado lo frágil que puede ser nuestro sistema de salud ante crisis sanitarias. A medida que los casos aumentaban, muchos hospitales se vieron abrumados, lo que afectó la atención que podían brindar a los pacientes. Esta situación no solo destacó la importancia de contar con recursos suficientes, sino que también puso de relieve la necesidad de herramientas que nos ayuden a gestionar mejor la atención médica en momentos críticos.

Este proyecto se propone abordar estos desafíos mediante el desarrollo de una herramienta interactiva que permita a los usuarios visualizar datos clave relacionados con la hospitalización y la incidencia de COVID-19. Además del análisis descriptivo, la herramienta no solo identifica en tiempo real qué hospitales están experimentando mayor saturación, sino que su capacidad para analizar la evolución temporal por provincias ofrece la oportunidad de elaborar hipótesis y realizar predicciones sobre futuros brotes. Esta información es crucial, ya que brinda la oportunidad de tomar decisiones administrativas para aliviar la presión en los centros más sobrecargados y permite a los pacientes elegir un hospital alternativo.

## 2. Fuentes de datos

En el desarrollo de este proyecto se han utilizado los siguientes conjuntos de datos disponibles en el Portal de Datos Abiertos de la Junta de Castilla y León:

- [situacion-epidemiologica-coronavirus-en-castilla-y-leon.csv](#): Proporciona una visión general de la evolución epidemiológica en las provincias de Castilla y León, incluyendo casos confirmados, nuevos positivos, altas y fallecimientos.
- [situacion-de-hospitalizados-por-coronavirus-en-castilla-y-leon.csv](#): Conjunto de datos principal para la herramienta interactiva, que contiene información detallada sobre la situación hospitalaria, como el número de hospitalizados en planta y UCI, altas y fallecimientos. Este conjunto es seleccionado por su riqueza en parámetros relevantes para visualizar la carga sobre el sistema de salud.
- [situacion-afectados-por-coronavirus-en-atencion-primaria.csv](#): Se centra en la atención primaria, proporcionando detalles sobre el número de enfermos y la tasa de enfermedad por provincia.

## 3. Desarrollo del Proyecto

Esta sección aborda la creación de una herramienta interactiva para visualizar y analizar datos de COVID-19 en Castilla y León, dividida en dos partes: primero, un análisis descriptivo de los datos que proporcionó las bases para entender las tendencias clave; y segundo, el desarrollo de la herramienta interactiva que permite a los usuarios explorar la información de manera accesible. A continuación, se presentan estos dos aspectos.

### 3.1. Análisis descriptivo

Previo al desarrollo de la aplicación, se realizó un análisis descriptivo de los datos de COVID-19 para identificar patrones y tendencias clave en las hospitalizaciones y la incidencia del virus en diversas provincias y hospitales.

Se realizó un preprocesamiento de datos para estandarizar formatos de fecha y se generaron visualizaciones temporales de indicadores clave como casos confirmados y fallecimientos. También se exploraron las relaciones entre variables, como la correlación entre la ocupación de UCI y los fallecimientos. Este análisis proporcionó una comprensión más profunda de los factores que afectan la capacidad hospitalaria y es esencial para guiar el diseño de la herramienta y facilitar el acceso a métricas relevantes para los usuarios. En las siguientes gráficas se pueden ver los datos representados temporalmente en función de algunas variables.

La [Figura 1](#) muestra patrones claros de las ola de la pandemia por hospital, destacando periodos críticos donde intervenciones oportunas podrían haber reducido la propagación.

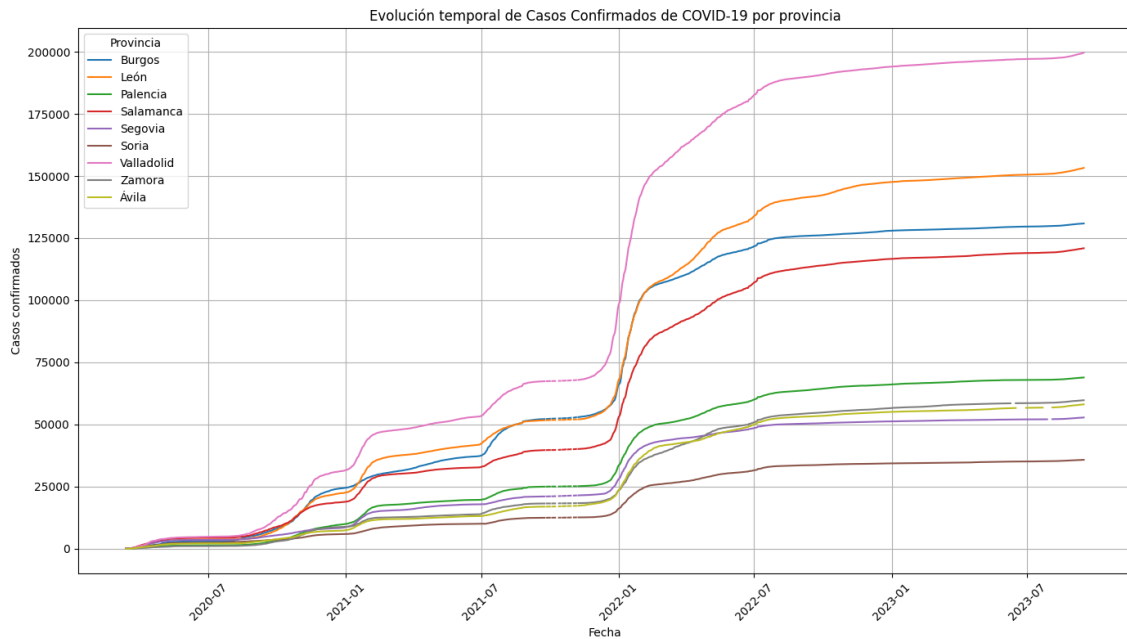


Figura 1: Evolución temporal de casos confirmados por provincia

Por otra parte, la [Figura 2](#) representa cómo el número de hospitalizados en planta decayó sucesivamente a medida que avanzaba la pandemia.

Además, en la [Figura 3](#) se puede observar que el pico de enfermos ocurrió en enero de 2022, momento en el cual los hospitales de Valladolid registraron la mayor cantidad de pacientes.

### 3.2. Desarrollo de la herramienta

A partir del análisis descriptivo, se desarrolló una herramienta interactiva para poder visualizar y analizar datos sobre COVID-19 en Castilla y León. La herramienta permite a los usuarios explorar diferentes métricas relacionadas con la situación hospitalaria, tales como el número de nuevos hospitalizados en planta y UCI, el porcentaje de ocupación en UCI, el número de altas y fallecimientos, entre otros.

Las características principales de esta aplicación son:

1. La interactividad: Los usuarios pueden seleccionar la fecha específica que desean analizar utilizando un control deslizante o un selector de calendario, permitiendo así un fácil análisis de la evolución de la pandemia día a día.
2. La amplia selección de métricas: La herramienta ofrece la posibilidad de elegir entre diferentes datos para representar en el mapa. Las métricas disponibles incluyen nuevos hospitalizados en planta y UCI, hospitalizados actuales, nuevas altas y fallecimientos, lo que brinda una visión completa de la situación hospitalaria.

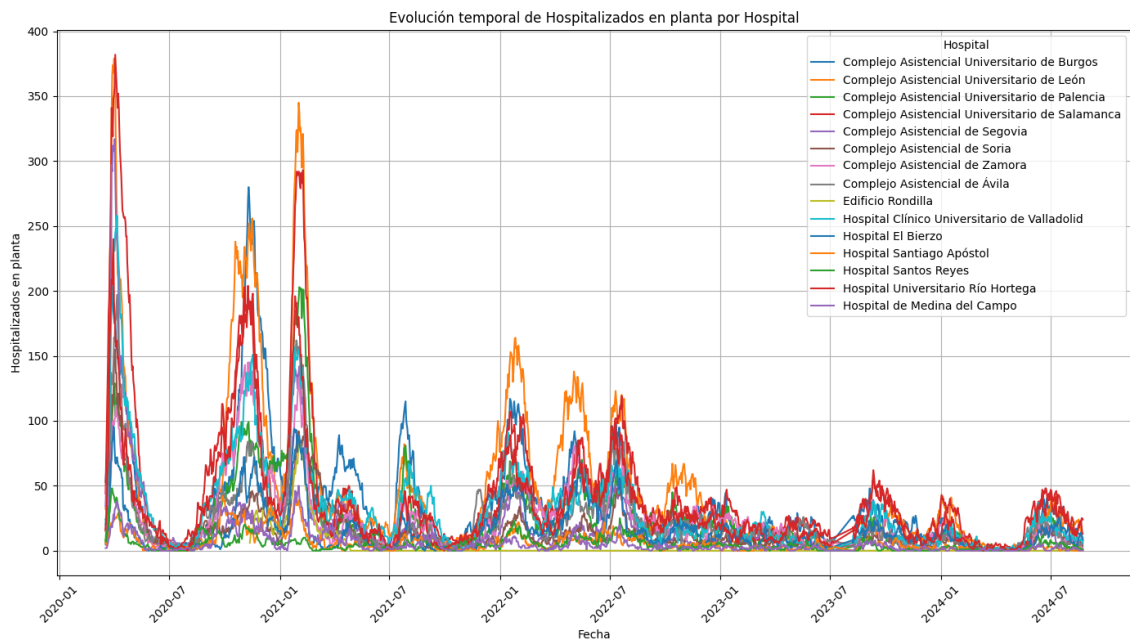


Figura 2: Evolución temporal de número de hospitalizados por hospital

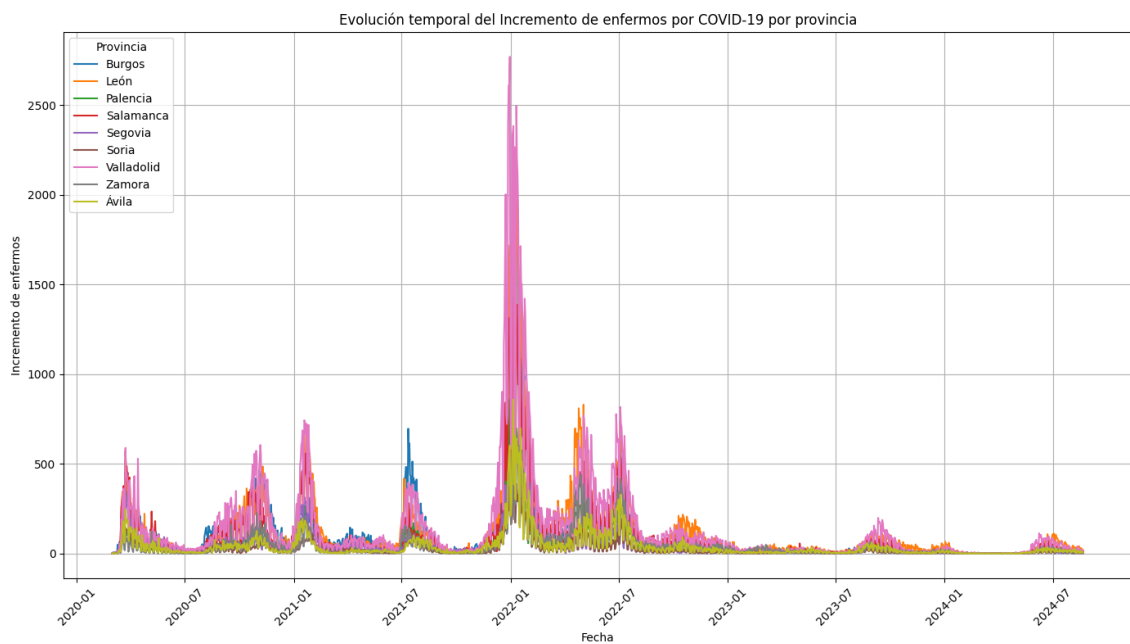
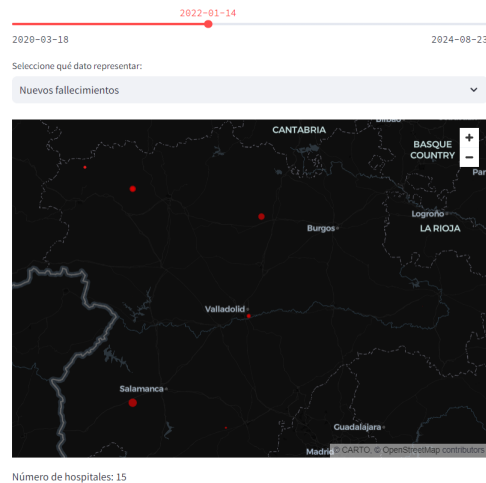


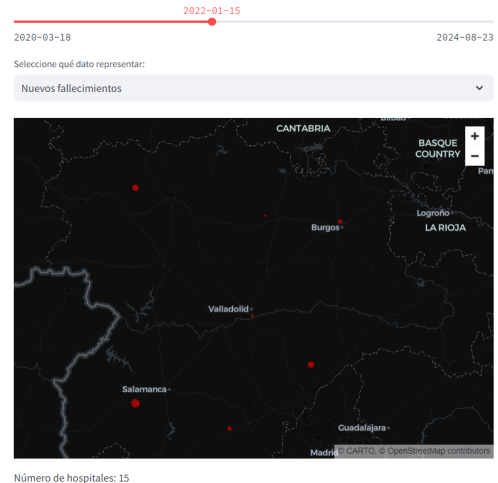
Figura 3: Evolución temporal del incremento de enfermos por provincia

3. La visualización geográfica: Los datos seleccionados se visualizan en un mapa de Castilla y León. Los hospitales se representan mediante círculos rojos cuyo tamaño varía en función de la densidad de la variable seleccionada.
4. La normalización de los valores: Las métricas han sido normalizadas para garantizar que las variaciones en el tamaño de los círculos sean proporcionales a la magnitud de los datos y se permita verlos adecuadamente.

Las siguientes imágenes [Figura 4](#) muestran el mapa interactivo en dos fechas consecutivas representando el número de nuevos fallecidos, lo que permitiría identificar patrones en la evolución de la pandemia y cómo varía entre hospitales.



(a) Mapa del número de nuevos fallecidos el 14/01/2022.



(b) Mapa del número de nuevos fallecidos el 15/01/2022.

Figura 4: Mapa interactivo de características cuantificables por hospital.

### 3.3. Futuro trabajo

Como líneas de trabajo futuras, se planea incorporar un modelo predictivo que, a partir de la posición geográfica de un usuario (latitud y longitud), recomiende el hospital más adecuado en función de su menor saturación, tomando como referencia las métricas proporcionadas. Este modelo utilizará técnicas de clustering y algoritmos de recomendación para proporcionar sugerencias en tiempo real, ayudando a mejorar la distribución de pacientes y aliviar la carga de los centros hospitalarios más saturados durante situaciones epidémicas.

Además, para la gestión por parte de las autoridades sanitarias y los gestores de hospitales, se abre la posibilidad de utilizar modelos predictivos basados en series temporales. Dado un conjunto de datos entre dos fechas consecutivas, se podrá predecir cuántos nuevos pacientes se estima que llegarán al hospital.

## 4. Impacto y valor del proyecto

Este proyecto tiene un impacto significativo en la gestión hospitalaria durante epidemias al ofrecer datos en tiempo real sobre la saturación de los hospitales. Facilita decisiones informadas para optimizar la distribución de recursos y pacientes, mejorando la calidad de la atención médica. Al permitir identificar hospitales menos saturados, contribuye a aliviar la presión en aquellos que están sobrecargados, lo que resulta en una respuesta más ágil durante situaciones críticas.

Desde una perspectiva social, el proyecto promueve la transparencia en la gestión de datos públicos y empodera a los pacientes, permitiéndoles elegir hospitales que se ajusten mejor a sus necesidades. Su originalidad radica en el uso de tecnologías interactivas para visualizar datos de manera accesible, lo que no solo es una solución técnica, sino que también aporta un valor social crucial para enfrentar futuras emergencias de salud.

---

## 5. Accesibilidad

La herramienta está disponible a través de una URL (<https://pandem-ia-cyl.streamlit.app/>), facilitando el acceso a investigadores y responsables de salud pública. Esto permite que el proyecto tenga un impacto directo en la toma de decisiones, especialmente en crisis sanitarias como la pandemia de COVID-19.

Además, el código está liberado bajo una licencia open source, por lo que otras personas pueden mejorarlo y contribuir al proyecto a través de este enlace: <https://github.com/Alvaro8gb/COVID-CyL>.

## Agradecimientos

Agradecemos a la Junta de Castilla y León por su compromiso con la transparencia y la disponibilidad de datos abiertos, especialmente de este tipo de datos de alto impacto social. Su aportación ha sido fundamental para el desarrollo de este proyecto y el análisis de la situación sanitaria. La calidad de los datos es muy alta, y el procesamiento necesario ha sido mínimo.