

PREVENCIÓN DE INUNDACIONES SOBRE INFRAESTRUCTURAS DE LA CIUDAD DE VALENCIA



Rubén Sierra Punter
Alberto Dalmau García
Hugo Frasquet García
Pablo Pons Del Río
Maksym Myronenko

Visualización de Datos 2025 2º Grado Ciencia de Datos



Contenido

1. Introducción.....	3
2. Metodología.....	4
2.1. Análisis exploratorio de datos	4
2.2. Preprocesado de datos	4
2.3. Preprocesado de la información geográfica	5
2.4. Elección de gráficas para los distintos tipos de datos.....	7
2.5. Diseño del mapa e interactividad	9
2.6. Diseño del cuadro de mandos.....	11
2.7. Implementación.....	11
3. Resultados	12
4. Discusión.....	17
5. Conclusiones	19
6. Referencias	20



1. Introducción

En el año pasado tuvo lugar la Dana 2024, un episodio muy dramático para los valencianos y tuvo, como consecuencia, daños materiales en varias zonas de Valencia y víctimas mortales. Dicho episodio ha hecho que los valencianos tengamos que tomar más conciencia sobre el tema de las inundaciones y tomar medidas de prevención.

Por tanto, en este trabajo buscamos hacer un análisis de las infraestructuras como colegios, hospitales, centros de salud... que pueden verse afectadas en posteriores inundaciones y proponer soluciones para prevenir estos riesgos y que no pase como en la Dana 2024. Para ello, nos hemos centrado en controlar el riesgo de inundación en las zonas de la provincia de Valencia.

Para reflejar los objetivos que nos hemos puesto, hemos empleado QGIS para hacer un análisis espacial y visualizar los datos de la provincia de Valencia y R para crear los mapas interactivos con los que tratamos de que se visualicen las infraestructuras en riesgo dentro de un entorno Shiny.

En definitiva, trataremos de dar a los ciudadanos valencianos un entorno en el cual podrá ver las infraestructuras que se encuentran en riesgo y posibles soluciones ante una hipotética inundación.



2. Metodología

2.1. Análisis exploratorio de datos

Los datos utilizados para este trabajo son:

- **Zonas de inundación frecuente (T=10 años):** Capa que sirve para ver las zonas más expuestas a inundación en un periodo de retorno de 10 años, es decir, aquellas zonas en las que hay un 10% de probabilidad de que ocurra anualmente y un 92.8% de que ocurra una vez en un periodo de 25 años. Dato georreferenciado (shapefile).
- **Zonas inundables de probabilidad media u ocasional (T=100 años):** Capa que sirve para ver las zonas expuestas a inundación en un periodo de retorno de 100 años, es decir, aquellas zonas en las que hay un 1% de probabilidad de que ocurra anualmente. Dato georreferenciado (shapefile).
- **Zonas inundables de probabilidad baja o excepcional (T=500 años):** Capa que sirve para ver las zonas expuestas a inundación en un periodo de retorno de 500 años, es decir, aquellas zonas en las que hay un 0.2% de probabilidades de que ocurra anualmente. Dato georreferenciado (shapefile).
- **Provincias de España:** Capa que sirve para poder filtrar las zonas expuestas a inundaciones de la provincia de Valencia. Dato georreferenciado(shapefile).
- **Colegios de la provincia de Valencia:** Colegios de la provincia de Valencia para, posteriormente, ver que colegios están en riesgo. Dato no georreferenciado (csv).
- **Hospitales y centros de salud de la provincia de Valencia:** Centro de salud de la provincia de Valencia para, posteriormente, ver que colegios están en riesgo. Dato no georreferenciado (csv).

2.2. Preprocesado de datos

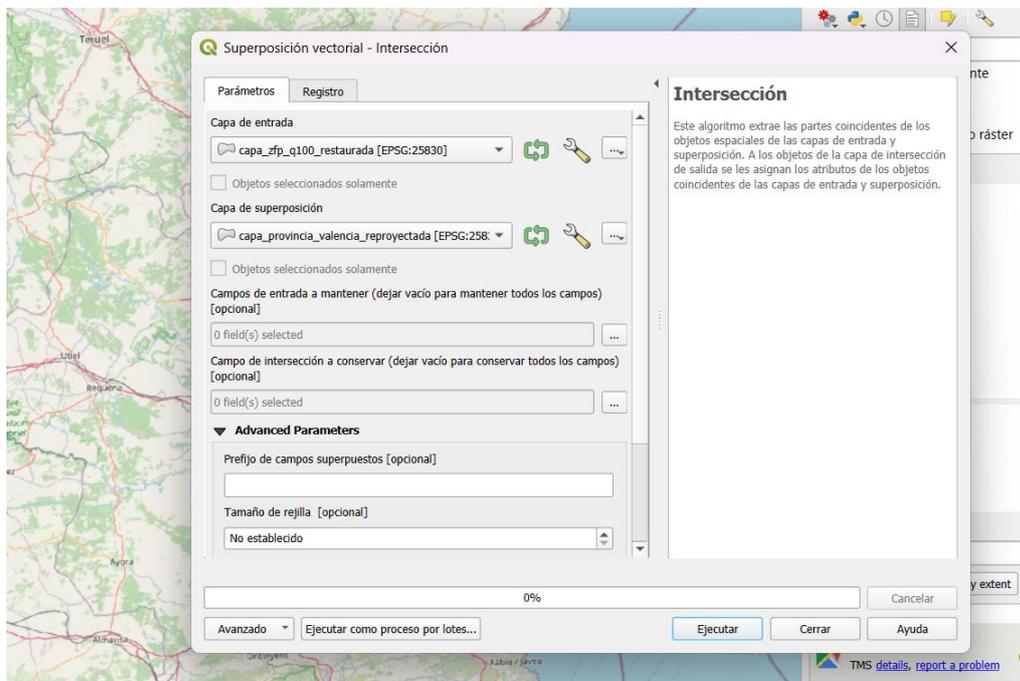
Para llevar a cabo el preprocesado de datos hemos hecho:

- Transformar atributos como códigos postales a caracteres para facilitar uniones con otras tablas (**as.character**).
- Renombrar variables para que sean entendibles por los miembros del trabajo y nos faciliten el procedimiento de la aplicación (**rename**).
- Unión de hospitales y centros de salud con los códigos postales y municipios para añadir información geográfica (**left_join**).
- Transformar el sistema de coordenadas a EPSG:4326 ya que es el sistema compatible con las aplicaciones cartográficas y permiten la visualización correcta de las capas en los mapas (**st_transform**).
- Convertir a minúsculas todos los nombres para evitar que un municipio pueda salir dos veces por diferencias en la escritura (**tolower**).

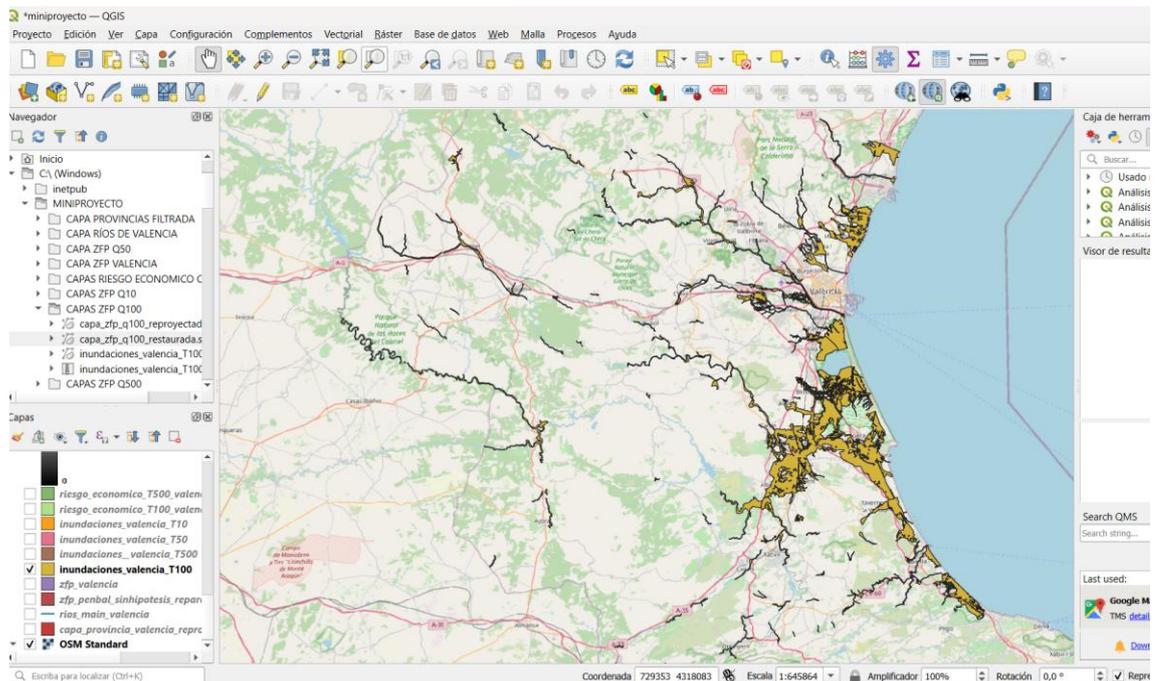
2.3. Preprocesado de la información geográfica

Los pasos realizados para la el preprocesado de la información geográfica son:

- 1. Carga y exploración de capas:** Hemos accedido a la tabla de atributos de cada capa verificando que la información estuviera completa y estructurada correctamente. Tras este análisis, hemos comprobado que no existen valores faltantes y que, por tanto, no ha hecho falta realizar limpieza de datos en esta etapa. Este paso es clave para evitar posibles errores y garantizar que los datos sean fiables.
- 2. Reproyección de capas:** Hemos reprojectado todas las capas al sistema de coordenadas EPSG:25830 (ETRS89 / UTM zona 30N). Esta operación ha sido necesaria para poder hacer las operaciones espaciales necesarias con coherencia ya que algunas capas tenían proyecciones diferentes.
- 3. Corrección de geometrías:** Al añadir las capas en qgis, hemos detectado imperfecciones y errores geométricos en algunos polígonos. Estos errores han sido por estructuras inconsistentes que lo que hacen es dificultar los procesos de análisis espacial. Para solucionarlo, hemos decidido utilizar la herramienta de “corrección de geometrías”, la cual ha hecho que todas las geometrías sean válidas. Por tanto, al aplicar esta solución ya podemos realizar la operación espacial que pretendemos.
- 4. Superposición espacial:** Al tener todas las capas preparadas, hemos realizados una superposición espacial a través de la intersección. Para ello, hemos utilizado como capa de entrada las diferentes capas de zonas de riesgo de inundación y la hemos intersecado con la capa de la provincia de Valencia. Por tanto, hemos obtenido una capa que refleja las áreas dentro de la provincia que presentan riesgo de inundación. En definitiva, la intersección nos ha permitido delimitar el análisis geométrico a la zona de interés.

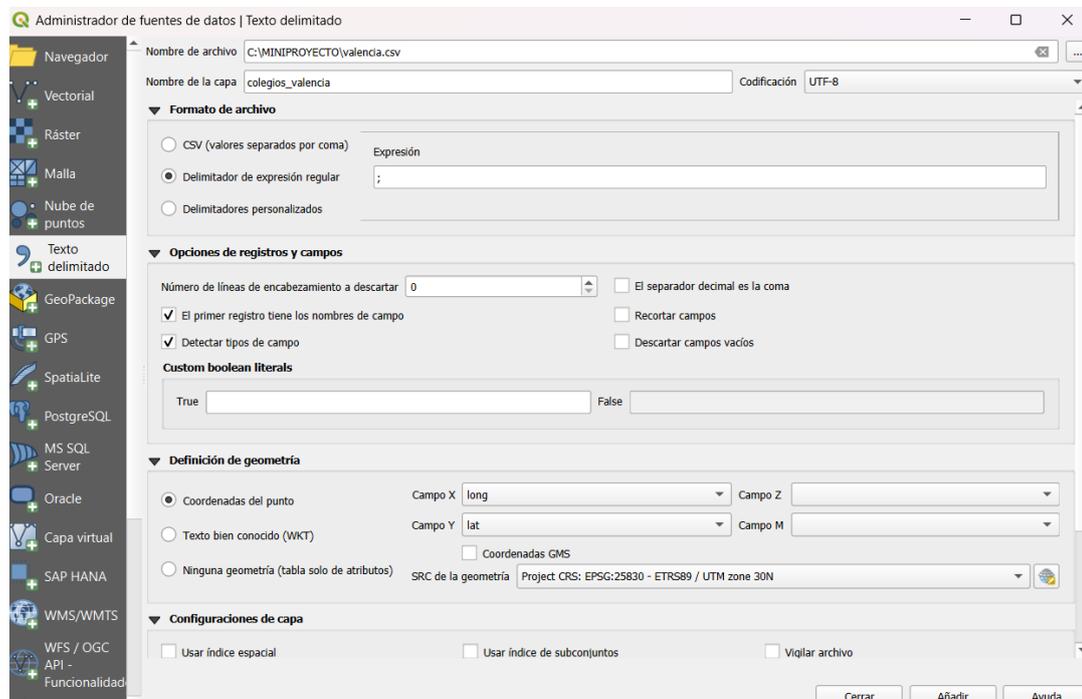


Intersección realizada de ambas capas



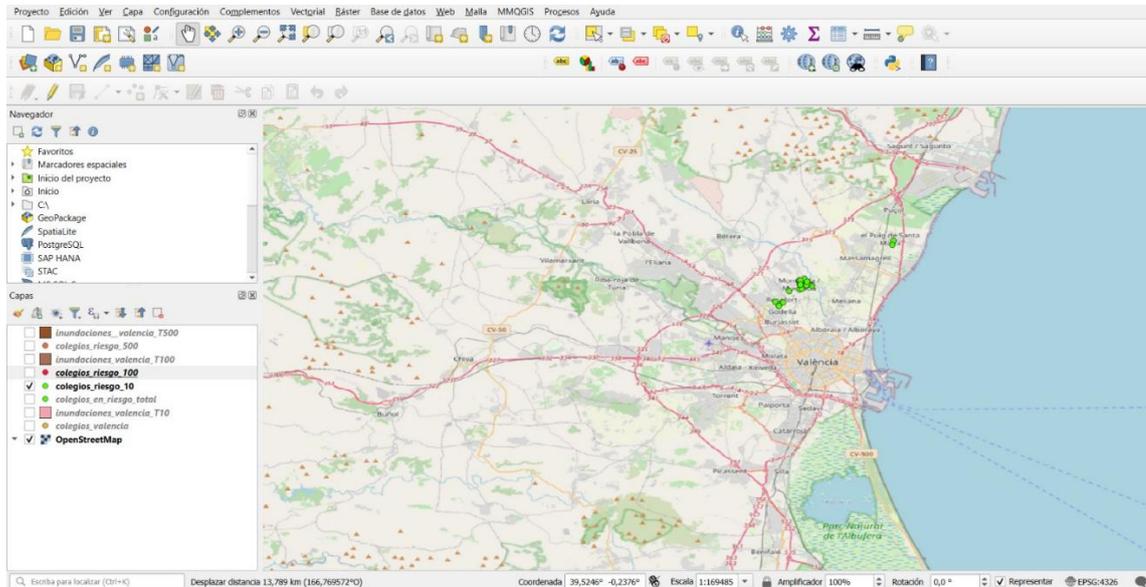
Resultado de la intersección realizada

- Añadir capas de infraestructuras:** Una vez ya hemos tenido todas las capas de las zonas de riesgo para cada periodo de retorno con la superposición espacial aplicadas, hemos añadido la capa de los colegios, hospitales y centros de salud. Para ello, hemos añadido capa de texto delimitado y hemos metido el csv.



En esta imagen se ve cómo hemos añadido la capa de colegios, este procedimiento lo hemos aplicado tanto para los hospitales como para los centros de salud.

- 6. Superposició espacial:** Tras tener la capa de colegios lista, realizamos una intersección con las capas de zonas de riesgo de inundación para cada periodo de retorno. De esta forma, clasificamos a cada colegio, hospital y centro de salud según su nivel de riesgo de inundación y, por tanto, identificar qué infraestructuras tienen mayor exposición a inundaciones y evaluar posibles medidas preventivas para dichas zonas.



En esta imagen podemos ver los colegios de la provincia de Valencia con riesgo de inundación de retorno 10. Esto mismo lo hemos hecho con los demás riesgos que tenemos en nuestro trabajo, tal y como se ve en la imagen.

2.4. Elección de gráficas para los distintos tipos de datos

En esta sección, se describen las gráficas seleccionadas para representar los diferentes tipos de datos utilizados en el análisis de prevención de inundaciones en la provincia de Valencia. La elección de las gráficas se basa en la naturaleza de los datos y los objetivos de visualización

En primer lugar, para representar las zonas de riesgo de inundación, hemos utilizado mapas de riesgo de inundación. Estos mapas muestran las áreas con diferentes niveles de probabilidad de inundación, utilizando capas geoespaciales que indican zonas de inundación para periodos de retorno de (T=10 años), (T=100 años), (T=500 años).

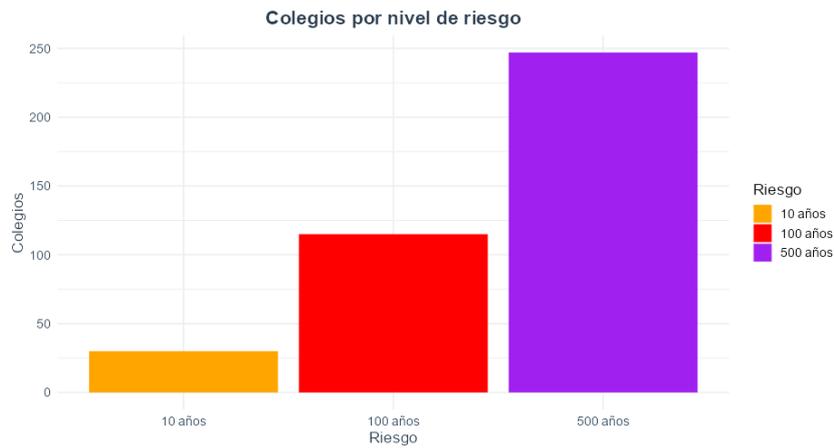
Los conjuntos de datos de las infraestructuras que hemos estudiado tenían la misma naturaleza y eran del mismo tipo, por lo que hemos utilizado gráficos de barras, por lo que hemos representado dos gráficos de barras para cada tipo de infraestructura. Los gráficos eran los siguientes:

- Hemos representado el número de hospitales, centros de salud o colegios que hay en cada municipio, por lo que teníamos una variable categórica (municipios) y una variable de conteo o frecuencia discreta (nº de hospitales, colegios...).
- Hemos representado el número de hospitales, centros de salud o colegios en riesgo según el periodo de retorno (T), por lo que tenemos una variable categórica de nuevo (T) y una variable de frecuencia que es discreta (nº de colegios, centros de salud...).



Debido a estas características y los tipos de datos que teníamos hemos escogido este tipo de gráficos.

Estos serían dos de las gráficas que os he comentado:

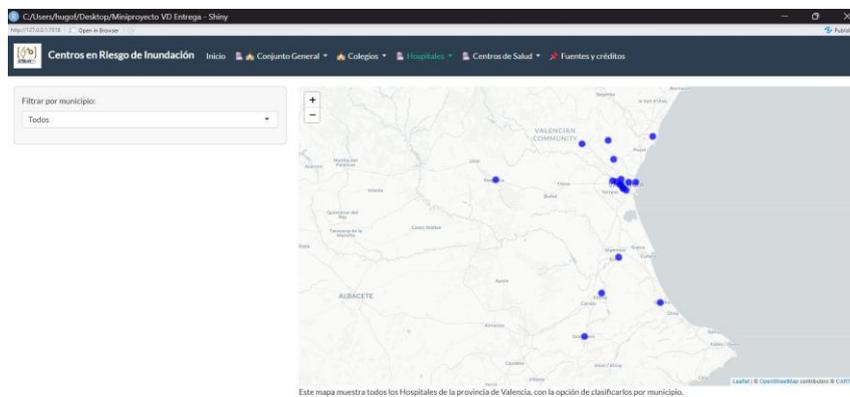


2.5. Diseño del mapa e interactividad

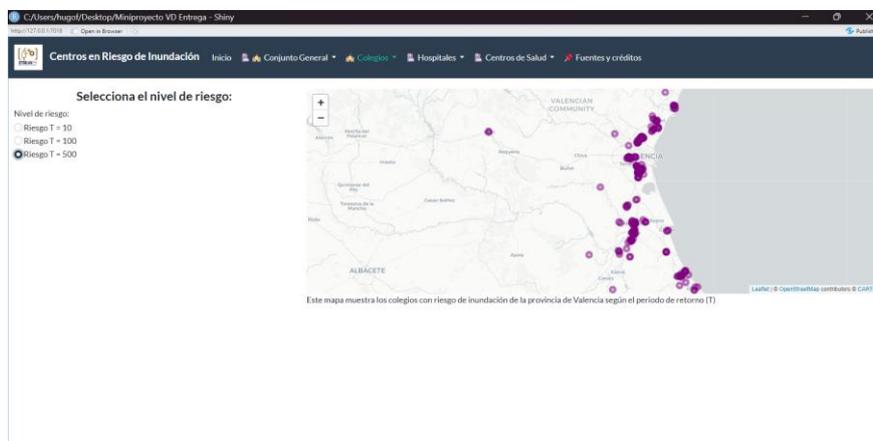
El diseño del mapa se ha centrado en ofrecer una visualización clara, accesible y funcional de nuestro estudio. Se ha empleado la librería *Leaflet* en R, que permite representar información geoespacial de forma interactiva.

En cuanto a los mapas e interactividad, hemos diseñado tres tipos de mapas distintos:

En primer lugar, hemos diseñado un mapa general para cada apartado del estudio, de modo que hemos hecho un mapa con los tres tipos de centros en la provincia de Valencia, y uno por separado para cada tipo, estos centros representados por puntos. Estos mapas tenían la interactividad de poder seleccionar los centros por municipio, de modo que se facilita mucho la búsqueda de estos centros. A parte se puede navegar y hacer zoom con total libertad por el mapa y si dejas el cursor del ratón sobre un centro te dice su nombre, así se puede diferenciar y buscar centros de una forma óptima.

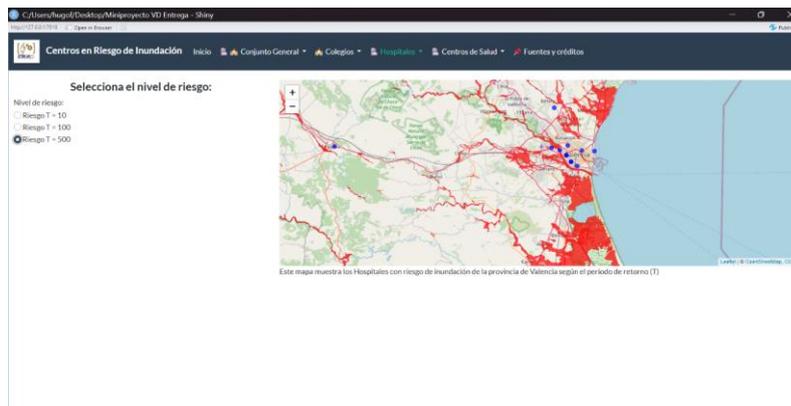


En segundo lugar, hemos hecho un mapa de centros con riesgos por periodo de retorno, de modo que tenemos una interfaz donde se puede seleccionar un valor de T (10, 100 o 500 años). Este mapa esta por separado para cada tipo de centros. A parte se puede navegar y hacer zoom con total libertad por el mapa y si dejas el cursor del ratón sobre un centro te dice su nombre, así se puede diferenciar y buscar centros de una forma óptima.

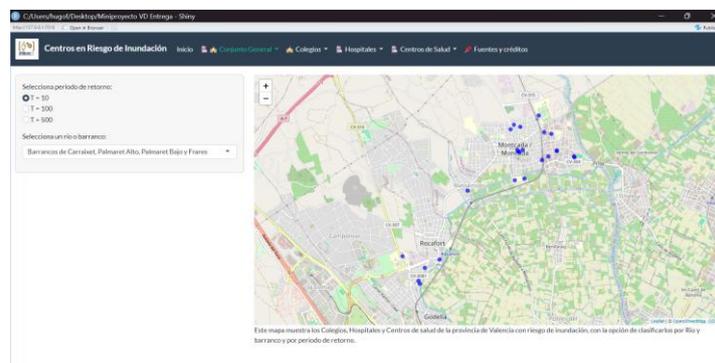




En este tipo de mapa hemos aplicado ciertas diferencias al mapa de hospitales y centros, mientras que en los colegios simplemente hemos añadido los colegios en riesgo marcados con un punto para cada valor de T, en hospitales y en centros de salud, como para algunos valor de T no había ningún edificio afectado hemos añadido todos los edificios de la provincia y hemos añadido también las zonas con riesgo de inundación coloreadas en rojo, de modo que se puede ampliar y navegar por el mapa para así poder observar que ningún punto está sobre las zonas de riesgo.

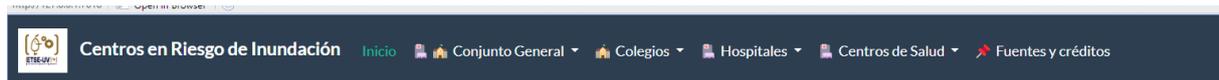


Por último, hemos hecho un mapa en conjunto para los tres tipos de centros donde se pueden observar los edificios en riesgo para cada valor de T y además con la posibilidad de poder clasificar las zonas de riesgo por Barrancos y Ríos.

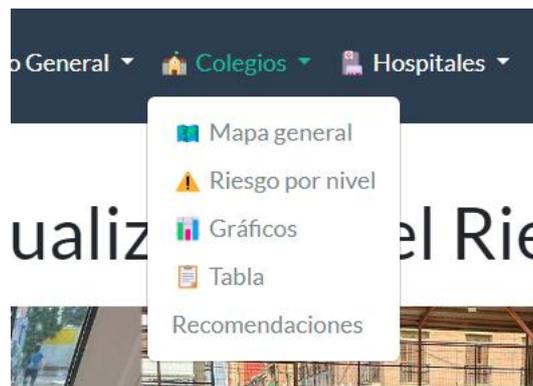


2.6. Diseño del cuadro de mandos

El cuadro de mandos ha sido diseñado con el objetivo de facilitar la navegación entre diferentes capas de información relevantes. La interfaz presenta una estructura clara y jerarquizada, con una barra de navegación superior que organiza el contenido en distintas pestañas: *Inicio*, *Conjunto General*, *Colegios*, *Hospitales*, *Centros de Salud* y *Fuentes y créditos*. Este enfoque permite al usuario acceder de forma segmentada a los distintos tipos de infraestructuras críticas, manteniendo en todo momento una experiencia coherente y fluida.



Cada sección del cuadro de mandos incorpora funcionalidades interactivas adaptadas al tipo de contenido. Por ejemplo, en la pestaña de colegios se incluye un filtro por municipio en la barra lateral izquierda, lo cual permite adaptar el contenido visualizado en el mapa a las necesidades del usuario. Además, se ha procurado un diseño limpio y con buen contraste visual para favorecer la accesibilidad, utilizando iconos y colores representativos que refuerzan la comprensión inmediata de los datos.



2.7. Implementación

Como ya hemos explicado anteriormente para la implementación en R, hemos desarrollado una aplicación web interactiva utilizando principalmente los paquetes **Shiny** y **Leaflet**. El objetivo ha sido facilitar la consulta y visualización de los **centros educativos y sanitarios** en riesgo de inundación en la provincia de Valencia, según distintos escenarios de peligrosidad (periodos de retorno de 10, 100 y 500 años).

La interfaz de la aplicación ha sido construida con el paquete shiny, que permite el desarrollo de aplicaciones web interactivas directamente desde R. Para facilitar la navegación, la app está organizada mediante un sistema de **pestañas** (navbarPage), cada una dedicada a un aspecto concreto del estudio: vista general, análisis por tipo de centro, mapas dinámicos por escenario de riesgo, rankings de municipios y recomendaciones.

- Para la interactividad hemos incluido distintos elementos como: **Radio buttons** para seleccionar el periodo de retorno ($T = 10, 100, 500$ años).
- **Selectores de municipios** para filtrar la información.
- **Pestañas internas** por tipo de centro (colegios, hospitales, centros de salud).
- **Textos dinámicos** que muestran recomendaciones adaptadas al nivel de riesgo seleccionado.

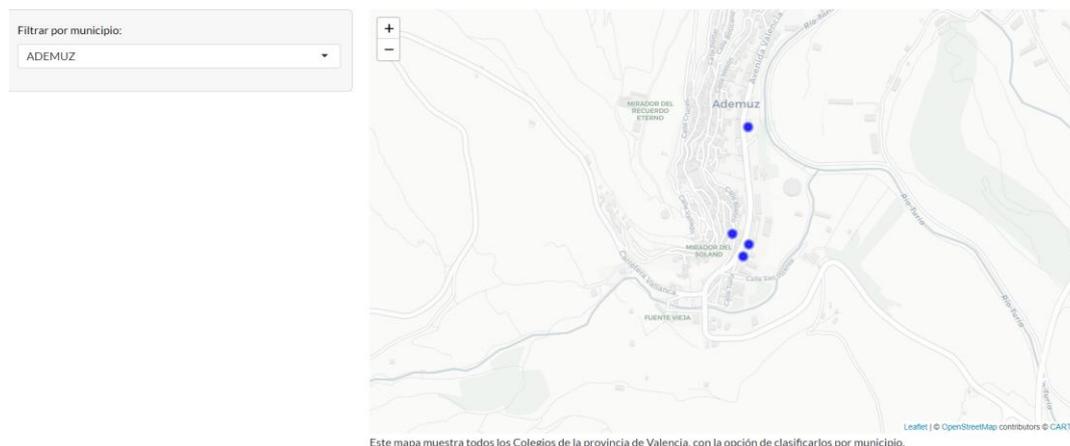
Para la visualización geográfica de los resultados se ha empleado el paquete leaflet, que permite representar los datos espaciales sobre un mapa interactivo. En la aplicación:

- Cada centro se muestra como un punto geolocalizado.
- Se utilizan colores y estilos diferenciados para representar el tipo de centro o el nivel de riesgo.
- Los mapas se actualizan automáticamente al cambiar el escenario o el filtro de municipio.

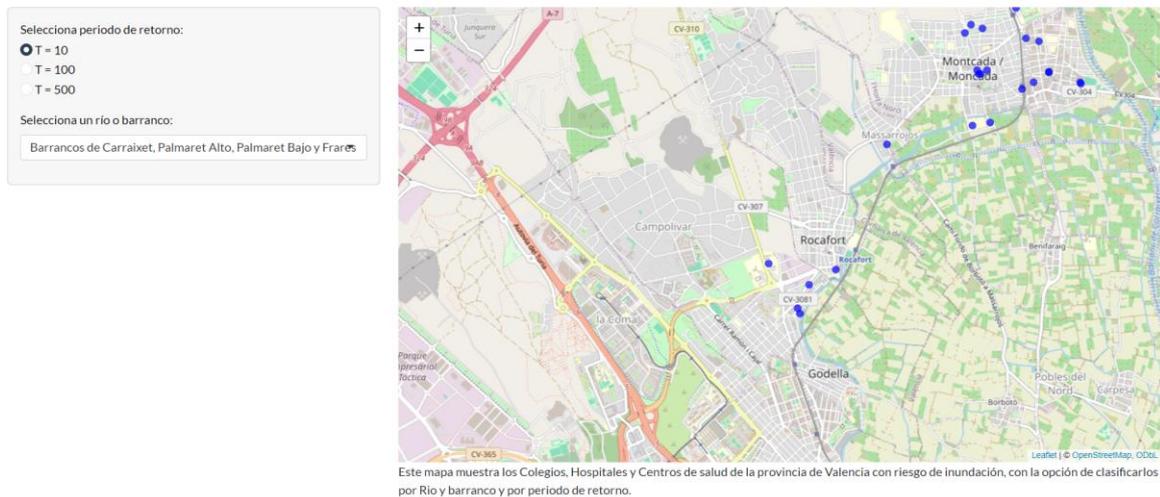
Gracias a esta implementación obtenemos una app muy dinámica y con una interfaz muy fácil de usar, donde hemos podido mostrar a la perfección los objetivos y resultados del estudio.

3. Resultados

MAPAS GENERALES



Mapa que muestra todos los centros (colegios, centros de salud, hospitales) de la provincia de Valencia con la opción de filtrarlos por municipio. En la imagen se ha seleccionado el municipio de Ademuz, por tanto, se visualizan solamente los colegios situados en esa zona representándose con un punto azul. El cuadro de selección situado a la izquierda permite cambiar el municipio a analizar.

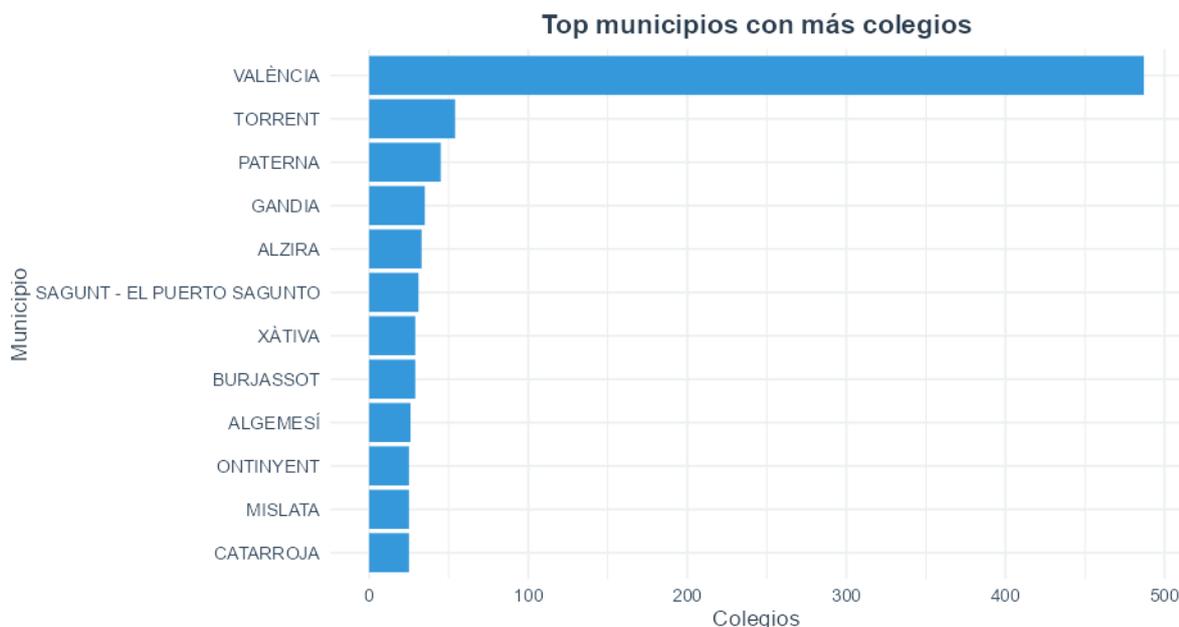


Este mapa refleja los centros de la provincia de Valencia que están en riesgo de inundación según su proximidad a ríos o barrancos para cada valor de T (10, 100, 500 años). Los puntos marcados expresan los centros dentro de la provincia que pueden verse afectados por crecidas en base al periodo de retorno que se elija.

COLEGIOS



Este mapa muestra los colegios de la provincia de Valencia que están en riesgo de inundación según el periodo de retorno (T), el cual indica la probabilidad de que ocurra un evento extremo en un periodo determinado. En la imagen podemos ver unos puntos rojos que representan los colegios con riesgo de inundación para cada periodo de retorno, el cual se puede cambiar gracias al cuadro de selección situado a la izquierda.



Este gráfico de barras muestra los municipios con mayor cantidad de colegios en la provincia de Valencia. Las conclusiones que se extraen son:

- Valencia es el municipio que más colegios tiene con mucha diferencia respecto a los demás.
- Torrent, Paterna y Gandía son los siguientes que más colegios tienen, pero con cantidades bastante menores.

En definitiva, el gráfico refleja que a mayor población hay mayor necesidad educativa, por tanto, mayor número de colegios.

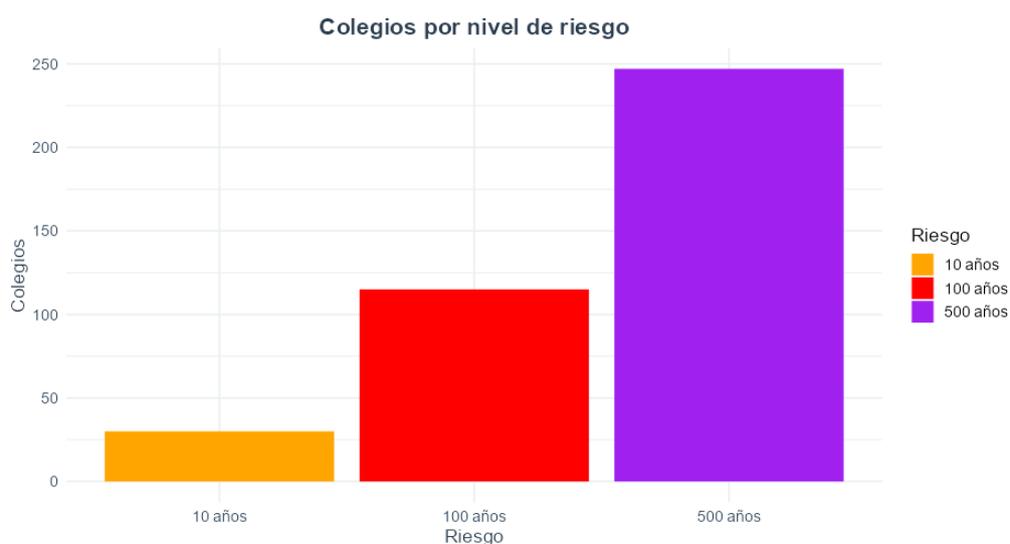
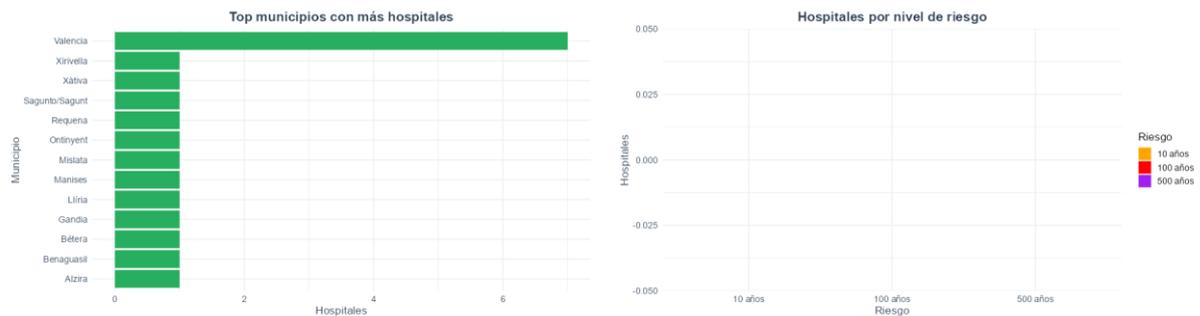


Gráfico de barras que refleja la cantidad de colegios en la provincia de Valencia clasificados según su nivel de riesgo de inundación. Observando el gráfico, podemos ver que el número

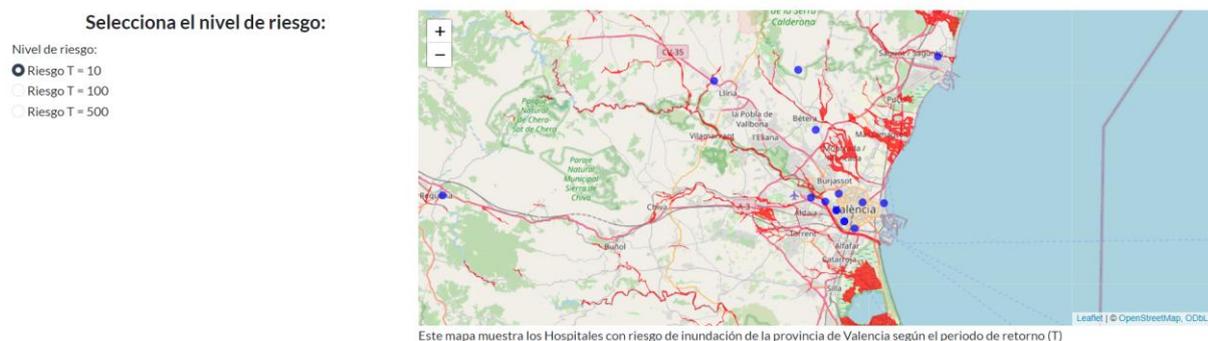
de colegios en riesgo aumenta cuando hay un mayor periodo de retorno, por tanto, en caso de evento extremo se verían afectados una gran cantidad de centros educativos.

HOSPITALES



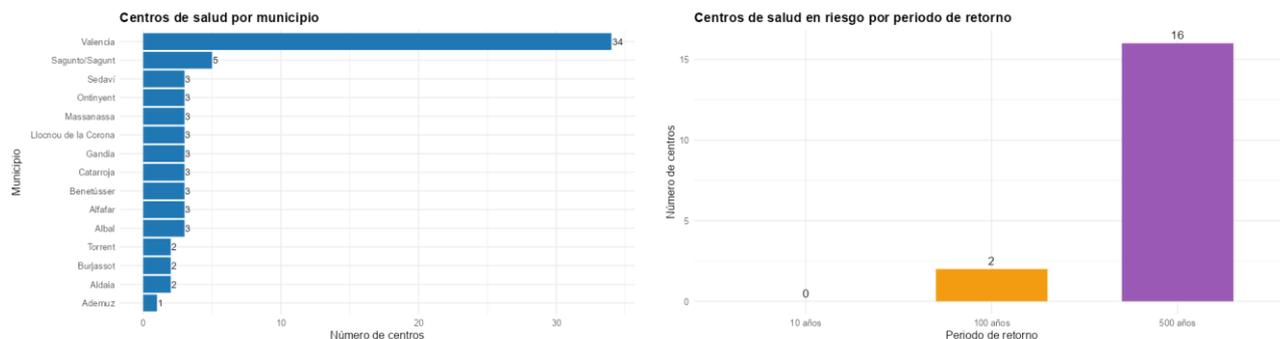
Por un lado, la primera gráfica muestra los municipios con mayor número de hospitales y se puede observar que la que más hospitales tiene es Valencia con 7 hospitales y las demás zonas tienen un hospital.

Por otro lado, hemos observado que no hay ningún hospital que esté expuesto a riesgo de inundación. Esto es debido a que los hospitales suelen construirse en zonas menos vulnerables para garantizar su labor durante situaciones catastróficas.



Este mapa muestra los hospitales (puntos azules) con la zona de riesgo delimitada (polígonos rojos). Haciendo zoom sobre el mapa, podemos comprobar que ningún hospital se encuentra dentro de las zonas de riesgo por inundación. De esta forma, se confirman los resultados obtenidos en la gráfica anterior.

CENTROS DE SALUD



Por un lado, la primera gráfica muestra los municipios con más centros de salud y se puede observar que la que más centros de salud tiene es Valencia con 34 siguiéndoles Sagunto con 5.

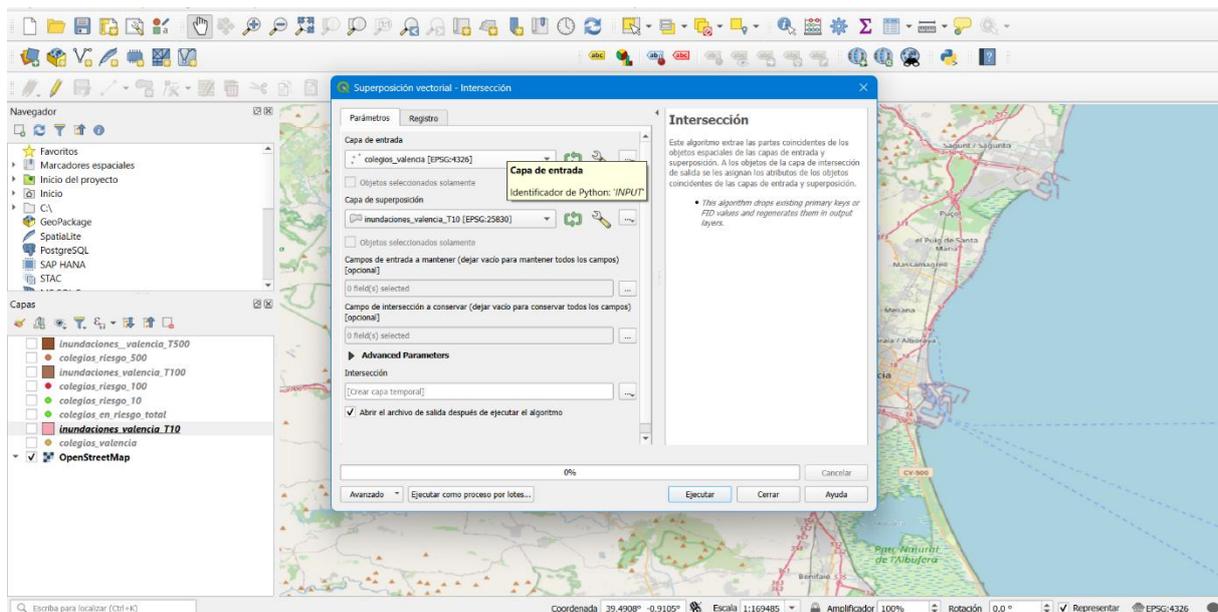
Por otro lado, en el segundo gráfico podemos observar la cantidad de centros de salud que se encuentran en riesgo por periodo de retorno de 10, 100 y 500 años. En T=10 no hay centros en riesgo, pero para T=100 hay 2 y para T=500 hay 16.



Este mapa muestra los centros de salud que tienen riesgo de inundación con un periodo de retorno de 500 años. Si hacemos zoom, podemos ver que algunos de los centros de salud que se encuentran en riesgo están situados por Gandía, Catarroja (zona afectada por la Dana 2024).

4. Discusión

A lo largo del desarrollo del proyecto, nos hemos enfrentado a distintos retos tanto técnicos como de organización. Uno de los aspectos más importantes fue preparar bien los datos espaciales, ya que al trabajar con distintas fuentes y formatos (shapefiles, CSV, GeoJSON...) fue necesario unificar proyecciones, corregir errores de geometría y asegurarnos de que todo encajara correctamente sobre el mapa. Esta parte la realizamos en QGIS y fue clave para que después en R todo funcionara como esperábamos.



También tuvimos que tomar decisiones sobre cómo mostrar la información para que fuera clara y fácil de entender. Por eso organizamos la app por tipos de infraestructura (colegios, hospitales y centros de salud) y añadimos funciones como filtros por municipio, gráficos explicativos, tablas descargables y un apartado de recomendaciones. Pensamos que todo esto hace que la aplicación no solo sea informativa, sino también útil para la prevención.

Uno de los puntos más interesantes fue la parte de análisis espacial. Al hacer intersecciones entre las zonas de riesgo y las ubicaciones de los centros, pudimos detectar con precisión cuáles están realmente en zonas afectadas y clasificarlos según el periodo de retorno. Esto nos permitió generar gráficas comparativas y ofrecer recomendaciones diferentes para cada nivel de riesgo.

Además, creemos que el hecho de haber fusionado en una sola app los tres tipos de centros (educativos, hospitalarios y de salud) le da más valor al trabajo. Así se puede tener una visión más global del problema y tomar decisiones más completas, por ejemplo, a nivel municipal o autonómico.



En definitiva, esta fase del proyecto nos ha servido para consolidar conocimientos sobre el uso combinado de SIG y visualización de datos, y para entender mejor cómo estructurar un cuadro de mando eficaz con datos espaciales. También hemos visto la importancia de adaptar las visualizaciones y funcionalidades al tipo de público al que van dirigidas, ya que no es lo mismo un análisis técnico para expertos que una herramienta pensada para la ciudadanía o para responsables de gestión. Esta reflexión nos ha ayudado a tomar decisiones más orientadas a la utilidad real de la app, más allá del puro desarrollo técnico.



5. Conclusiones

Este mini proyecto nos ha servido para aplicar de forma práctica todo lo que hemos aprendido sobre visualización de datos y análisis espacial. A partir de datos abiertos y oficiales, hemos podido detectar qué colegios, hospitales y centros de salud de la provincia de Valencia están en zonas con riesgo de inundación, diferenciando entre riesgos bajos, medios y altos (según el periodo de retorno).

Gracias al trabajo previo en QGIS, pudimos preparar las capas adecuadamente para después crear una aplicación en R Shiny que muestra esa información de manera clara y útil. La aplicación permite al usuario explorar los datos por tipo de infraestructura, filtrar por municipio, ver mapas interactivos con colores por nivel de riesgo, descargar tablas e incluso leer recomendaciones adaptadas a cada caso.

Uno de los puntos fuertes del trabajo es haber integrado distintos tipos de datos (puntos, polígonos, tablas) y presentarlos de forma ordenada, para que cualquier persona pueda entender fácilmente qué zonas están más expuestas. También creemos que el hecho de añadir gráficos comparativos y recomendaciones personalizadas da un valor añadido a la app y puede ser útil en la toma de decisiones.

Además, el proyecto nos ha hecho ser más conscientes de lo importante que es planificar con antelación y tener sistemas que nos ayuden a prevenir situaciones de emergencia. Si este tipo de herramientas se usaran en la administración o en centros educativos, podrían mejorar mucho la preparación frente a episodios como la Dana de 2024.

En resumen, creemos que hemos cumplido con los objetivos planteados y que el resultado final es un trabajo útil, completo y bien estructurado, que no solo nos ha servido para aprender, sino también para aportar una posible solución a un problema real.



6. Referencias

- **Instituto Cartográfico Valenciano (ICV).** (s.f). Información territorial base y límites provinciales. [Conjunto de datos]. Recuperado de <https://icv.gva.es>
- **Ministerio para la Transición Ecológica y el Reto Demográfico (MITECO).** (s.f). Capas de zonas inundables con periodos de retorno de 10,100 y 500 años.[Conjunto de datos]. Recuperado de <https://www.miteco.gob.es/es/agua/temas/gestion-de-los-riesgos-de-inundacion>
- **Confederación Hidrográfica del Júcar (CHJ).** (s.f). Gestión del riesgo de inundación en la demarcación del Júcar. [Conjunto de datos]. Recuperado de <https://www.chj.es>
- **QGIS (Versión 3.x).** (s.f). Software libre para el análisis espacial: reproyección de capas, corrección de geometrías, intersecciones y generación de nuevas capas. [Software]. Recuperado de <https://qgis.org>
- **R (y entorno RStudio).** (s.f). Lenguaje que usamos para el análisis de datos y creación de gráficos. [Software]. Recuperado de <https://www.r-project.org>
- **Shiny, Leaflet, ggplot2, DT.** (s.f). Librerías en R utilizadas para crear la aplicación interactiva, mostrar mapas, gráficos y tablas. [Software]. Recuperado de
 - Shiny: <https://shiny.posit.co>
 - Leaflet: <https://rstudio.github.io/leaflet/>
 - ggplot2: <https://ggplot2.tidyverse.org>
 - DT: <https://rstudio.github.io/DT>
- **Base de datos de centros educativos, hospitales y centros de salud.** (s.f). Datos georreferenciados mediante fuentes oficiales. [Conjunto de datos]. Recuperado de
 - Centros educativos: <https://ceice.gva.es/es/web/centros-docentes/descarga-base-de-datos>
 - Hospitales: <https://dadesobertes.gva.es/es/dataset/sanidad-sip-hospitales>
 - Centros de Salud Comunidad Valenciana: <https://dadesobertes.gva.es/es/dataset/sanidad-sip-centros-salud>
 - Códigos Postales / municipios: https://github.com/walterleonardo/codigos_postales_espa-a/blob/main/codigos_postales.json