



Islas Canarias  
Del 15 al 19 de noviembre de 2021



## **CARACTERIZACIÓN DE LA POBLACIÓN ANDALUZA MEDIANTE PYTHON Y QGIS**

**Luís Ángel Moya Ruano**

Consejería de Salud y Familias – Junta de Andalucía  
langel.moya@juntadeandalucia.es

**M<sup>a</sup> Encarnación Madrid Verdugo**

Consejería de Salud y Familias – Junta de Andalucía  
encarnacion.madrid@juntadeandalucia.es

**Francisco Javier Rodríguez Rasero**

Consejería de Salud y Familias – Junta de Andalucía  
francisco.rodriguez.rasero@juntadeandalucia.es

**Miguel Antonio Sanz Pérez**

Consejería de Salud y Familias – Junta de Andalucía  
miguel-7755@hotmail.es

### **Introducción**

En la actualidad se está llevando a cabo un gran proceso de digitalización de los trámites administrativos, tanto los realizados por entidades públicas como privadas. Esto permite reducir los tiempos empleados en realizar dichas actividades, permitiendo reducir con ello sus costes asociados, y por otra parte, ofrece la oportunidad de dar un mejor servicio al destinatario de dichos trámites.

Es en esta coyuntura donde se decide implementar un algoritmo en QGIS (sistema de información geográfica) que permita elaborar mapas por personal de la Consejería de Salud de la Junta de Andalucía, tanto con formación en este ámbito, como por personas sin conocimientos en el ámbito de la georreferenciación. Este algoritmo/plugin, permite caracterizar la población andaluza en función de distintos indicadores sociales como la tasa de paro, porcentaje de población infantil o ratios de mortalidad. Esta información se muestra gráficamente en un mapa.

## Objetivo

La Junta de Andalucía diseñó el Decreto 169/2014 mediante el cual establecía que cualquier proyecto perteneciente a sectores como la minería, urbanismo, etc. que se encontrase a menos de 1 km de población tendría que realizar un informe VIS (Valoración del Impacto en Salud). La entidad pública competente, en nuestro caso, la Consejería de Salud y Familias se encargaría por su parte de elaborar la Evaluación de Impacto en Salud (EIS). Para ello, es necesario determinar si existe población afectada, y en caso afirmativo, evaluar las características de dicha población. Es aquí donde surge la necesidad de acelerar el proceso de evaluación del proyecto. Además, la evaluación del impacto en la salud (EIS) de proyectos, planes y urbanismo sobre la población andaluza lleva asociada la necesidad de caracterizar y representar espacialmente indicadores sociales que permitan facilitar la elaboración de informes administrativos.

Por ello, el principal objetivo es comparar indicadores sociales que permitan caracterizar la población respecto a cuatro tipologías geográficas: Andalucía, sus ocho provincias, cualquier municipio o respecto a un buffer en torno al punto de interés. Esto ofrece una gran versatilidad al plugin, ya que permite caracterizar la población afectada por un proyecto con la población de su propio municipio, su provincia o el conjunto de la comunidad autónoma.

A su vez, con este plugin se busca simplificar el uso de datos estadísticos para que puedan ser empleados por usuarios sin conocimientos técnicos. Como se ha comentado anteriormente, este trabajo es desarrollado tanto para personas con conocimientos en herramientas informáticas como QGIS, como para personas con escasos conocimientos en la materia, para las que las tareas de elaboración de mapas pueden conllevar elevados tiempos de trabajo. Con este plugin, este problema queda resuelto al permitir que toda la información necesaria sea introducida a través de una interfaz a forma de formulario, y tras finalizar la ejecución, el mapa resultante con la parametrización puede ser incluido en el informe EIS a través de una imagen en formato png.

Resumiendo, los principales objetivos que tiene el plugin son:

- Caracterizar la población andaluza según indicadores sociales
- Representación georreferenciada de los indicadores sociales
- Facilitar la elaboración de VIS y EIS
- Agilizar el proceso de obtención de información a partir de la localización de un proyecto

## Metodología

Una vez expuesta la motivación y el objetivo de este trabajo, se va a comenzar a explicar en mayor profundidad cómo se consigue llevar a cabo la implementación de este plugin. Para ello, es necesario separar este trabajo en distintos segmentos:

- Obtención de la información
- Tratamiento de datos
- Almacenamiento de datos postprocesados
- Diseño de la interfaz y algoritmia

En primer lugar, y posiblemente el pilar principal de este trabajo sea la obtención de datos. Para poder llevar a cabo un trabajo tan intensivo en datos georreferenciados es necesario tener el soporte de un organismo cuyos datos sean fiables, públicos, y sobre todo, sigan una estructura fija a partir de la cual se pueda obtener toda la información necesaria.

Todos los datos empleados en la elaboración de indicadores sociales se han obtenido a partir de los datos públicos del IECA (Instituto de Estadística y Cartografía de Andalucía). Este organismo provee de una gran gama de indicadores socio-económicos de la región de Andalucía distribuidos espacialmente en mallas de distintos tamaños (250 x 250 metros y 1x1 km). Las capas empleadas son las siguientes (250 x 250 metros):

- Distribución Espacial de la Población en Andalucía
- Razón de mortalidad general
- Distribución Espacial de las Tipologías Constructivas Catastrales en Andalucía
- Caracterización y distribución del espacio construido en Andalucía: Viviendas

Para este trabajo se decidió emplear la malla de 250x250 metros, ya que permite conocer con mayor precisión las características principales de una determinada región de Andalucía, principalmente a pequeña escala. Sin embargo, esto conlleva por otra parte una serie de problemas.

El reducido tamaño del mallado implica que en zonas rurales y despobladas, cada celda de la malla no tenga una suficiente cantidad de personas viviendo ( $< 5$ ), por lo que por motivos de protección de datos y anonimato de los datos, el IECA asigna un valor de -1 al atributo cuya cantidad total sea inferior a 5. Por ejemplo, para evaluar la cantidad de población extranjera en una celda, se distingue entre nacionalidad española, sudamericana, europea, magrebí u otros. En caso de que el total de personas extranjeras

sea inferior a 5, se asigna un valor de -1 como mínimo a dos atributos (por ejemplo nacionalidad sudamericana y europea). Esto provoca que a la hora de analizar esta información y tratarla, estas celdas sean descartadas cuando se esté evaluando el porcentaje de población extranjera (por ejemplo).

Por otra parte, emplear una malla 250x250 metros en un territorio tan grande como es la superficie de la comunidad autónoma de Andalucía provoca que la cantidad de datos e información se vea incrementada, y por lo tanto, las necesidades computacionales. Esto provoca que sea necesario emplear etapas intermedias para tratar los datos obtenidos, reduciendo la cantidad de información a la mínima necesaria, y almacenándola para poder trabajar posteriormente con los datos postprocesados.

Es por esta razón y por varias más, que se hace indispensable realizar un tratamiento de los datos ofrecidos por el IECA. Entre esos motivos destaca que los atributos poblacionales escogidos por el IECA no se corresponden directamente con indicadores sociales, pero a partir de ellos se pueden obtener. Por ejemplo, los datos originales permiten conocer la población total en cada celda y la población masculina en cada celda, por lo que si se quiere comparar entre las distintas celdas cuál es el porcentaje de mujeres, es necesario realizar una serie de cálculos previos.

Además, en esta etapa se realiza un cruzado de datos entre los valores georreferenciados obtenidos a partir de las capas en formato WFS y una tabla con los municipios y la provincia a la que corresponden, ya que esta información es de especial valor a la hora de realizar las comparativas entre regiones, como se verá más adelante. Este cruce de información se debe a la ausencia de un campo ‘provincia’ en las capas del IECA.

La salida de esta etapa es una tabla en Excel donde cada fila es una de las celdas de la malla y las columnas son los indicadores sociales. A continuación se muestra una tabla con los indicadores sociales empleados en este trabajo.

*Tabla 1. Indicadores sociales empleados*

| Nombre               | Descripción   |
|----------------------|---|
| Población total      | Se representa por código de colores la población total de cada malla  |
| Población infantil   | $100 * (\text{Población} < 16 \text{ años} / \text{Población Total})$ → Permite conocer el % de población infantil en cada malla                    |
| Población mayor      | $100 * (\text{Población} > 65 \text{ años} / \text{Población Total})$ → Permite conocer el % de población mayor en cada malla                       |
| Población extranjera | $100 * (\text{Población Total-Española-UE15} / \text{Población Total})$ → Permite conocer el % de población extranjera (excepto UE15) en cada malla |

|   |   |
|---|---|
| Población activa                            | $100 * (\text{Población activa} / \text{Población 15-65 años}) \rightarrow$<br>Permite conocer el % de población activa en cada malla |
| Paro registrado                             | Permite conocer el % de población que se encuentra en paro en cada malla  |
| Mortalidad RMEs                             | Índice RMEs que permite conocer la mortalidad en una malla, una vez corregido el factor poblacional.                                  |
| % Mujer en la población                     | $100 * (\text{Población mujer} / \text{Población Total}) \rightarrow$ % de mujeres respecto del total de la población                 |
| % Suelo viviendas urbanas                   | Permite conocer el porcentaje de viviendas urbanas existente en una celda de la malla   |
| % Suelo industrial                          | Permite conocer el porcentaje de industria existente en una celda de la malla   |
| % Suelo oficinas                            | Permite conocer el porcentaje de suelo destinado a oficinas existente en una celda de la malla  |
| % Suelo comercial                           | Permite conocer el porcentaje de suelo comercial existente en una celda de la malla   |
| % Suelo deportivo                           | Permite conocer el porcentaje de suelo destinado a instalaciones deportivas existente en cada celda de la malla                       |
| % Suelo sanitario, asistencial y beneficios | Permite conocer el porcentaje de suelo destinado a instalaciones sanitarias y asistenciales existente en cada celda de la malla       |
| % Suelo cultural                            | Permite conocer el porcentaje de suelo destinado a instalaciones culturales existente en cada celda de la malla                       |

En siguiente lugar es necesario almacenar toda la información generada en las etapas anteriores, ya que estas etapas iniciales son muy costosas computacionalmente y no es práctico que cada vez que se realice una simulación, o se inicie el plugin, estas etapas tengan que llevarse a cabo. Esto permite ahorrar tiempo a los usuarios del plugin a costa de incrementar el uso de memoria del ordenador, pero al ser almacenados estos valores en Excel en vez de archivos WFS, tanto su acceso por parte de terceras aplicaciones como por el propio plugin es mucho más rápido, siendo a su vez una fuente de información para otros posibles proyectos que usen este archivo Excel como base de datos.

Para ello, se ha diseñado un instalador que cuando se inicia por primera vez el plugin permite recoger todos los indicadores sociales a partir de los datos del IECA y los almacena en un archivo Excel para que en futuras ejecuciones del plugin, estos valores

no tengan que ser recalculados. Las últimas versiones ya instalan automáticamente el archivo con los valores almacenados, pero sigue estando habilitada la opción de cambiar los datos originales en caso de que sea necesario actualizar los datos, iniciándose con ello el instalador del archivo Excel.

Una vez se han visto las etapas relacionadas más puramente con los datos, nos adentramos un poco más en la parte más visible del trabajo desarrollado y la cual permite satisfacer los principales objetivos de este proyecto. Esto no es otra cosa que la interfaz de usuario que ha sido desarrollada en el lenguaje de programación Python, al igual que el resto de plugin. Su principal característica es la capacidad de abstraer al usuario de la complejidad que conlleva realizar mapas elaborados en QGIS a través de un sencillo formulario que puede ser rellenado y ejecutado en menos de 1 minuto, ofreciendo un mapa preparado para su inclusión en un informe EIS en un tiempo mucho más reducido de lo que sería necesario en caso de llevar a cabo ese mismo proceso a mano. Y aunque QGIS ofrece herramientas muy útiles y versátiles para realizar mapas, esa amplia gama de posibilidades también aumenta la complejidad de elaborar estos mapas, que incluso para personas expertas en esta herramienta les llevaría entre 30 minutos y 1 hora, por lo que la justificación de este plugin está más que garantizada. Además, permite estandarizar las gráficas emitidas y reducir al mínimo la formación necesaria para poder elaborar estos mapas.

Con la simplificación y la reducción del tiempo empleado para elaborar estos mapas, ha surgido una nueva oportunidad, ya que el exceso de tiempo puede ser empleado en atender a indicadores sociales secundarios que anteriormente no se podían tener en cuenta por falta de recursos, pero que con la ayuda que ofrece esta herramienta informática, estos análisis en mayor profundidad pueden ser realizados con mayor facilidad.

Como se ha comentado anteriormente, uno de los objetivos es poder comparar una zona de interés (zona donde se va a elaborar un proyecto) respecto los valores del municipio en el que se encuentra, la provincia o la comunidad autónoma. Más concretamente, en la interfaz es necesario introducir las coordenadas en UTM Huso 30 y la región contra la que se quiere comparar:

- Andalucía
- Cualquier provincia (por defecto la provincia a la que pertenece las coordenadas introducidas)
- Cualquier municipio (por defecto el municipio al que pertenece las coordenadas introducidas)
- Buffer/región entorno a las coordenadas introducidas (en metros)

Otro parámetro a introducir es la distancia de visualización del mapa. Esto permite alejar o acercar la vista del mapa para que englobe a una mayor región o por el contrario, se centre en las coordenadas introducidas y las celdas de su entorno.

En siguiente lugar se escoge el indicador social que se desea mostrar en el mapa que se generará al final de la ejecución. Una vez seleccionado el indicador, se podrá seleccionar los colores y los rangos en los que se mostrarán esos colores. Esto se explicará con más detalle en el apartado de resultados.

Por último, se ofrece la opción de mantener las capas generadas en el proyecto de QGIS por si una persona con la formación necesaria para manipular capas en QGIS desea realizar mapas más complejos o elaborados que el plugin diseñado no permite. De esta forma se provee al usuario de una capa con los indicadores sociales que de otra forma no podría obtener, ya que el IECA no provee directamente estos valores. Una modificación reciente fue la posibilidad de emplear la capa CDAU (Callejero Digital de Andalucía Unificado) en vez de la ortofoto de Andalucía como capa de fondo.

Finalmente, existe una pestaña de configuración que permite a los usuarios modificar la dirección desde donde se toman las capas WFS y WMS, permitiendo actualizar estas capas cuando el IECA actualiza dichas capas en su página web. Esto provocará que se inicie el programa de instalación para poder actualizar y almacenar estas nuevas capas. Además, recientemente se ha añadido la posibilidad de cambiar el idioma a inglés, otorgándole una mayor capacidad de difusión al plugin.

## **Resultados**

El resultado que se obtiene de este complemento es una imagen en formato png con los valores mallados del indicador seleccionado en la interfaz del complemento y con un código de colores que categoriza cada celda de la malla.

Este código de colores se basa en la comparación entre el valor de la celda para el indicador social seleccionado, por ejemplo 25% de población extranjera, y el valor de la región con la que se compara (municipio, provincia, Andalucía o buffer). Para ello, por defecto se plantea una distribución de colores en quintiles desde verde hasta rojo, de tal forma que si por ejemplo un 25% de población extranjera se encuentra entre el 80% y el 100% de las celdas con más población extranjera (quinto quintil), el color de fondo asignado a la celda será rojo.

Esto permite de forma sencilla y rápida conocer qué zonas de las representadas en el mapa se encuentran en un quintil u en otro respecto a la referencia que se haya escogido previamente en la interfaz. Aunque por defecto la distribución es en quintiles, esta puede ser modificada por el usuario, cambiando tanto el rango como la cantidad de tramos en los que se quiere subdividir la categorización del mapa.

De igual forma, el color asignado a cada uno de los tramos en los que se subdivide la categorización también puede ser modificado, así como la opacidad de los mismos.

Tras exponer todo lo anterior, la imagen (.png) final que se obtiene con la ejecución del programa tendrá el mapa con toda la parametrización especificada anteriormente por el usuario y una leyenda en la que se reflejará la siguiente información:

- Indicador social
- Coordenadas del punto de interés
- Distancia de visualización (buffer)
- Región contra la que se comparan las celdas de la imagen
- Código de colores y rangos de los mismos (simbología)

## **Conclusión**

La implementación de este plugin ha cumplido con su objetivo inicial de ayudar en las tareas de elaboración de informes EIS, mediante tanto de una reducción del tiempo empleado en elaborar estos mapas como en la posibilidad de realizar análisis más profundos sobre los distintos proyectos a los que es necesario realizarles un informe de Evaluación de Impacto en Salud.

Pero los resultados han sido tan satisfactorios, que se ha promovido y se sigue promoviendo su uso en otros campos de aplicación por su fácil acceso, la casi nula necesidad de conocimientos de QGIS o sistemas de información geográfica en general, y la cantidad de información que es capaz de mostrar con muy poco tiempo de trabajo. Siendo esto ideal también para las propias empresas encargadas de llevar a cabo proyectos para una evaluación inicial de distintos emplazamientos en donde puedan llevar a cabo su actividad económica y poder comprobar los datos de la población afectada sin necesidad de tener que consultar al organismo público correspondiente, reduciendo con ello los costes de elaborar un informe inicial de la zona en cuestión.

Y por último, el uso de QGIS, Python y los datos públicos del IECA para elaborar esta aplicación permite facilitar su distribución y uso por parte de entidades públicas y privadas, sin necesidad de adquirir ninguna licencia. Por otra parte, al tratarse de código abierto, también existe la posibilidad de que este trabajo sea ampliado y personalizado para otras aplicaciones que puedan surgir.

Debido a la buena acogida que ha tenido por distintos organismos públicos a los que se les ha presentado este proyecto, se propone la posibilidad de ampliar este proyecto a otras comunidades autónomas donde el acceso a este tipo de información este restringida a personas con una cierta formación en la materia. Aunque el principal inconveniente es la necesidad de que existan datos bien estructurados y mallados como los ofrecidos por el IECA en Andalucía, ya que como se comentaba al inicio de este documento, uno de los

pilares en los que se basa esta aplicación informática, si no el pilar principal es la fuente de información que ofrece el Instituto de Estadística y Cartografía de Andalucía.

## **Bibliografía**

<https://www.juntadeandalucia.es/institutodeestadisticaycartografia/temas/index-geo.htm>