

# EXTRACCIÓN DE DATOS ESTADÍSTICOS ESPACIALES CON R A PARTIR DE INFORMACIÓN GEOGRÁFICA Y CATASTRAL

RAFAEL SIERRA REQUENA

Dirección General del Catastro

[rafael.sierra@catastro.minhfp.es](mailto:rafael.sierra@catastro.minhfp.es); [rasiere@upvnet.upv.es](mailto:rasiere@upvnet.upv.es)

JOSÉ CARLOS MARTÍNEZLLARIO

Escuela Técnica Superior Ing. en Geodésica, Cartográfica y Topográfica (Universitat Politècnica de València)

[jomarlla@cgf.upv.es](mailto:jomarlla@cgf.upv.es)

ELOÍNA COLL ARRIAGA

Escuela Técnica Superior Ing. en Geodésica, Cartográfica y Topográfica (Universitat Politècnica de València)

[ecoll@cgf.upv.es](mailto:ecoll@cgf.upv.es)

**RESUMEN:** La Dirección General de Catastro (DGC) proporciona la cartografía catastral y datos alfanuméricos catastrales, principalmente mediante servicios web muy orientados a ciudadanos y técnicos a través de la Sede Electrónica de Catastro (SEC - <https://www.sedecatastro.gob.es/>). Aunque también permite el acceso a los usuarios a la descarga de una gran parte de la información catastral.

La información catastral gráfica (geometría de parcelas y construcciones) y alfanumérica (atributos semánticos asociados a fincas y viviendas) es fundamental para la gestión de las Entidades Locales (EE. LL.). Por tanto, la explotación y análisis de esta información es básica para la Administración local (Municipios y Diputaciones Provinciales). Comenzando desde el número de bienes inmuebles hasta su tipología, usos, alturas y otros datos que se utilizan junto con la población por zonas para dimensionar infraestructuras (agua, electricidad, comunicaciones, residuos, equipamientos, etc.).

Con estas premisas, se han desarrollado unos algoritmos con el *software open source* estadístico R y sus paquetes de tratamiento de datos espaciales (R Spatial, dplyr, rgdal, etc), que permiten explotar la información suministrada por la DGC de una forma más cómoda y sencilla, pudiendo orientarse esos análisis a un perfil de técnicos mucho más amplio. Aunque desde el Departamento de Estadística de Catastro ya se utilizan múltiples técnicas y análisis estadísticos con diferentes *software*, e incluso se facilitan datos estadísticos y su uso a los usuarios (<http://www.catastro.minhap.gob.es/esp/estadisticas.asp>).

Concretamente este desarrollo experimental interactúa automáticamente con los ficheros de descarga de información catastral, tanto en formato gráfico o *shapefile*, junto a los datos alfanuméricos que se proporcionan en un formato de texto con extensión CAT. El resultado procesado son tablas gráficas y alfanuméricas de datos agregados estadísticamente y espacialmente, que permiten un análisis rápido de la información por unidad administrativa de ámbito municipal. También se facilita su comprensión confeccionando automáticamente informes en formato texto, gráficos, diagramas y mapas en formato imagen y visualizadores web de información espacial sobre el *software open source* Leaflet.

El objetivo inicial de estos trabajos es proporcionar desarrollos en forma de *scripts* a través de plataformas colaborativas como github. Desde el punto de vista técnico creemos que podría servir como punto de partida para enlazar con otros entornos de desarrollo rural o urbanos como son las ciudades inteligentes (*Smart cities*) e incluso otras tecnologías utilizadas en entornos de desarrollo *open source* (Python, Spark, Hadoop) aplicados a entornos Big Data. De hecho, el desarrollo se centra en la parte conocida como *Extract-Transform-Load* (ETL) dentro del ámbito de minería de datos (*data mining*); también estos métodos dejan la puerta abierta para enlazar con otro tipo de datos estadísticos municipales y ampliar los métodos de análisis a técnicas como *Business Intelligence*.

Finalmente, creemos que los desarrollos informáticos en entornos colaborativos *open source* y su interacción con datos abiertos (*open data*) pueden mejorar el tejido empresarial y de las AA. PP. Y aunque somos conscientes de que existen un gran número de compañías (*geomarketing*, mensajería, servicios, etc.) que ya explotan estos datos catastrales, es posible que otro tipo de técnicos profesionales y de la administración pública vean más fácil su uso.

**PALABRAS CLAVE:** R spatial, datos estadísticos espaciales, información catastral, *data mining*.