

XXI Jornadas de Estadística de las Comunidades Autónomas

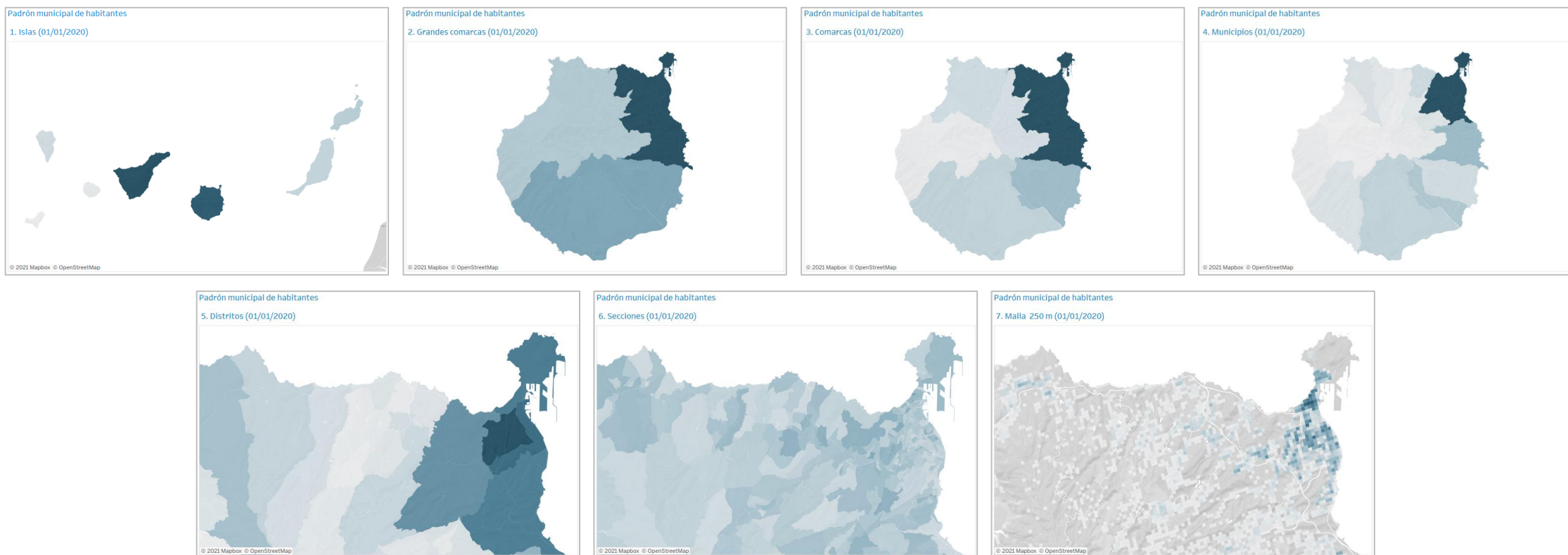
Indicadores demográficos de Canarias en mallas de 250 metros (2010-2020)

Noelia Martín Morales, ISTAC
Rafael Betancor Villalba, ISTAC

Índice:

- Objetivo
- Metodología
- Difusión

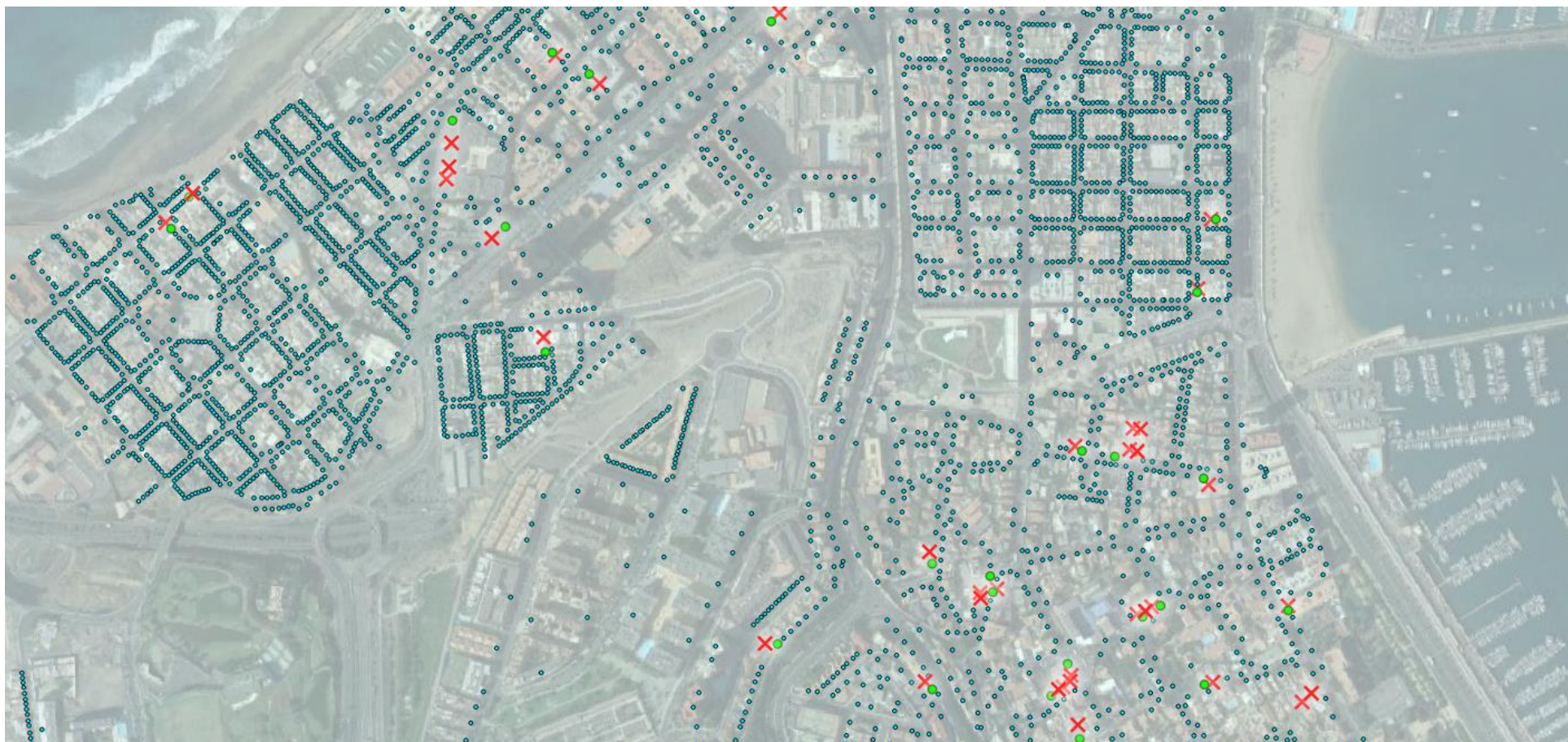
Objetivo: aproximar los datos a la realidad del territorio



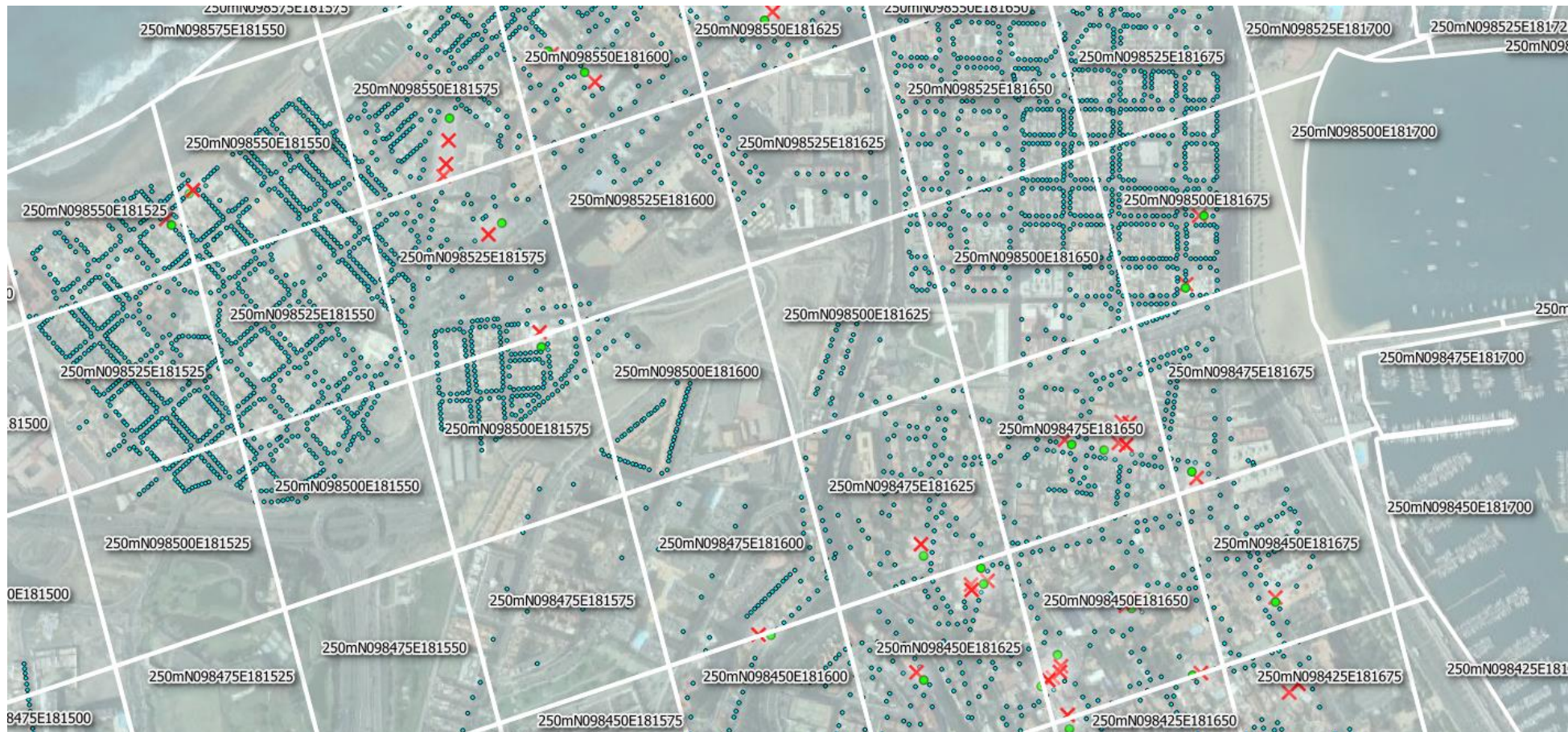
Metodología:

- Georreferenciación y geocodificación
- Identificación de registros
- Secreto estadístico

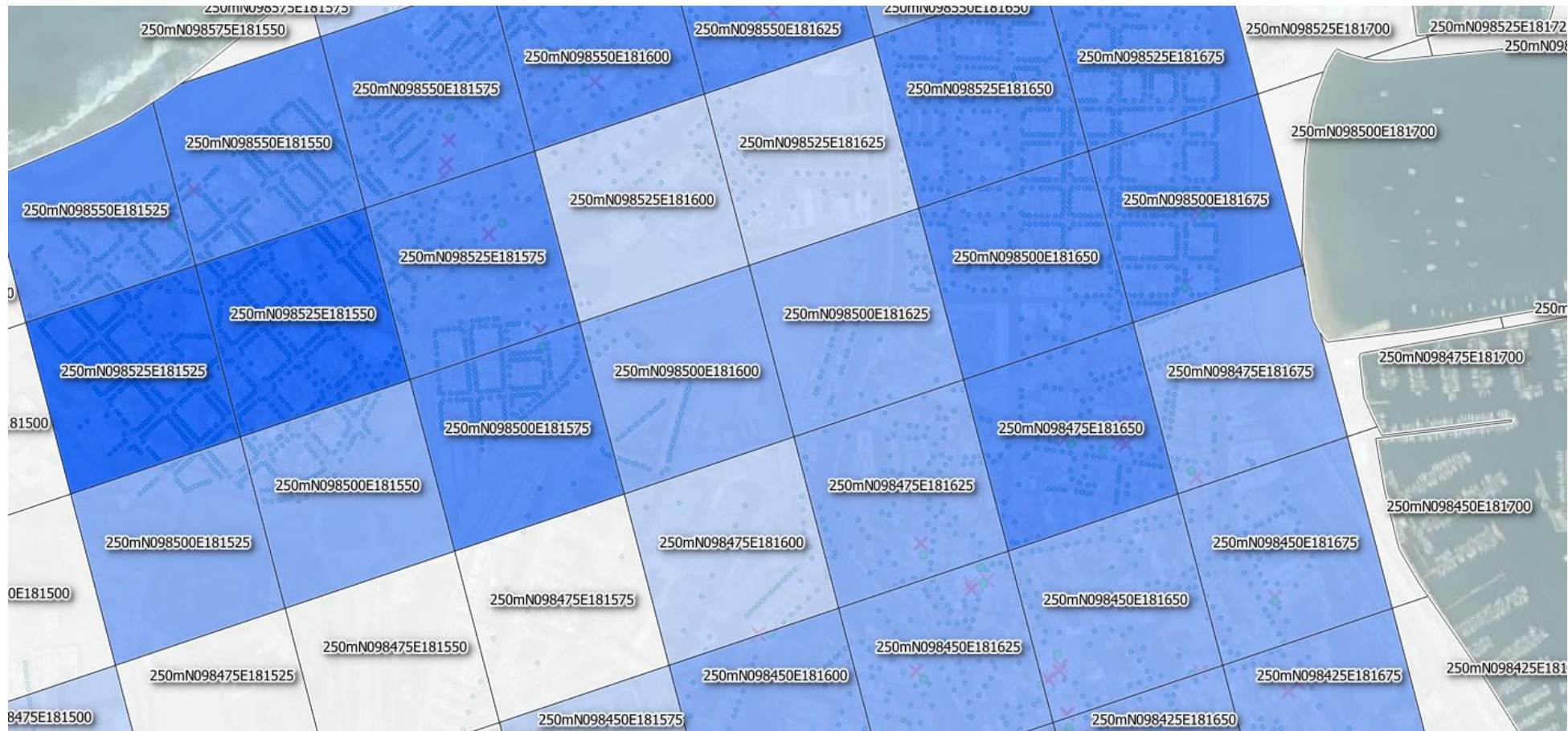
Georreferenciación y geocodificación



Georreferenciación y geocodificación



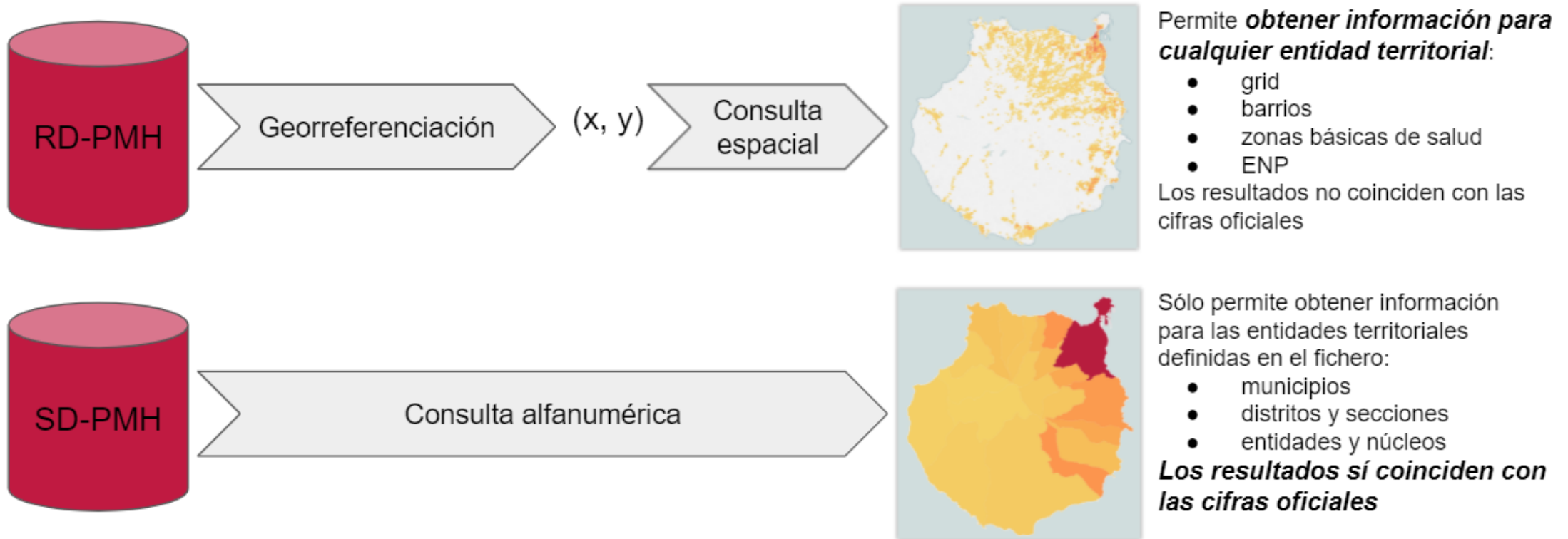
Georreferenciación y geocodificación



Identificación de registros

RD-PMH: Sin anonimizar		SD-PMH: Anonimizado
Con identificación		Sin identificación
No es coherente con las cifras oficiales publicadas		Coherente con las cifras oficiales publicadas
sobran registros		los registros oficiales
errores: entidad, núcleo, sección, etc.		depuración de errores
Dirección completa y normalizada		Sin dirección
permite la georreferenciación		no se puede georreferenciar
población para cualquier territorio: GRID, barrios, ENP...		población sólo para los territorios definidos en el propio fichero: municipios, secciones y núcleos

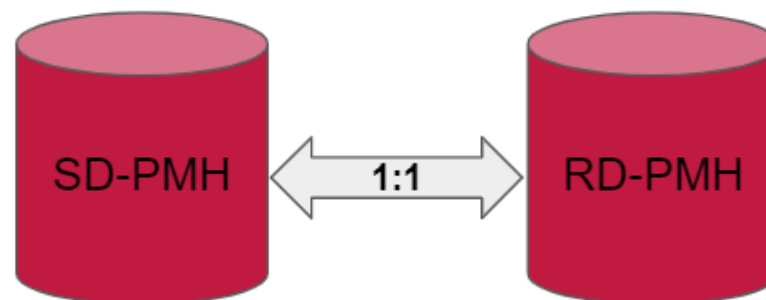
Identificación de registros



OBJETIVO: OBTENER LO MEJOR DE CADA FICHERO

Identificación de registros

Necesitamos trabajar con el fichero SD-PMH porque es el que permite obtener las cifras oficiales. Por lo tanto hay que relacionar cada registro del fichero SD-PMH con el correspondiente registro, ya georreferenciado, del fichero RD-PMH de forma unívoca.



No existe ninguna clave que permita relacionar ambos ficheros.
Sin embargo ambos ficheros comparten las variables:

Variables de enlace entre PMH_SD y PMH_RD													
cpro	cmun	dist	secc	centco	centci	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci

¿Podemos unir por todos estos campos y establecer una relación unívoca?

Parece poco probable que en cada fichero exista un registro que compartiendo todas estas características no se correspondan con la misma persona.

Identificación de registros

Empezamos por los más obvio

Variables de enlace entre PMH_SD y PMH_RD														ENLACE
cpro	cmun	dist	secc	centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	
cpro	cmun	dist	secc	centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	01

Datos 20170101							
ENLACE	Datos de cada enlace				Relación [1 / 1] para cada enlace		
	PMH_SD		PMH_SD Acumulado		PMH_SD		PMH_SD Acumulado
01	2.015.173	95,59%	2.015.173	95,59%	1.986.292	94,22%	1.986.292 94,22%

- Enlace
- Relación: [1 : 1], [n : m]
- Iteración

Los registros que se han conseguido enlazar se quitan de ambos ficheros para los pasos siguientes

Identificación de registros

Seguimos avanzando

Variables de enlace entre PMH_SD y PMH_RD														ENLACE
cpro	cmun	dist	secc	centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	
cpro	cmun	dist	secc	centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	01
cpro	cmun	dist	secc	centco	centsi		sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	02
cpro	cmun	dist	secc	centco			sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	03
cpro	cmun	dist		centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	04
cpro	cmun			centco	centsi	cnucle	sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	05
cpro	cmun	dist		centco	centsi		sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	06
cpro	cmun			centco			sexo	cpronac	cmunnac	anonac	mesnac	dianac	naci	07
cpro	cmun	dist	secc	centco			sexo	cpronac	cmunnac	anonac	mesnac	dianac		08
cpro	cmun	dist	secc	centco			sexo			anonac	mesnac	dianac	naci	09
cpro	cmun			centco			sexo			anonac	mesnac	dianac		10

Identificación de registros

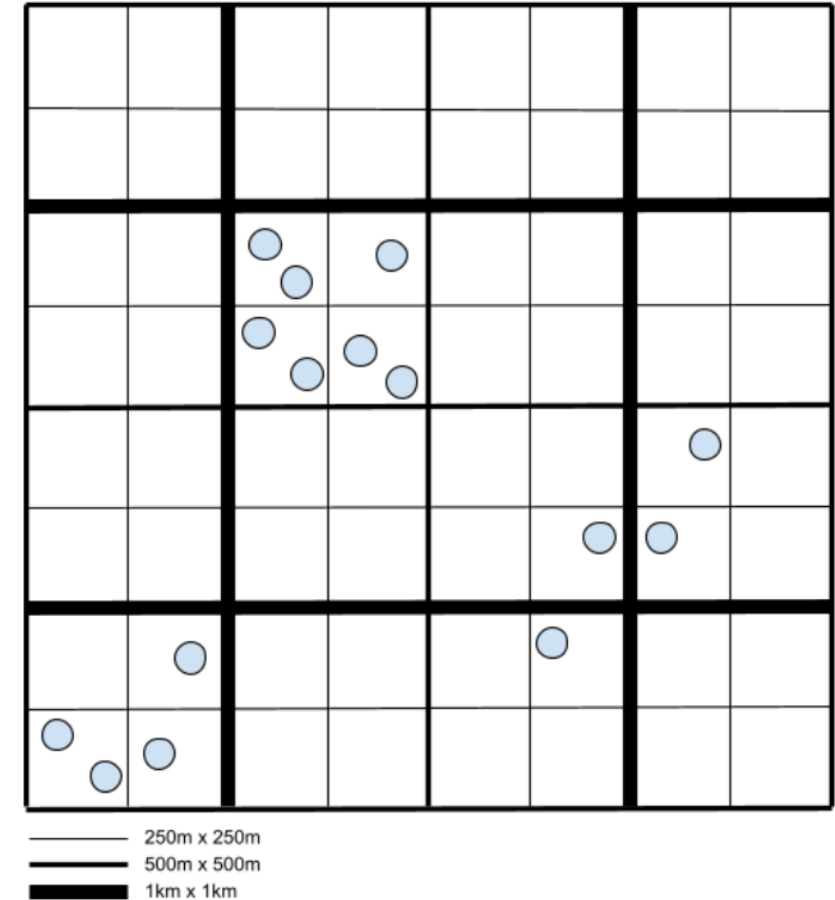
Datos 20170101								
ENLACE	Datos de cada enlace				Relación [1 / 1] para cada enlace			
	PMH SD		PMH SD Acumulado		PMH SD		PMH SD Acumulado	
01	2.015.173	95,59%	2.015.173	95,59%	1.986.292	94,22%	1.986.292	94,22%
02	43.565	2,07%	2.058.738	97,66%	43.422	2,06%	2.029.714	96,28%
03	5.579	0,26%	2.064.317	97,92%	5.525	0,26%	2.035.239	96,54%
04	16.046	0,76%	2.080.363	98,68%	15.895	0,75%	2.051.134	97,30%
05	7.286	0,35%	2.087.649	99,03%	7.150	0,34%	2.058.284	97,64%
06	2.809	0,13%	2.090.458	99,16%	2.798	0,13%	2.061.082	97,77%
07	2.069	0,10%	2.092.527	99,26%	2.031	0,10%	2.063.113	97,87%
08	447	0,02%	2.092.974	99,28%	445	0,02%	2.063.558	97,89%
09	2.970	0,14%	2.095.944	99,42%	2.958	0,14%	2.066.516	98,03%
10	745	0,04%	2.096.689	99,46%	616	0,03%	2.067.132	98,06%
TOTAL	2.108.121							

Secreto estadístico

- Fase 1: Garantizar que ninguna celda tiene menos de 10 residentes (umbral $t1 = 10$).
- Fase 2: Para cada categoría de cada variable de clasificación no se publica la correspondiente cifra si es menor que 3 (umbral $t2 = 3$)

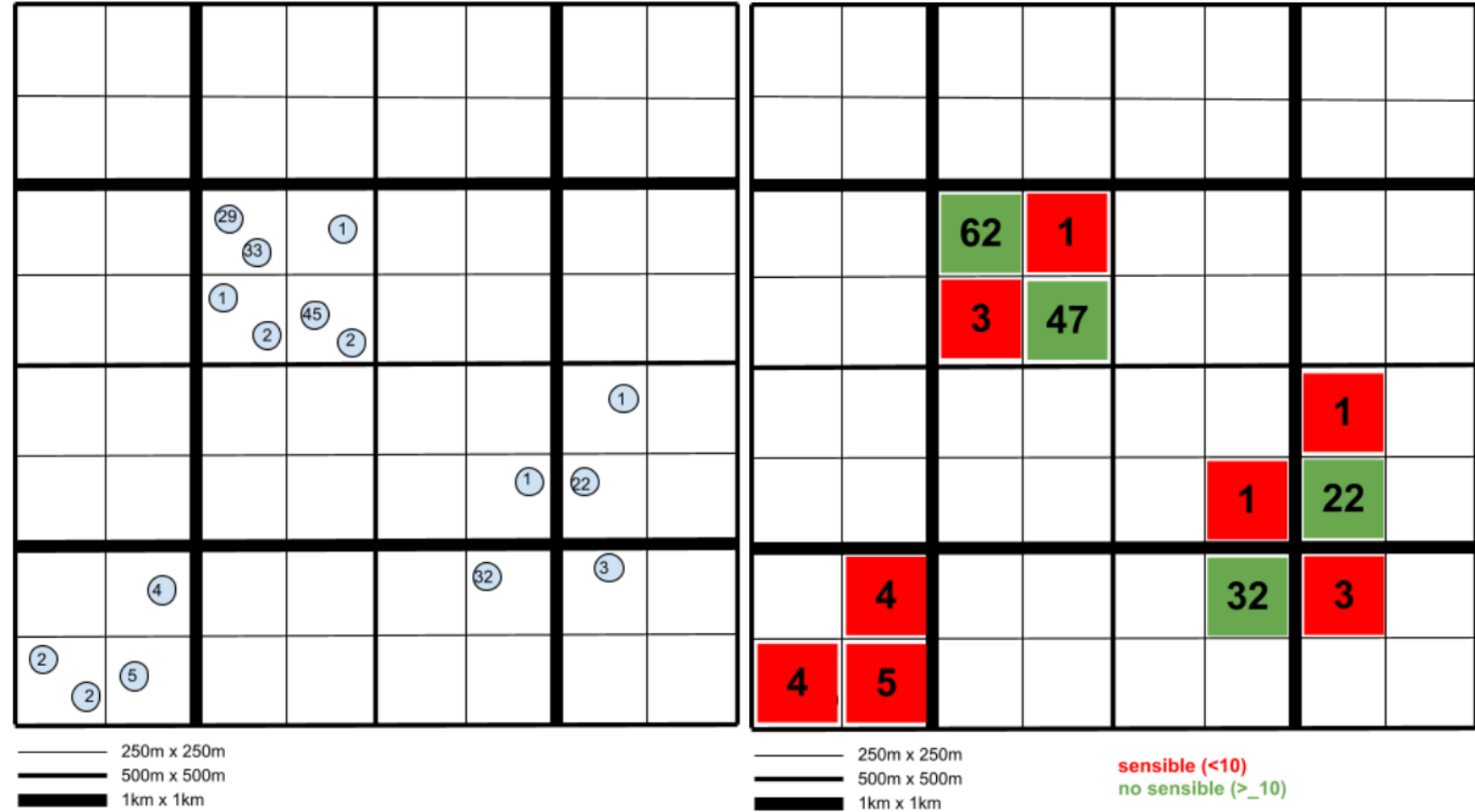
Secreto estadístico. Fase 1

Paso -1 (previo a este trabajo):
georreferenciar



Secreto estadístico. Fase 1

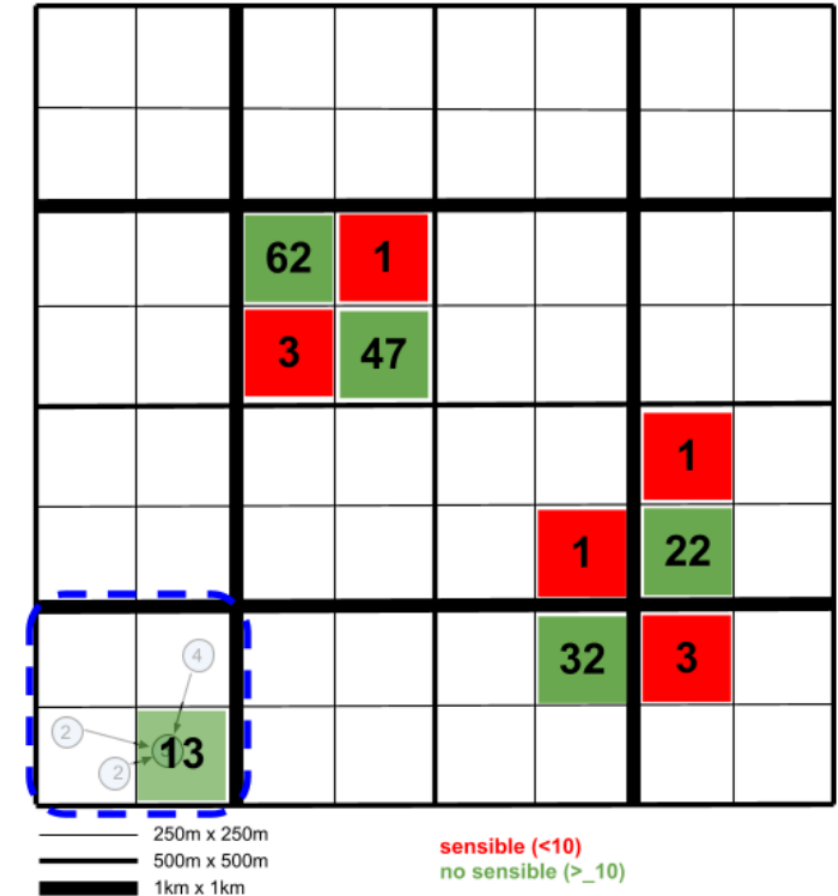
Paso 0 [detectar sensibles]:
calcular población XY y
población GRID



Secreto estadístico. Fase 1

Paso 1 [hermanos menores]:

todos los grid de 250m x 250m hermanos (hijos del mismo grid de 500m x 500m) son sensibles, sin embargo la suma no lo es: reasignar los hermanos a un único grid de 250m x 250m, el de mayor peso (reasignar a nivel de XY, no persona a persona)

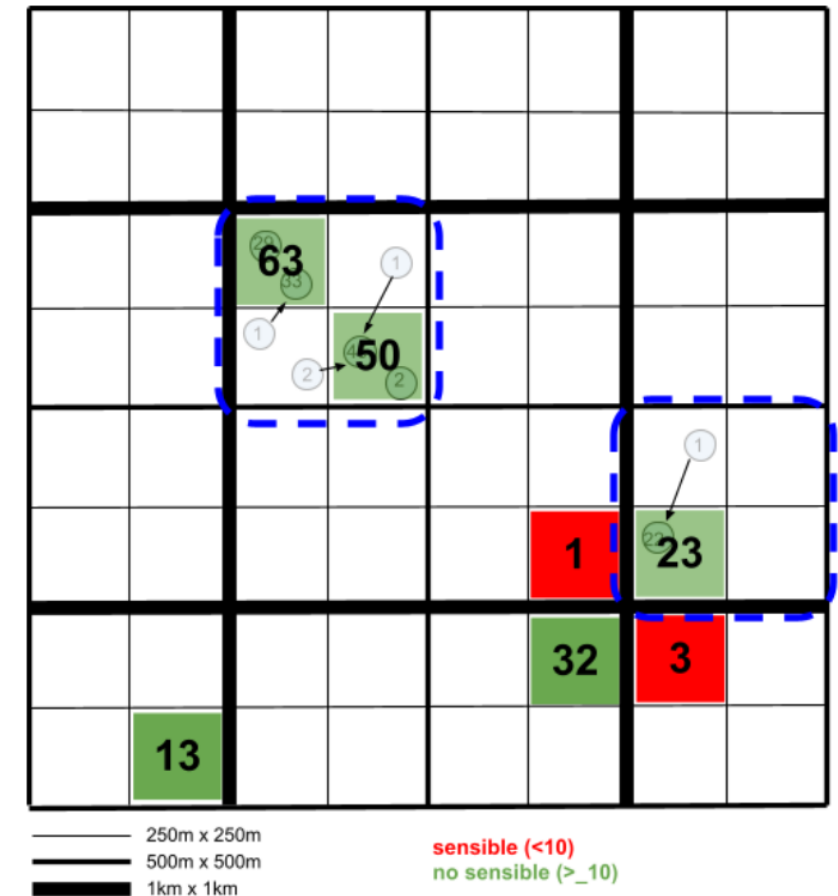


Secreto estadístico. Fase 1

Paso 2 [hermano mayor]:

entre los grid de 250m x 250m hermanos (hijos del mismo grid de 500m x 500m) existen sensibles y no sensibles:

- se vuelve al nivel de XY
- se separan dos conjuntos, uno con los XY de los grid sensibles y otros de los XY' de los grid no sensibles
- para los XY de las celdas sensibles se calcula el XY' más cercano en una celda no sensible hermana
- se le reasigna el grid de 250m x 250m de dicho XY' no sensible
- los grid de 250m x 250m sensibles quedan vacíos
- y se reparten sus pesos entre los hermanos no sensibles según el XY' más cercano

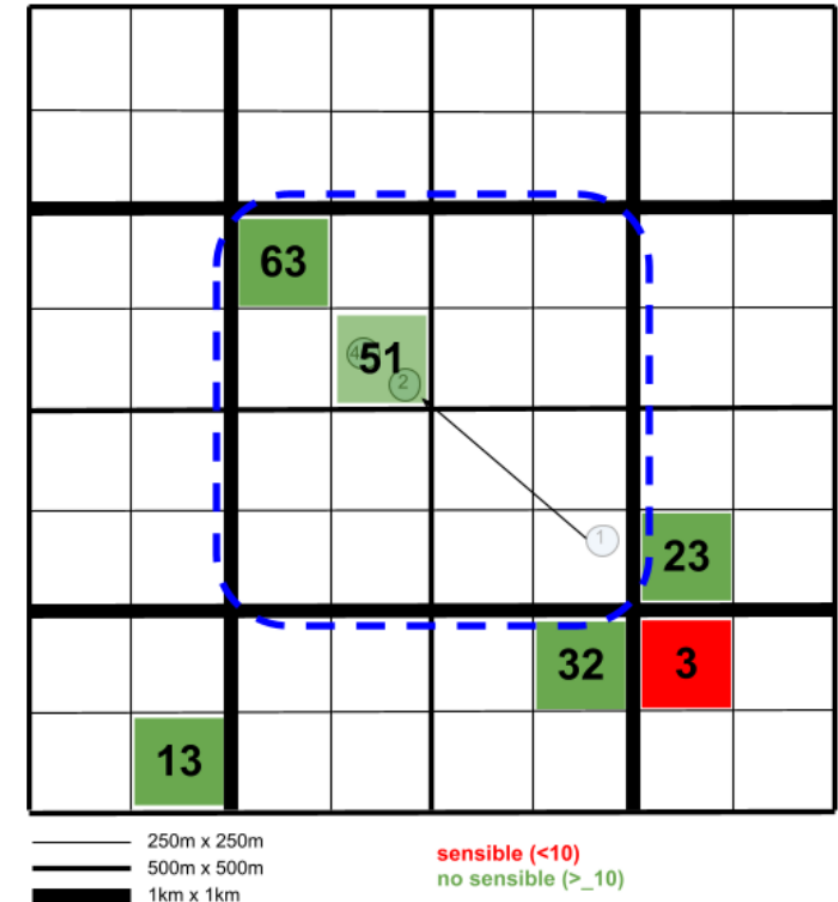


Secreto estadístico. Fase 1

Paso 3 [primo]:

entre los grid de 250m x 250m primos (nietos del mismo grid de 1km x 1km, pero no hermanos, es decir no hijos del mismo grid 500m x 500m) existen sensibles y no sensibles:

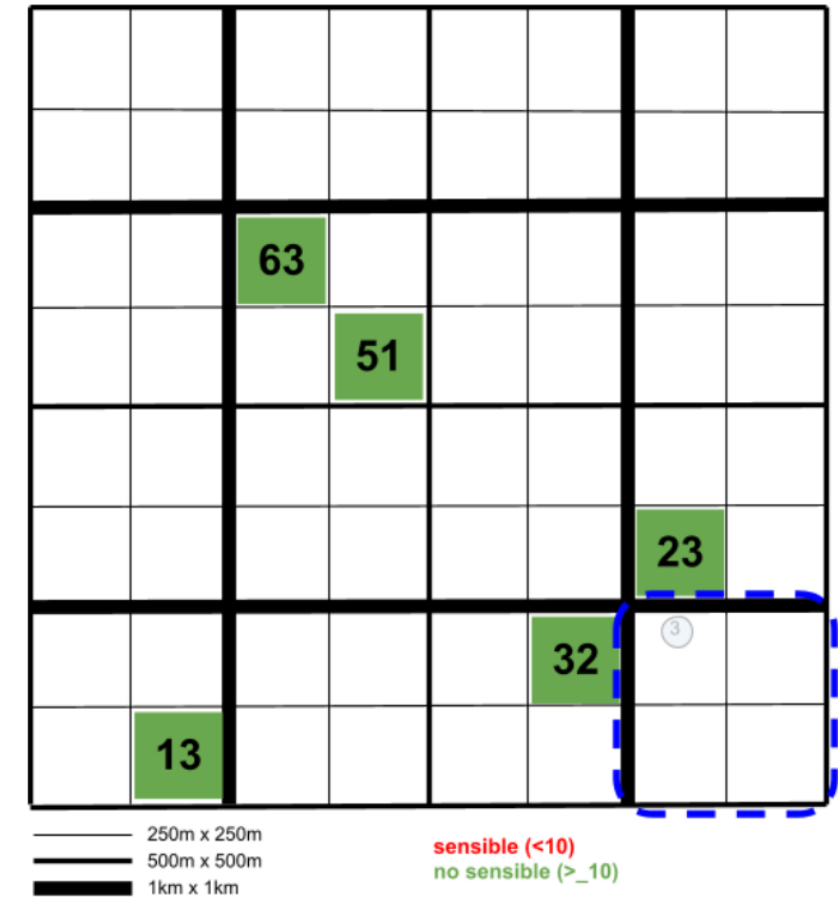
- se vuelve al nivel de XY
- se separan dos conjuntos, uno con los XY de los grid sensibles y otros de los XY' de los grid no sensibles
- para los XY de las celdas sensibles se calcula el XY' más cercano en una celda no sensible prima
- se le reasigna el grid de 250m x 250m de dicho XY' no sensible
- los grid de 250m x 250m sensibles quedan vacíos
- y se reparten sus pesos entre los primos no sensibles según el XY' más cercano



Secreto estadístico. Fase 1

Paso 4 [suprimir sensibles]:

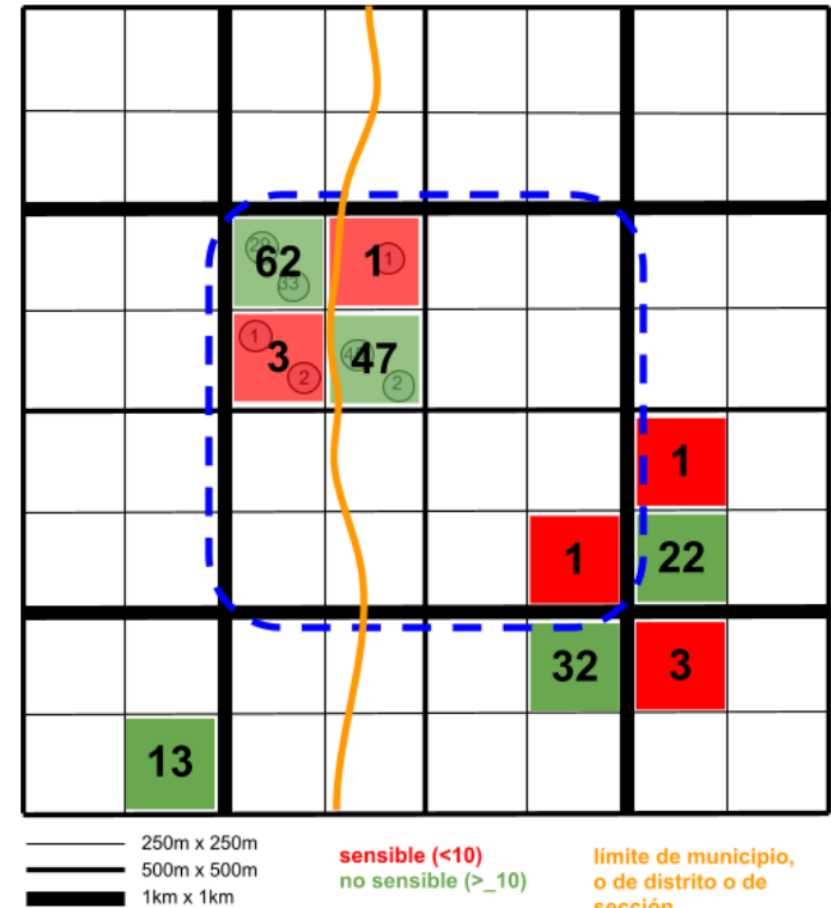
Se suprimen las celdas sensibles que no han sido procesadas anteriormente, perdiéndose mucha menos información, que si se hubiera aplicado la supresión de todas las celdas sensibles desde el inicio, sin aplicar esta metodología.



Secreto estadístico. Fase 1

Posibles mejoras:

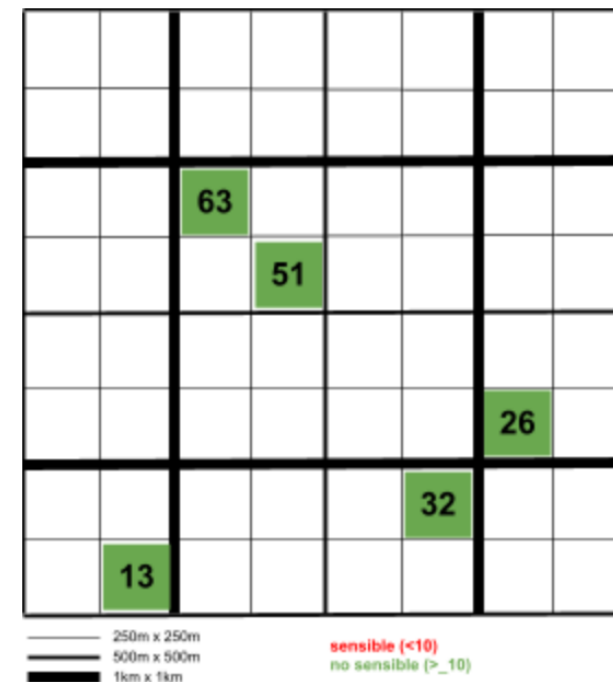
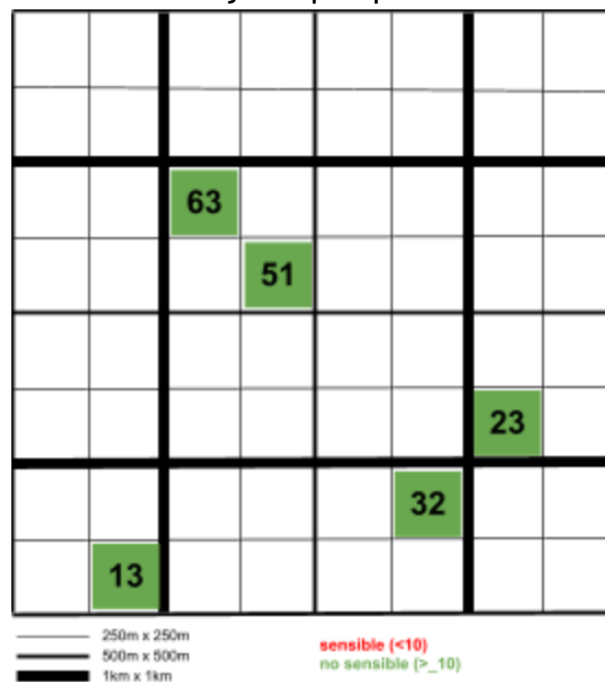
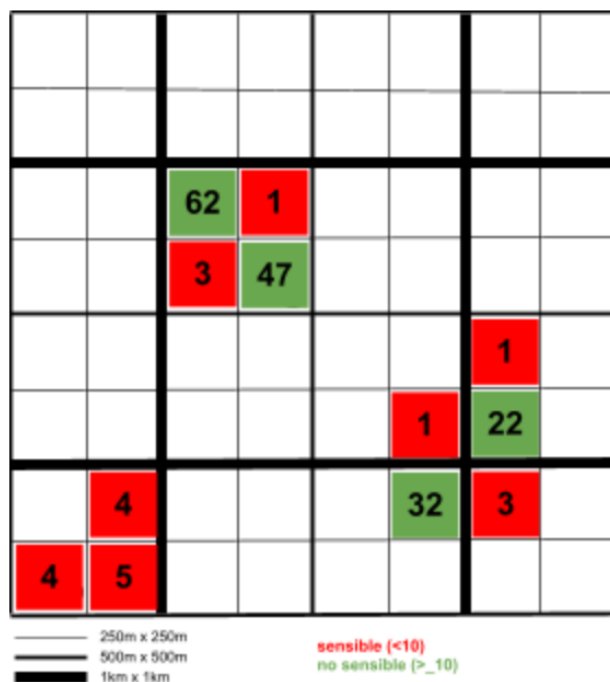
- añadir restricciones como obligar a que sean hermanos o primos del mismo municipio o incluso del mismo distrito o sección



Secreto estadístico. Fase 1

Comparando sin y con las mejoras:

- Inicialmente hay 185
- Si suprimimos las sensibles se pierden 22
- Con el método propuesto se reasignan 19 y se pierden 3
- Con la mejora propuesta se han recuperado esos 3



Secreto estadístico. Fase 2

Secreto estadístico para las variables demográficas:

Una vez garantizado que no existen celdas de 250m con menos de 10 personas pasamos a una segunda fase 2, preservar el secreto para las variables estadísticas sexo, edad, nacionalidad y lugar de nacimiento:

S.E. fase 1	S.E. fase 2			S.E. fase 1	
POBLACION	POBLACION_HOMBRES	POBLACION_MUJERES	POBLACION_INDICE_FEMINIDAD	POBLACION_HOMBRES	POBLACION_MUJERES
10				2	8
13	13	0	0	13	0
11				9	2

S.E. fase 1	S.E. fase 2			S.E. fase 1		
POBLACION	POBLACION_00A14	POBLACION_15A64	POBLACION_65MAS	POBLACION_00A14	POBLACION_15A64	POBLACION_65MAS
11	0	11	0	0	11	0
14		12		2	12	0
30		19		9	19	2

Variables de clasificación e indicadores publicados

Variables:

Las variables del Padrón Municipal de Habitantes que se han incluido en este análisis como variables de clasificación para la fase 2 del secreto estadístico son:

- Sexo: hombre, mujer
- Grupos de edad: menos de 15 años, de 15 a 64 años, de 65 y más años
- Nacionalidades: españoles, extranjeros
- Lugar de nacimiento: Canarias, resto de España, extranjero

Indicadores:

- Población
- Edad media
- Hombres
- Mujeres
- Índice de feminidad
- Población nacionalidad española
- Población nacionalidad extranjera
- Población extranjera (%)
- Población de 0 a 14 años
- Población de 15 a 64 años
- Población de 65 o más años
- Población de 0 a 14 años (%)
- Población de 15 a 64 años (%)
- Población de 65 o más años (%)
- Índice de dependencia
- Índice de dependencia mayores
- Índice de dependencia menores
- Índice de vejez
- Índice de juventud
- Densidad de población
- Población nacionalidad extranjera por km²
- Población nacida en Canarias
- Población nacida en el resto de España
- Población nacida en el extranjero
- Población nacida en Canarias (%)
- Población nacida en el resto de España (%)
- Población nacida en el extranjero (%)

Difusión:

- Catálogo de datos abiertos del ISTAC
- Atlas estadístico de Canarias

Catálogo de datos abiertos del ISTAC

The screenshot displays the ISTAC Open Data Catalogue interface. The header includes the ISTAC logo and the text 'INSTITUTO CANARIO DE ESTADÍSTICA'. A navigation bar at the top right contains a 'MÁS APLICACIONES' button. The main content area is titled 'Catálogo de datos abiertos' and features a sidebar with 'Conjuntos de datos' (12) and social media links for Twitter and Facebook. The main panel shows the dataset 'Indicadores demográficos para malla 250 m' with a large image of a young girl blowing bubbles. Below the image, there is a section for 'Cartografía y principales indicadores demográficos' with tabs for 'Cartografía', 'INSPIRE - Unidades', and 'Indicadores Demográficos'. A list of 11 resources is provided, detailing demographic indicators for the 250m grid for each year from 2010 to 2020.

istac INSTITUTO CANARIO DE ESTADÍSTICA

MÁS APLICACIONES

Catálogo de datos abiertos

Datos Colecciones

Conjuntos de datos

12

Social

Twitter

Facebook

Indicadores demográficos para malla 250 m

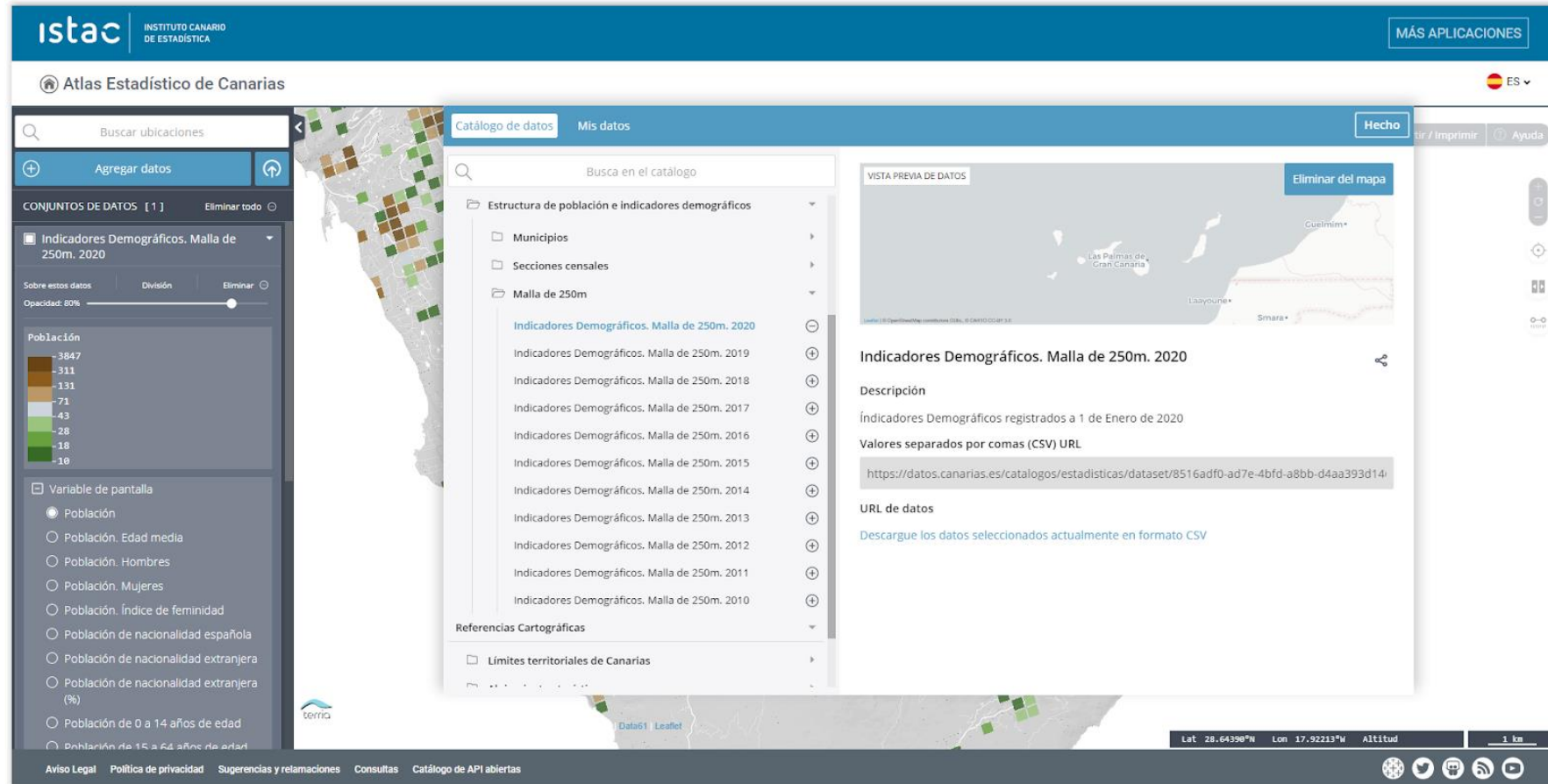
Cartografía y principales indicadores demográficos.

Cartografía INSPIRE - Unidades Indicadores Demográficos

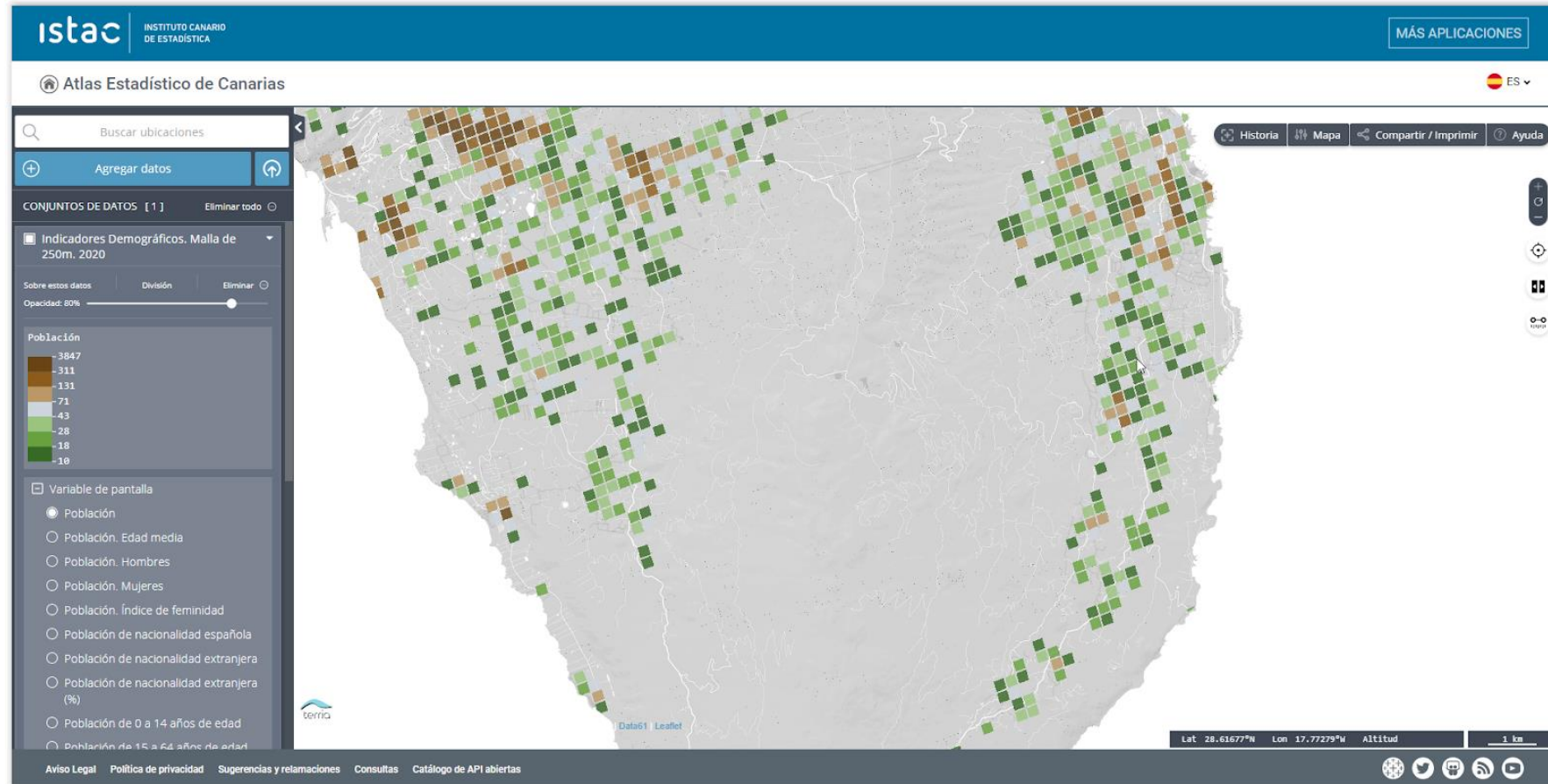
Recursos o servicios vinculados entre sí

1. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2010.
2. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2011.
3. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2012.
4. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2013.
5. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2014.
6. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2015.
7. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2016.
8. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2017.
9. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2018.
10. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2019.
11. Indicadores demográficos para la malla estadística de Canarias 250 m de lado. Año 2020.

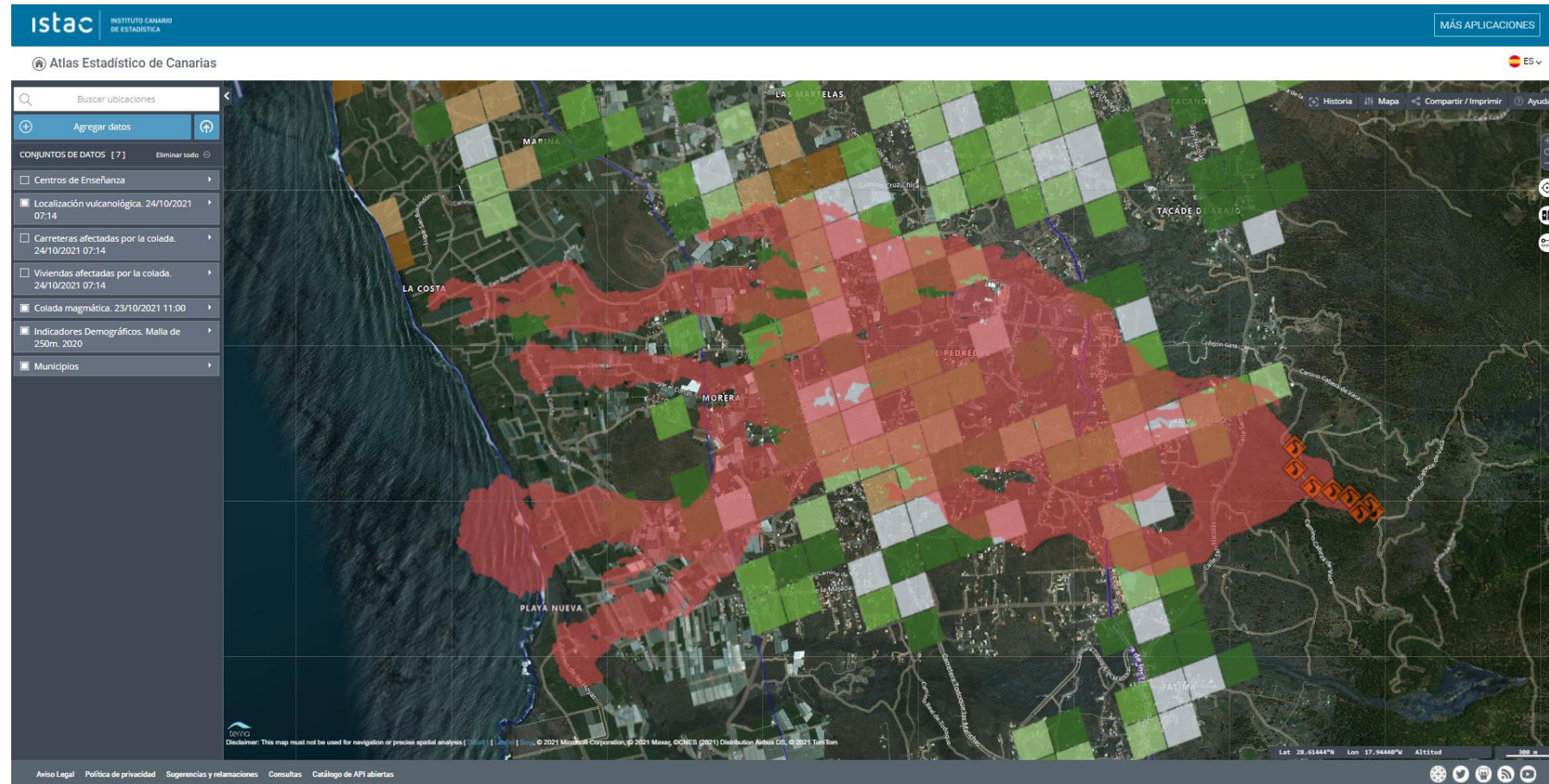
Atlas estadístico de Canarias



Atlas estadístico de Canarias



Atlas estadístico de Canarias



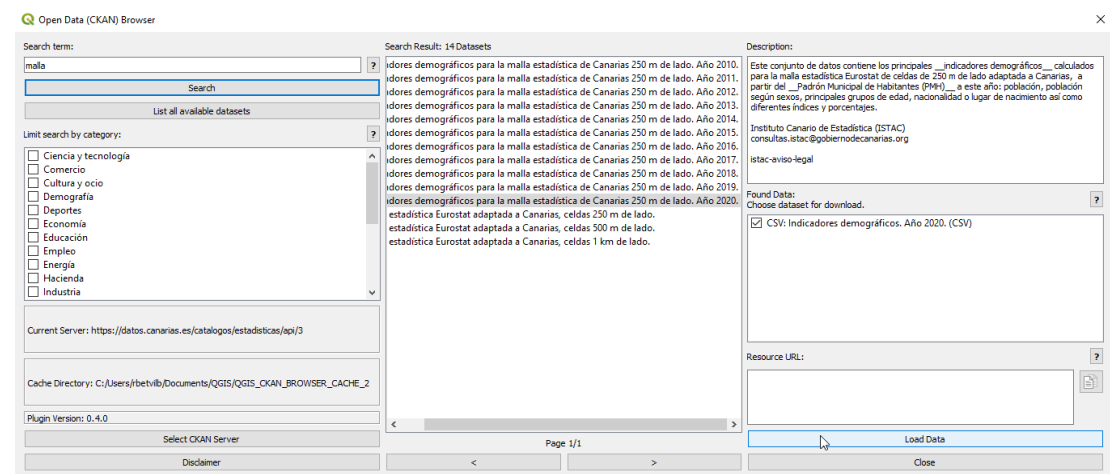
Reutilización:

- Desde QGIS, con el plugin [Open Data \(CKAN\) Browser](#)

QGIS: con el plugin Open Data (CKAN) Browser

Desde QGIS, con el plugin [Open Data \(CKAN\) Browser](#) es posible conectar con Open Data Gobierno de Canarias. ISTAC, usando la url:

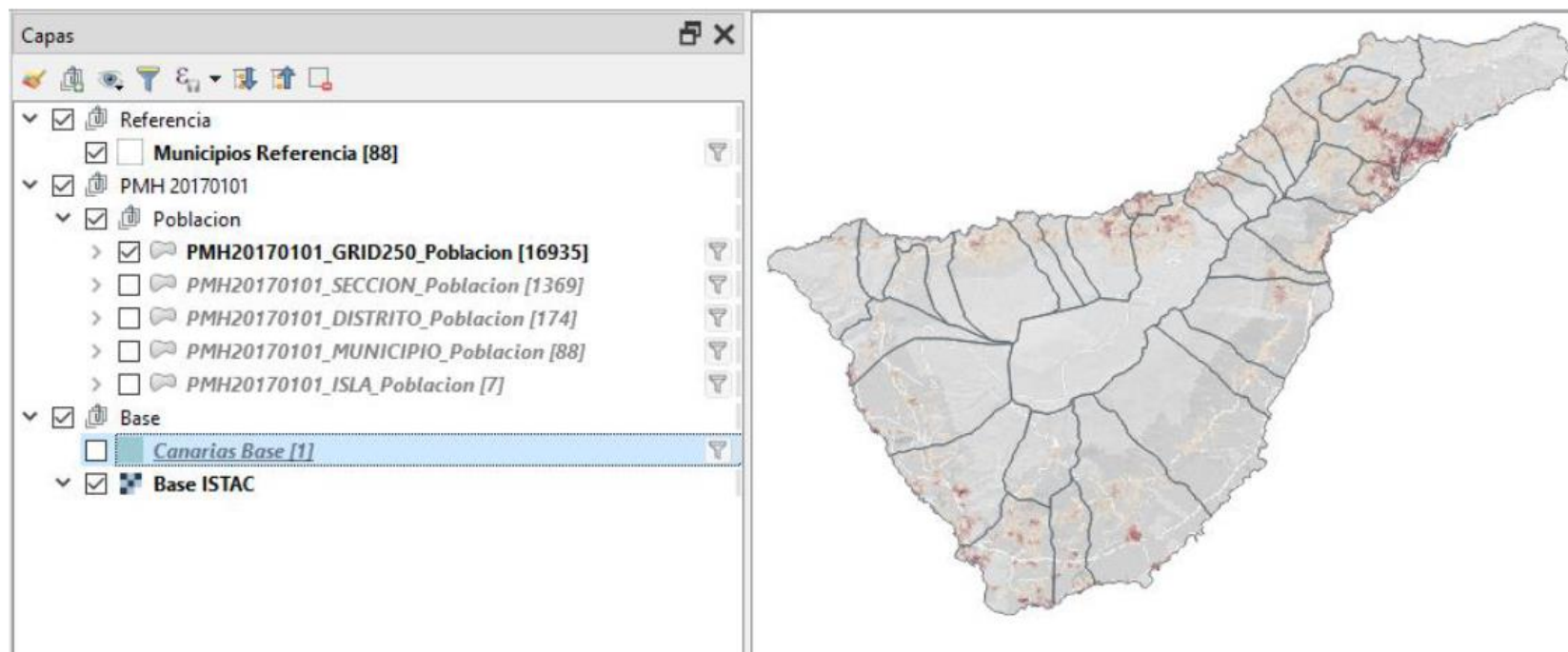
<https://datos.canarias.es/catalogos/estadisticas/api/3>



QGIS: con el plugin Open Data (CKAN) Browser

Desde QGIS, con el plugin [Open Data \(CKAN\) Browser](https://datos.canarias.es/catalogos/estadisticas/api/3) es posible conectar con Open Data Gobierno de Canarias. ISTAC, usando la url:

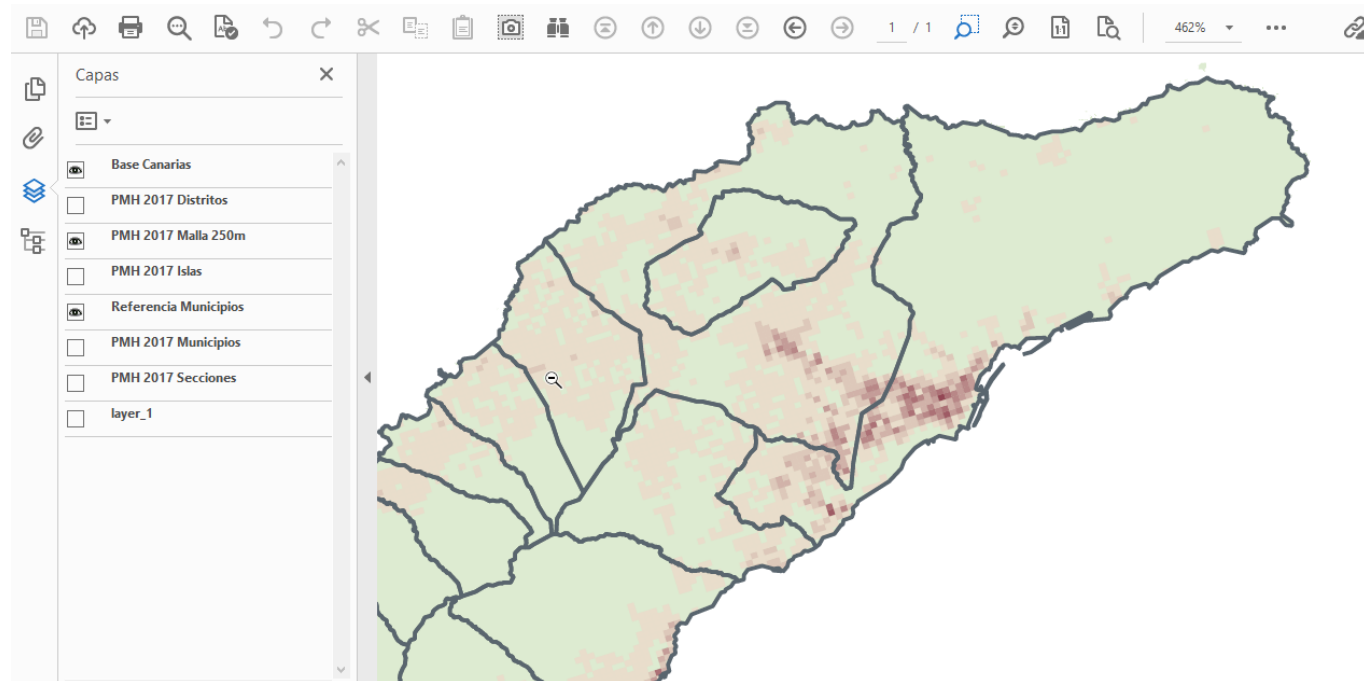
<https://datos.canarias.es/catalogos/estadisticas/api/3>



QGIS: con el plugin Open Data (CKAN) Browser

Desde QGIS, con el plugin [Open Data \(CKAN\) Browser](#) es posible conectar con Open Data Gobierno de Canarias. ISTAC, usando la url:

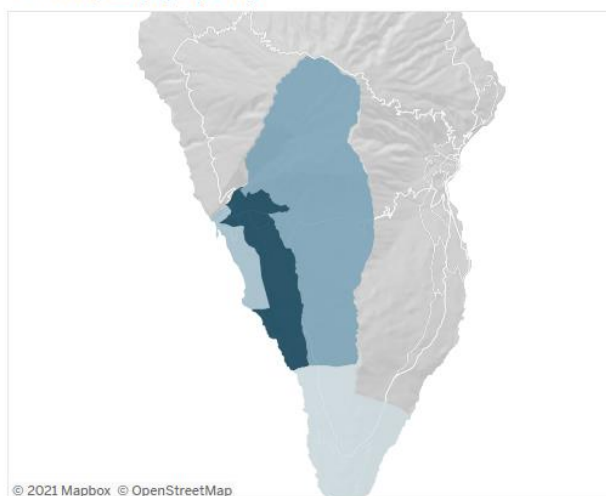
<https://datos.canarias.es/catalogos/estadisticas/api/3>



Otros ejemplos:

Padrón municipal de habitantes

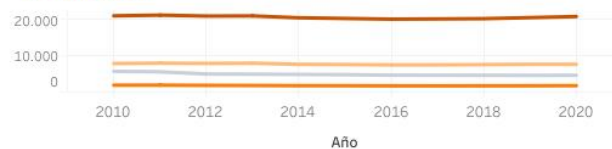
4. Municipios (01/01/2020)



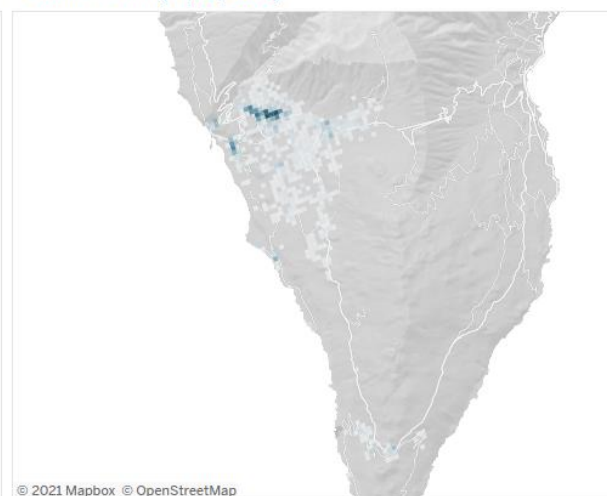
Leyenda



001 Población



7. Malla 250 m (01/01/2020)



Leyenda



Municipios

- El Paso
- Fuencaliente de La Palma
- Los Llanos de Aridane
- Tazacorte

Mapa

- ☐ 1. Islas
- ☐ 2. Grandes comarcas
- ☐ 3. Comarcas
- ☒ 4. Municipios
- ☐ 5. Distritos
- ☐ 6. Secciones
- ☐ 7. Malla 250 m

Mapa derecha

- ☐ 1. Islas
- ☐ 2. Grandes comarcas
- ☐ 3. Comarcas
- ☐ 4. Municipios
- ☐ 5. Distritos
- ☐ 6. Secciones
- ☒ 7. Malla 250 m

Zoom islas

La Palma

Zoom comarcas

La Palma - Valle de Aridane

Indicadores de población

001 Población

Año

< 2020 >



☒ Mostrar historial

+ a b | e a u

Otros ejemplos:

Datos anuales del fichero estadístico de afiliaciones según malla de 250x250 metros

Medias anuales de los cuatro trimestres

Mapa de: Afiliaciones. Año 2020



XXI Jornadas de Estadística de las Comunidades Autónomas

Indicadores demográficos de Canarias en mallas de 250 metros (2010-2020)

Noelia Martín Morales, ISTAC
Rafael Betancor Villalba, ISTAC

iii GRACIAS !!!