A population grid for Andalusia
Year 2013

Institute of Statistics and Cartography of Andalusia (IECA)

Sofia (BU), 24th October 2013

IECA project. Population grid cells sized 250 x 250m for Andalusia

1. Highlights & main information sources
2. Georeferentiation of the population
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

IECA project. Population grid cells sized 250 x 250m for Andalusia

1. **Highlights & main information sources**
2. Georeferentiation of the population
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

1. Highlights & main information sources

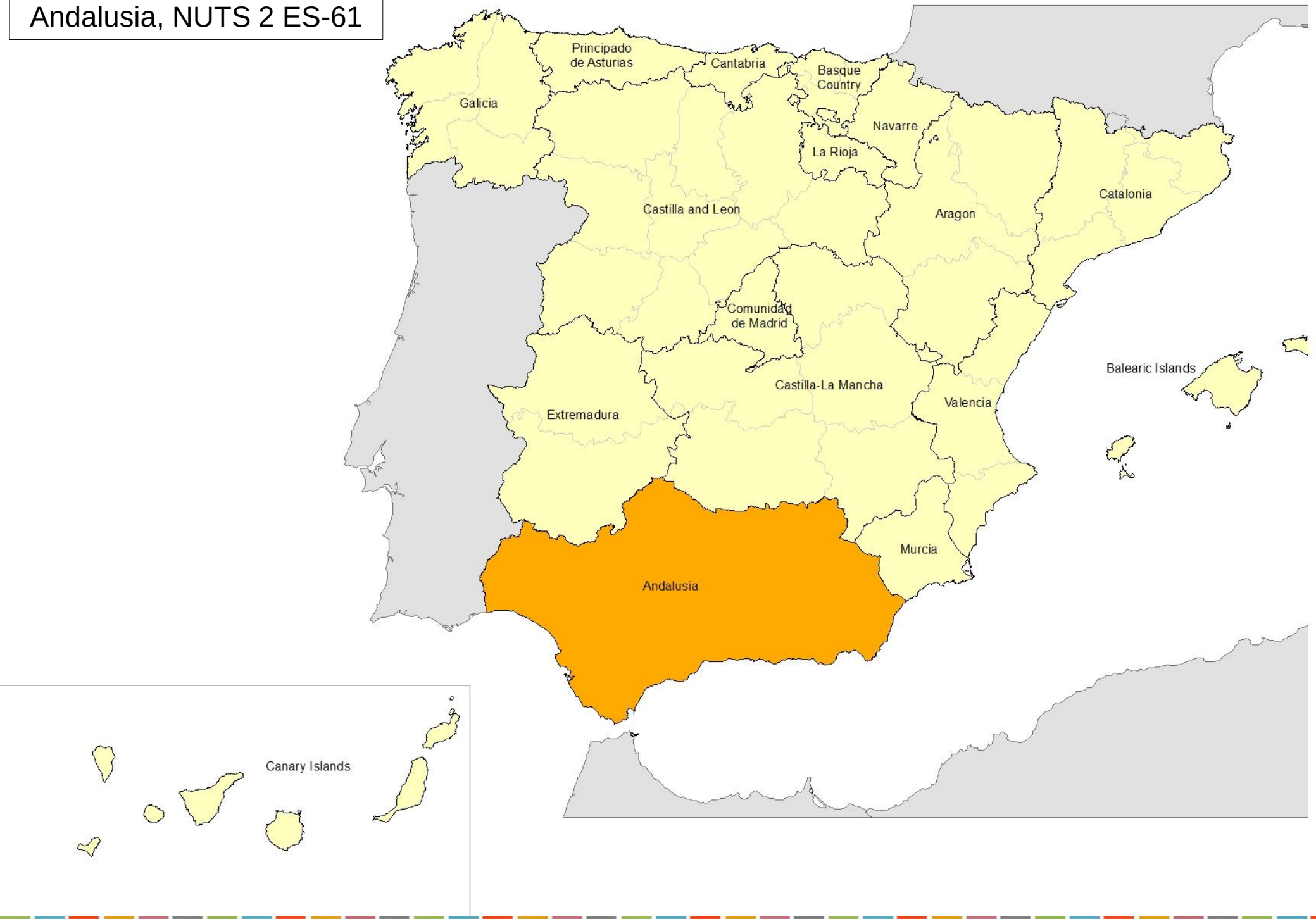
- The Institute of Statistics and Cartography of Andalusia (IECA), official body member of the Regional Government of Andalusia, has developed a project to allocate the population of Andalusia
- A layer of cells has been defined using the Eurostat Grid Generation Tool for ArcGIS. We ran this add-on in the Spatial Analysis module of ArcGIS 10. Our region has been divided generating a grid of cells sized 250 x 250 m
- We have used as main source of information for population, administrative data coming from Population Register of Andalusia (RPA), dated January 1st, 2013

1. Highlights & main information sources

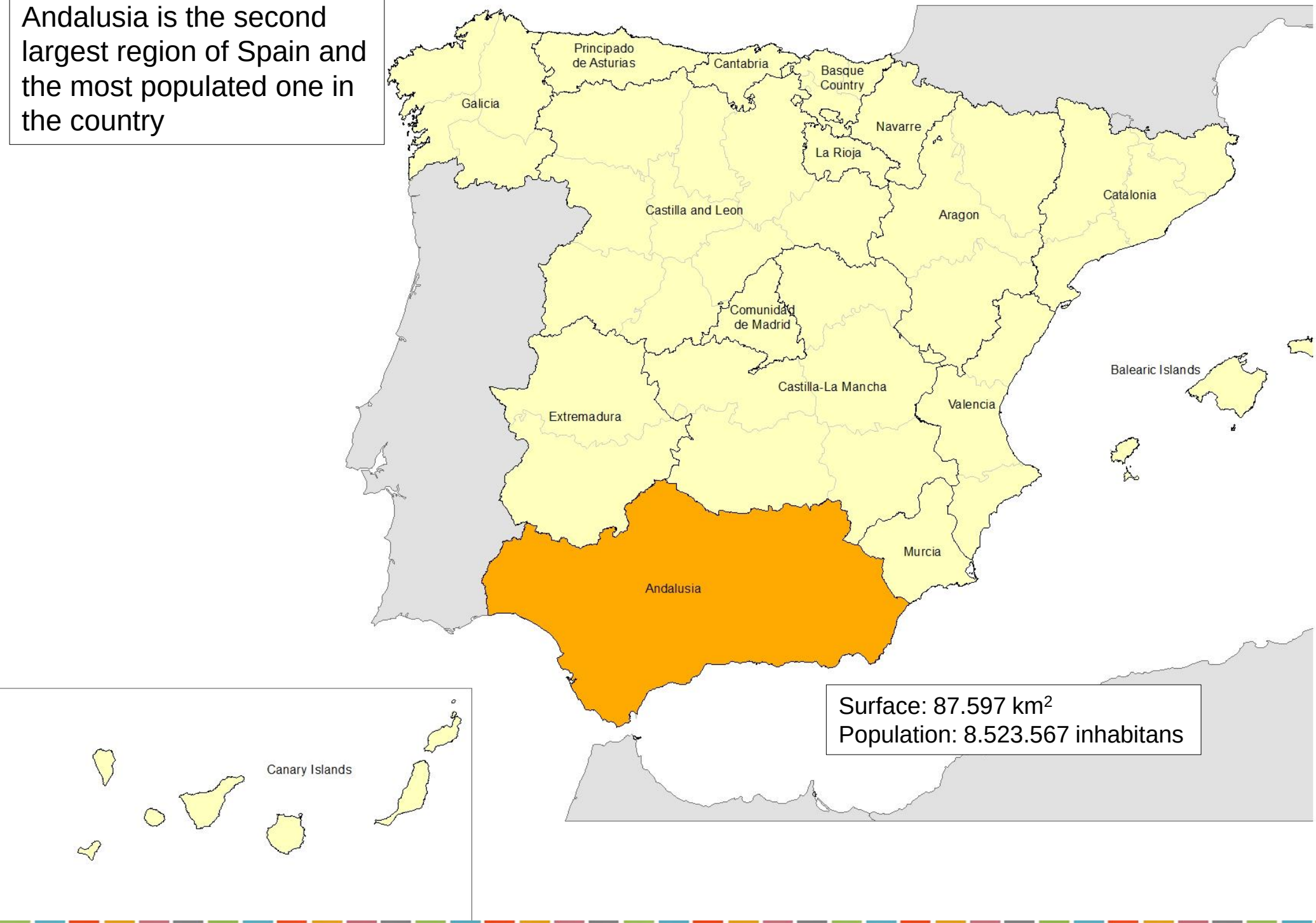
- **Population Register of Andalusia (RPA)**: It accounts for all households and individuals settled in Andalusia. It provides information about postal address and demographic characteristics such as sex, age and nationality of household members. This database is created and maintained by IECA and its data come from administrative files
- **Street & Building Directory of Andalusia**: Directory created and maintained by IECA in cooperation with town and city halls, municipalities and province governments. It contains info on street and building entrances coming from National Statistical Institute (Spain) files, cadastre files, etc.
- **Urban Cadastre**: Administrative files provided by Cadastre Unit from the Ministry of Finance and Public Administration (Spain) containing cadastre information, related to residential area, building characteristics, road characteristics, etc.

NUTS 2 in Spain

Andalusia, NUTS 2 ES-61



Andalusia is the second largest region of Spain and the most populated one in the country



Surface: 87.597 km^2
Population: 8.523.567 inhabitans

IECA project. Population grid cells sized 250 x 250m for Andalusia

1. Highlights & main information sources
2. **Georeferentiation of the population**
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

IECA project. Population grid cells sized 250 x 250m for Andalusia

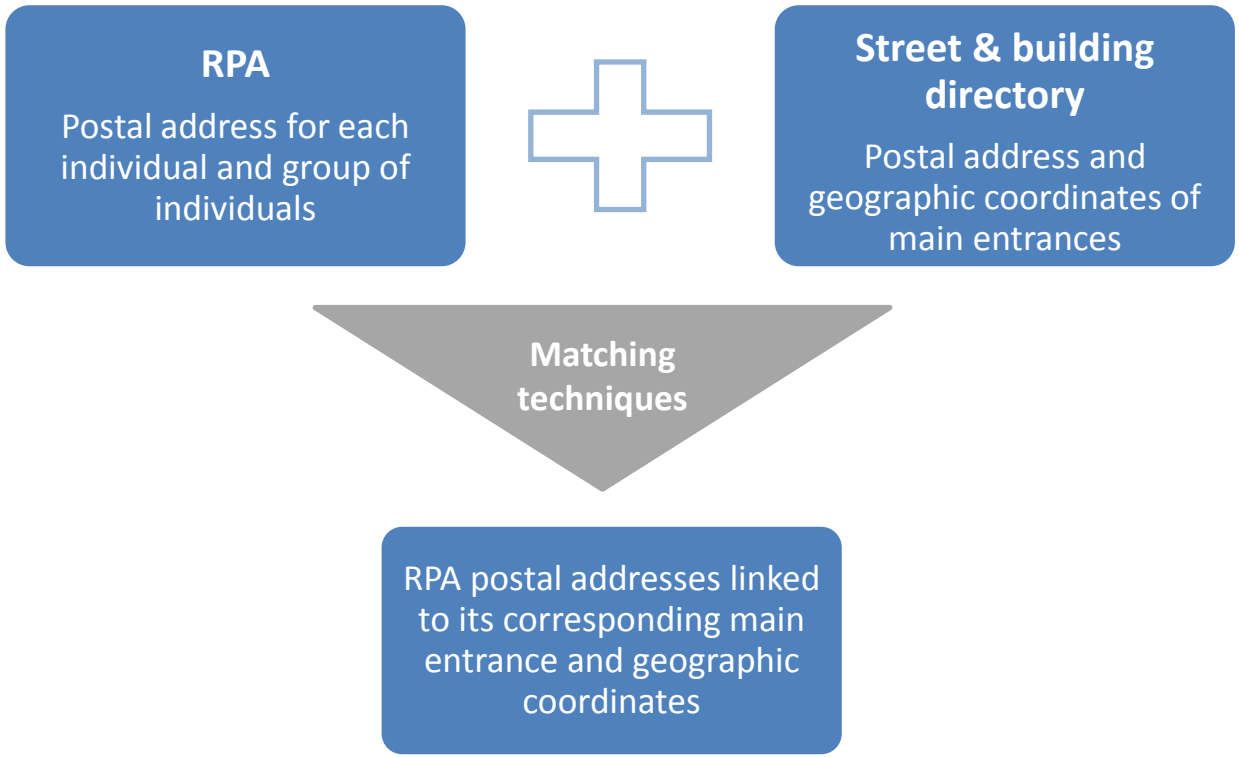
1. Highlights & main information sources
2. **Georeferentiation of the population**
   - **2.1. Bottom-up approach: the geocoding of buildings main entrances**
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

2.1. Bottom-up approach: the geocoding of buildings main entrances

Georeferentiation of population living in buildings is possible through GIS once coordinates are assigned.

The linking process informs about the degree of coincidence (reliability of matching)

A layer of points is generated representing buildings main entrances and associated number of inhabitants and their demographic characteristics



Population Register of Andalusia

Street & building directory

ORTAL	RPA_CPMR	RPA_DSEC	RPA_TVIA	RPA_NVIAC	RPA_NUMER	POB_TOTAL	RPA_TVIA	RPA_NVIAC	RPA_NUMER	PORTALERO	X	Y
702978	21001	1001	ALDEA	COLLADO EL	6	1	ALDEA	COLLADO EL	6	S	176854,84	4198111,19
702979	21001	1001	ALDEA	COLLADO EL	12	13	ALDEA	COLLADO EL	12	S	176880,68	4198077,28
702980	21001	1001	ALDEA	COLLADO EL	13	1	ALDEA	COLLADO EL	13	S	176895,93	4198070,13
702981	21001	1001	ALDEA	COLLADO EL	24	1	ALDEA	COLLADO EL	24	S	176878,73	4198097,9
702982	21001	1001	ALDEA	COLLADO EL	27	15	ALDEA	COLLADO EL	27	S	176852,66	4198125,27
702983	21001	1001	AVDA	EL CACHON	1	7				0	0	0
702984	21001	1001	AVDA	EL CACHON	16	4				0	0	0
702985	21001	1001	AVDA	EL CACHON	20	1				0	0	0
702986	21001	1001	CALLE	ALEGRIA	6	3	CALLE	ALEGRIA	6	S	177673,78	4198294,6
702987	21001	1001	CALLE	ALEGRIA	8	4	CALLE	ALEGRIA	8	S	177684,37	4198300,6
702988	21001	1001	CALLE	ALEGRIA	10	4	CALLE	ALEGRIA	10	S	177699,25	4198309,01
702989	21001	1001	CALLE	ALTA	1	5	CALLE	ALTA	1	S	177698,74	4198436,31
702990	21001	1001	CALLE	ALTA	2	3	CALLE	ALTA	2	S	177691,64	4198424,56
702991	21001	1001	CALLE	ALTA	7	1	CALLE	ALTA	7	S	177710,26	4198446,86
702992	21001	1001	CALLE	ALTA	8	4	CALLE	ALTA	8	S	177714,45	4198434,21
702993	21001	1001	CALLE	ALTA	15	1	CALLE	ALTA	15	S	177724,42	4198447,09
702994	21001	1001	CALLE	ANDALUCIA	3	2	CALLE	ANDALUCIA	3	S	177696,55	4198284,2
702995	21001	1001	CALLE	ANDALUCIA	5	3	CALLE	ANDALUCIA	5	S	177683,38	4198272,82
702996	21001	1001	CALLE	ANDALUCIA	7	3	CALLE	ANDALUCIA			9,17	4198269,18
702997	21001	1001	CALLE	ANDALUCIA	9	3	CALLE	ANDALUCIA			5,85	4198257,68
702998	21001	1001	CALLE	ANIMAS	6	2	CALLE	ANIMAS			6,71	4198342,37
702999	21001	1001	CALLE	ANIMAS	8	2	CALLE	ANIMAS			0,12	4198342,32
703000	21001	1001	CALLE	ANIMAS	10	2	CALLE	ANIMAS			7,18	4198342,23
703001	21001	1001	CALLE	ARIAS MONTANO	5	2	CALLE	ARIAS MONTANO	5	S	178022,11	4198486,61
703002	21001	1001	CALLE	ARIAS MONTANO	7	1	CALLE	ARIAS MONTANO	7	S	178022,78	4198492,58
703003	21001	1001	CALLE	ARIAS MONTANO	9	4	CALLE	ARIAS MONTANO	9	S	178023,56	4198499,43
703004	21001	1001	CALLE	ARIAS MONTANO	11	1	CALLE	ARIAS MONTANO	11	S	178025,08	4198514,63
703005	21001	1001	CALLE	ARIAS MONTANO	13	1	CALLE	ARIAS MONTANO	13	S	178025,32	4198527,16
703006	21001	1001	CALLE	ARIAS MONTANO	15	2	CALLE	ARIAS MONTANO	15	S	178025,31	4198539,83
703007	21001	1001	CALLE	ARIAS MONTANO	18	2	CALLE	ARIAS MONTANO	18	S	178030,8	4198590,02
703008	21001	1001	CALLE	ARIAS MONTANO	20	1	CALLE	ARIAS MONTANO	20	S	178031,73	4198576,48
703009	21001	1001	CALLE	ARIAS MONTANO	22	3	CALLE	ARIAS MONTANO	22	S	178029,48	4198607,65
703010	21001	1001	CALLE	ARIAS MONTANO	24	2	CALLE	ARIAS MONTANO	24	S	178027,67	4198625,53
703011	21001	1001	CALLE	ARIAS MONTANO	26	3	CALLE	ARIAS MONTANO	26	S	178030,29	4198597,5
703012	21001	1001	CALLE	ARIAS MONTANO	28	2	CALLE	ARIAS MONTANO	28	S	178029,88	4198603,42
703013	21001	1001	CALLE	ARIAS MONTANO	30	1	CALLE	ARIAS MONTANO	30	S	178016,63	4198700,94
703014	21001	1001	CALLE	ARIAS MONTANO	34	3	CALLE	ARIAS MONTANO	34	S	178001,32	4198789,37

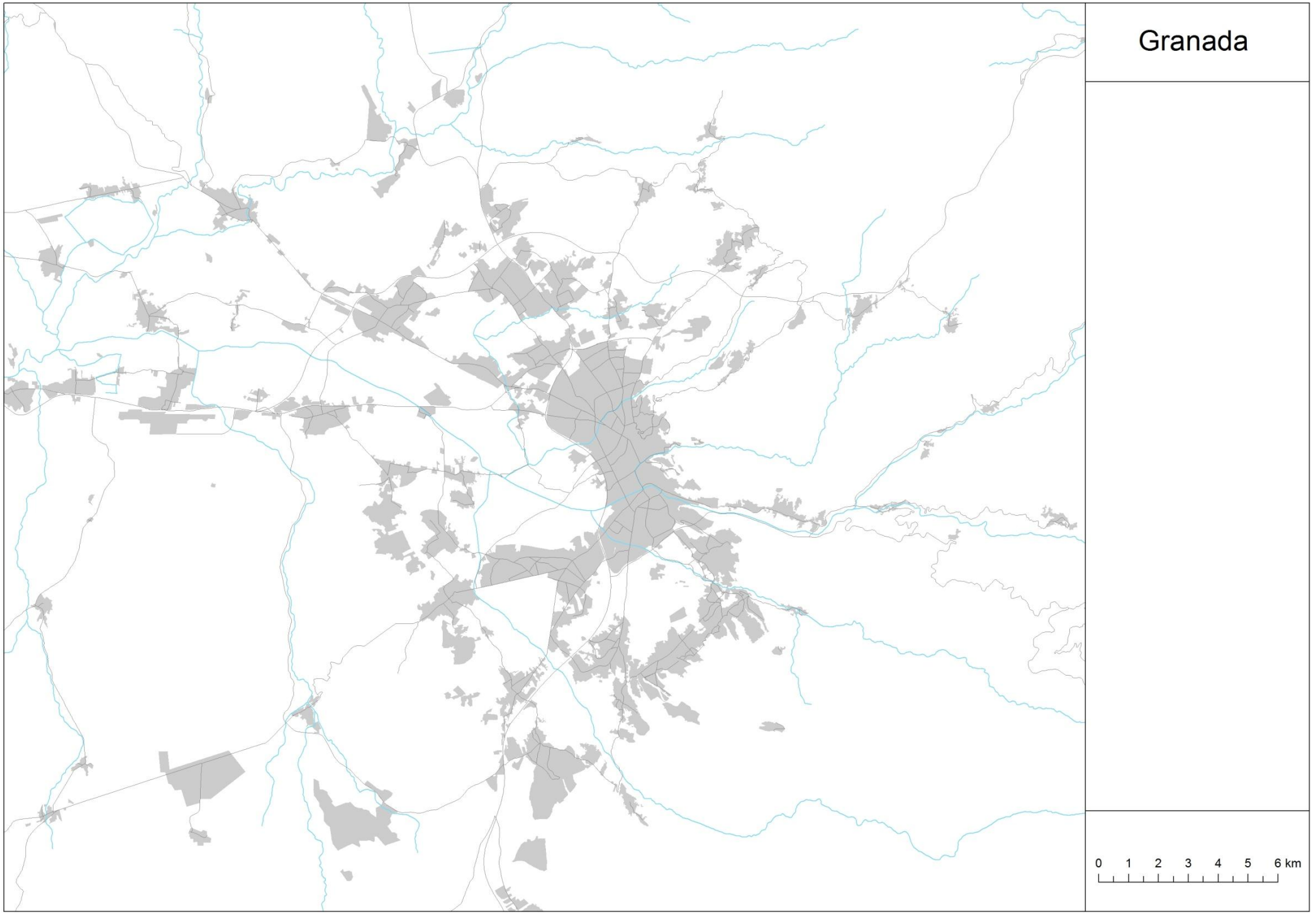
7,453,107 inhabitants

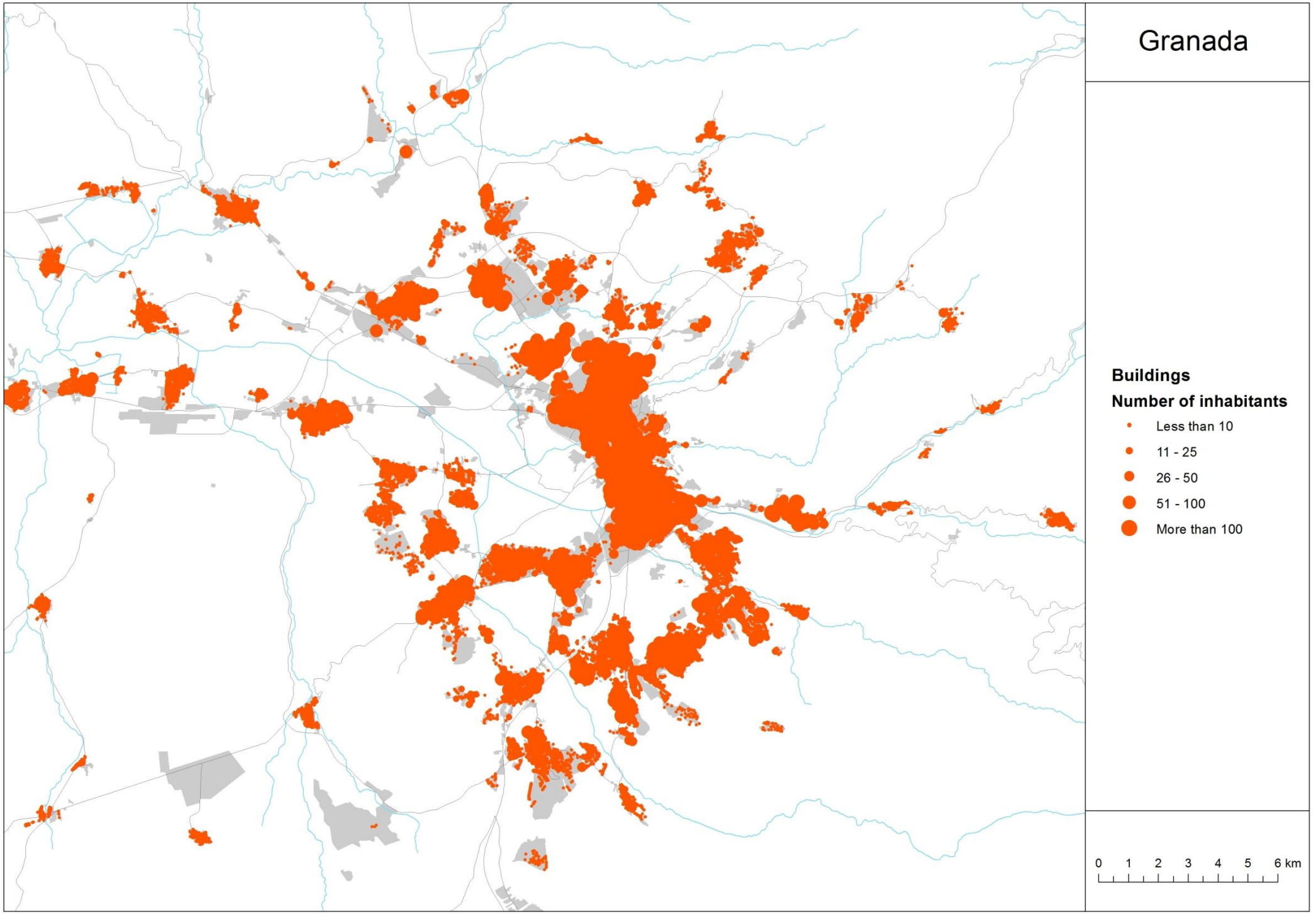
(87.4% of the population registered in RPA)

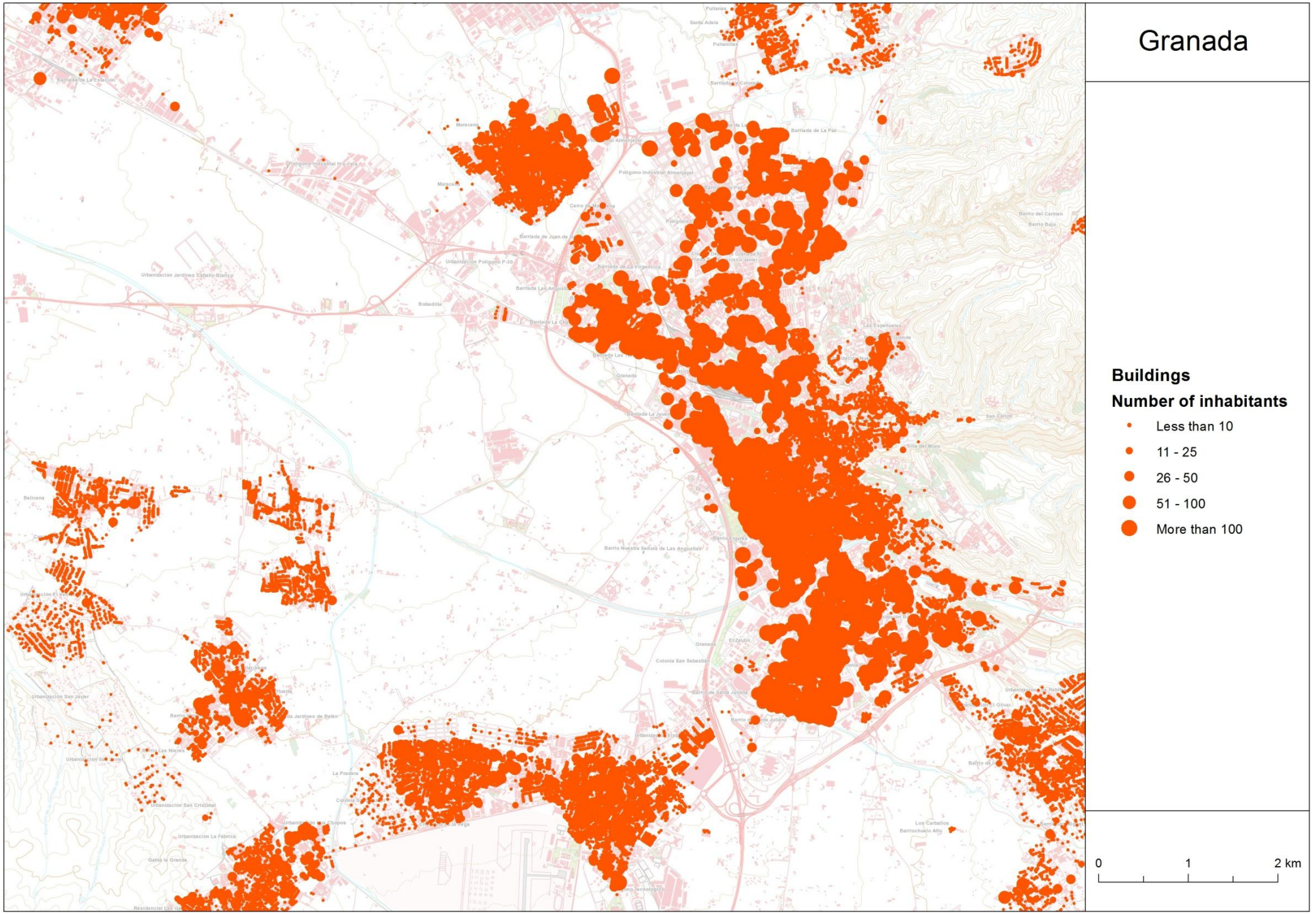
2.1. Bottom-up approach: the geocoding of buildings main entrances

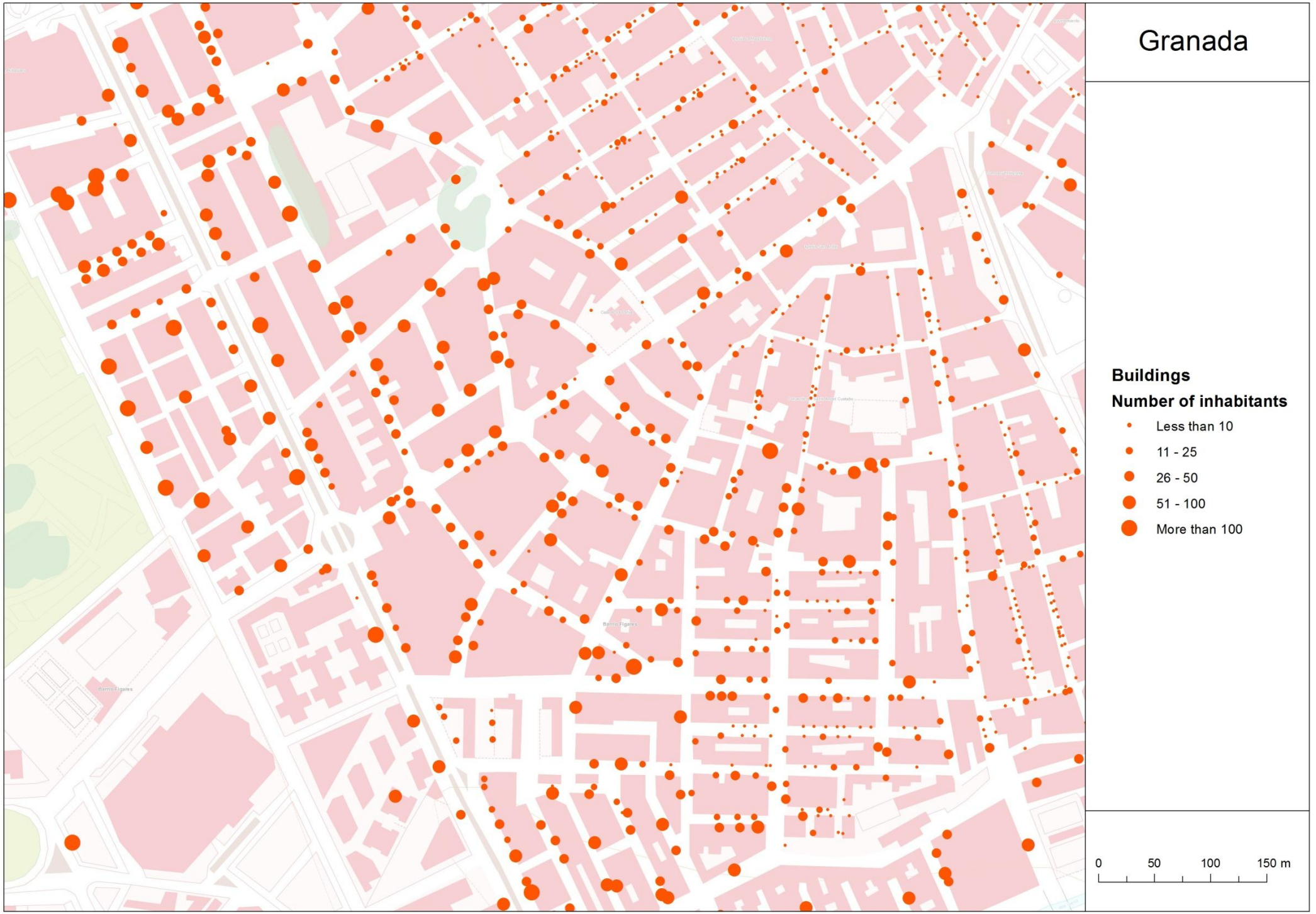
- Next step is to aggregate the information within a cell, i.e. add all points (main entrance and its population) located in a cell. This is a spatial join operation
- Each main entrance is added within a cell according to its spatial location. Each cell ends up with a total number of inhabitants and main entrances located inside the cell according to the location of the points (main entrances)
- This procedure assigned 7,453,107 inhabitants (87.4% of the population registered in RPA) to their corresponding cells. A total of 28,806 cells registered population at this stage of the process

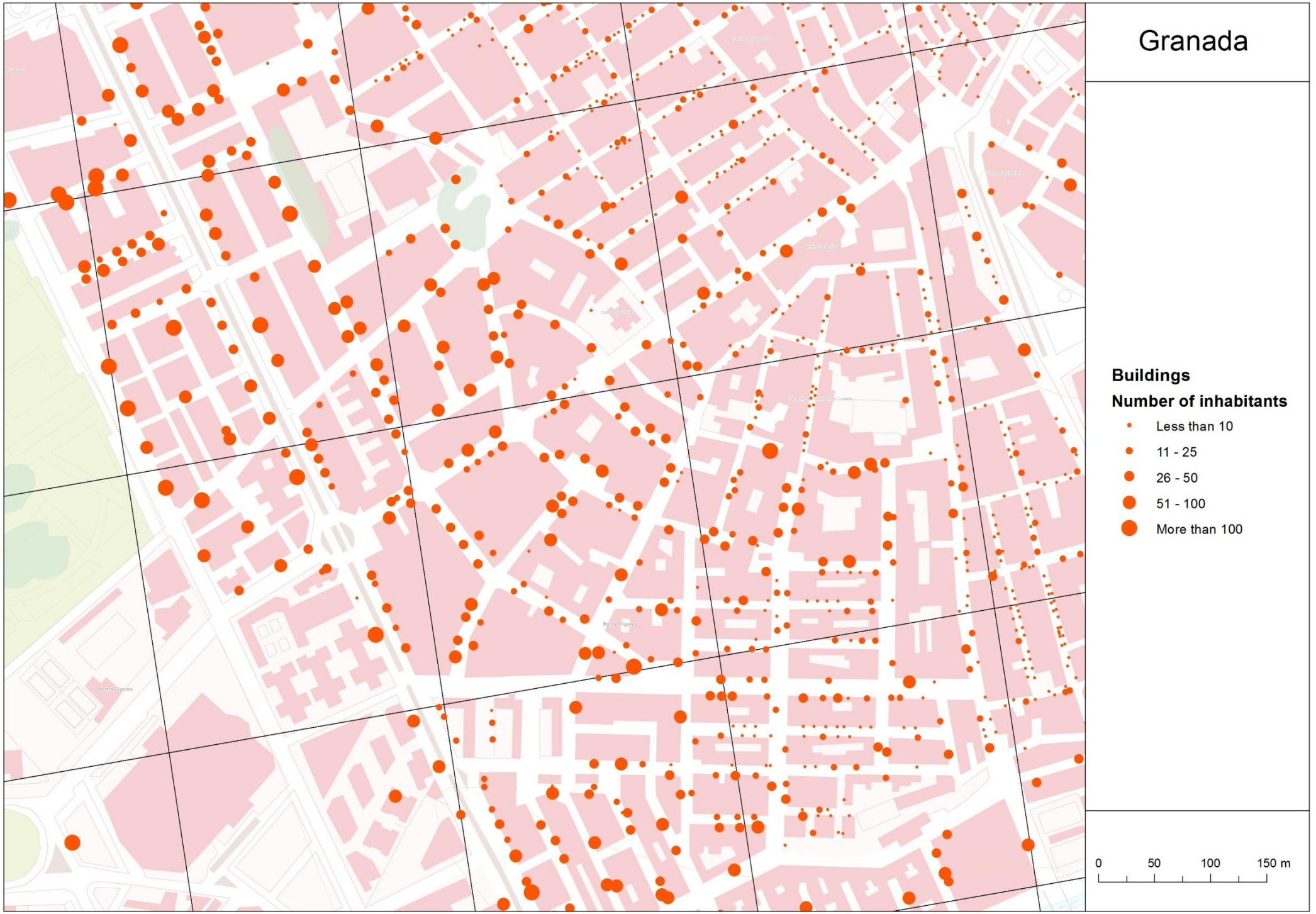
An example in the Metropolitan Area of Granada


Granada
0 1 2 3 4 5 6 km

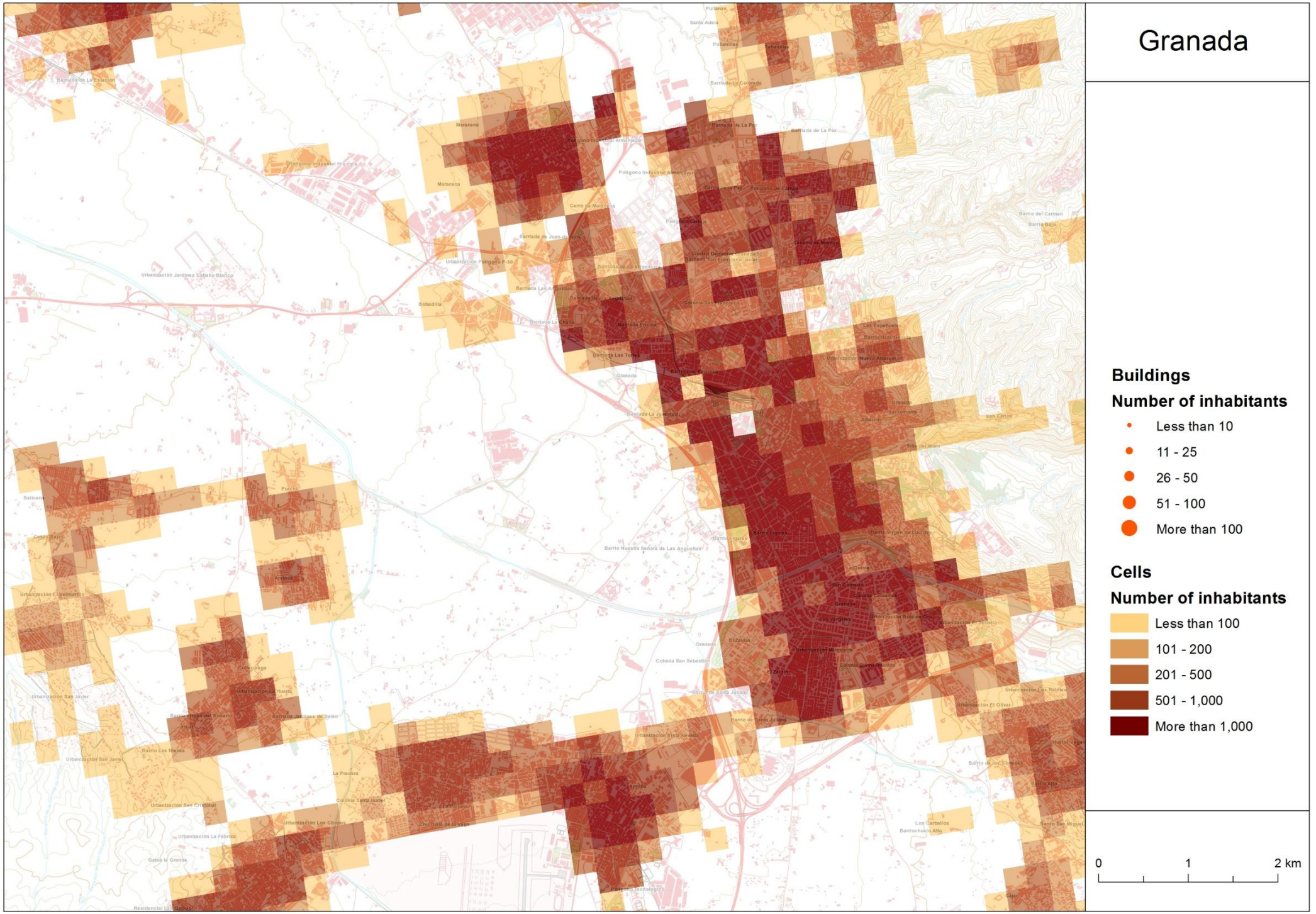

Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
0 1 2 3 4 5 6 km

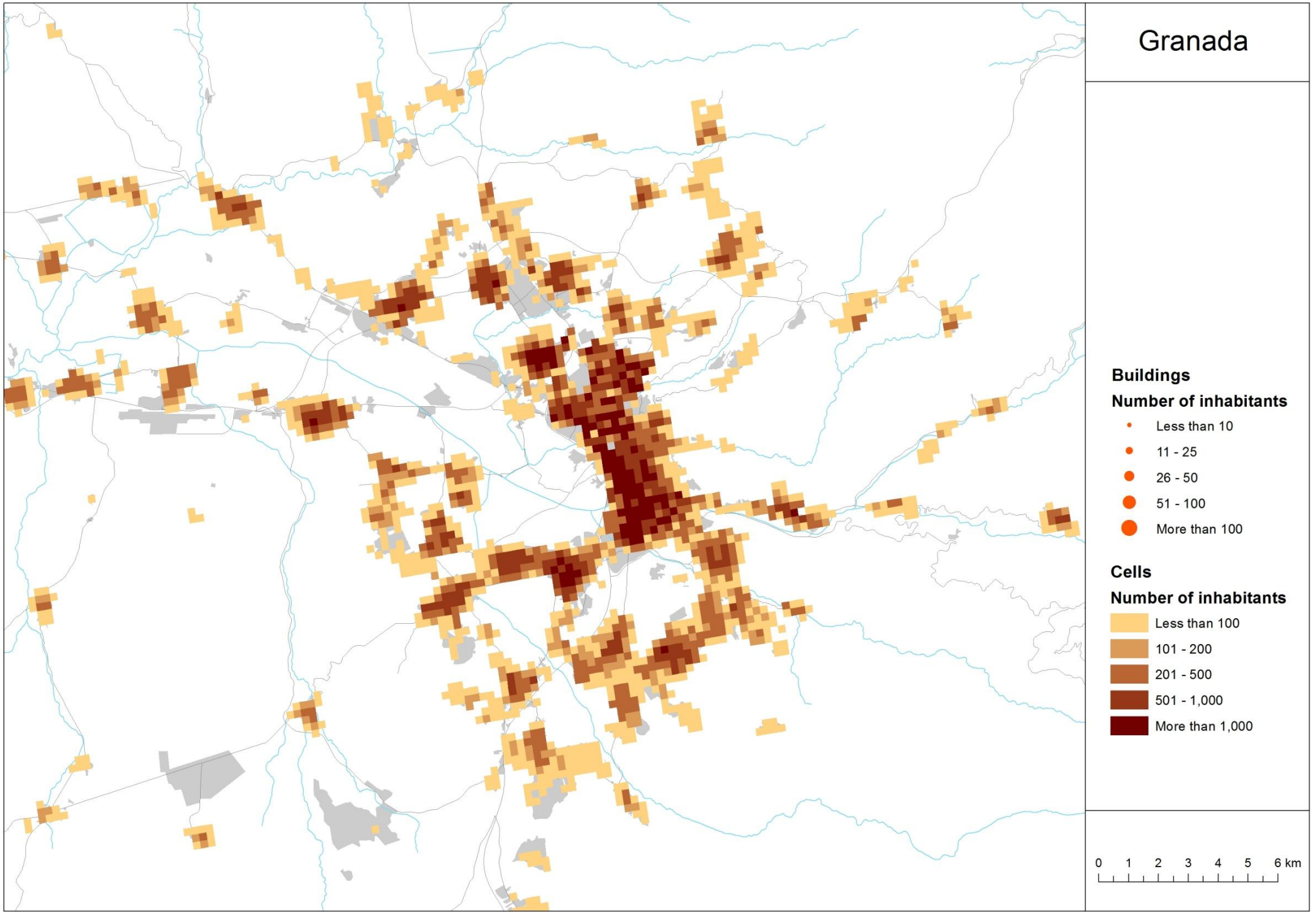

Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
0
1
2 km


Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
0
50
100
150 m


Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
0
50
100
150 m


Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
Cells
Number of inhabitants
Less than 100
101 - 200
201 - 500
501 - 1,000
More than 1,000
0
50
100
150 m


Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
Cells
Number of inhabitants
Less than 100
101 - 200
201 - 500
501 - 1,000
More than 1,000
0
1
2 km


Granada
Buildings
Number of inhabitants
Less than 10
11 - 25
26 - 50
51 - 100
More than 100
Cells
Number of inhabitants
Less than 100
101 - 200
201 - 500
501 - 1,000
More than 1,000
0 1 2 3 4 5 6 km

IECA project. Population grid cells sized 250 x 250m for Andalusia

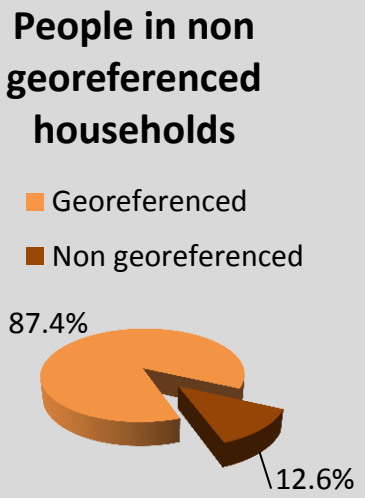
1. Highlights & main information sources
2. **Georeferentiation of the population**
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

IECA project. Population grid cells sized 250 x 250m for Andalusia

1. Highlights & main information sources
2. **Georeferentiation of the population**
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - **2.2. Top-down approach: the spatial disaggregation of non georeferenced households**
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

Question…



68% of the census-tracts have non-georeferenced households

- 12.6% of the population living in households could not be exactly georeferenced via bottom-up approach. Therefore a top-down approach was designed to allocate them spatially to a cell
- Census-tract level is the maximum territorial disaggregation available for non-georeferenced households
- Based on statistical techniques and ancillary information, the aim of the top-down approach is to transfer these data (non-georeferenced households) from the census-tracts zoning system to the grid zoning system

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

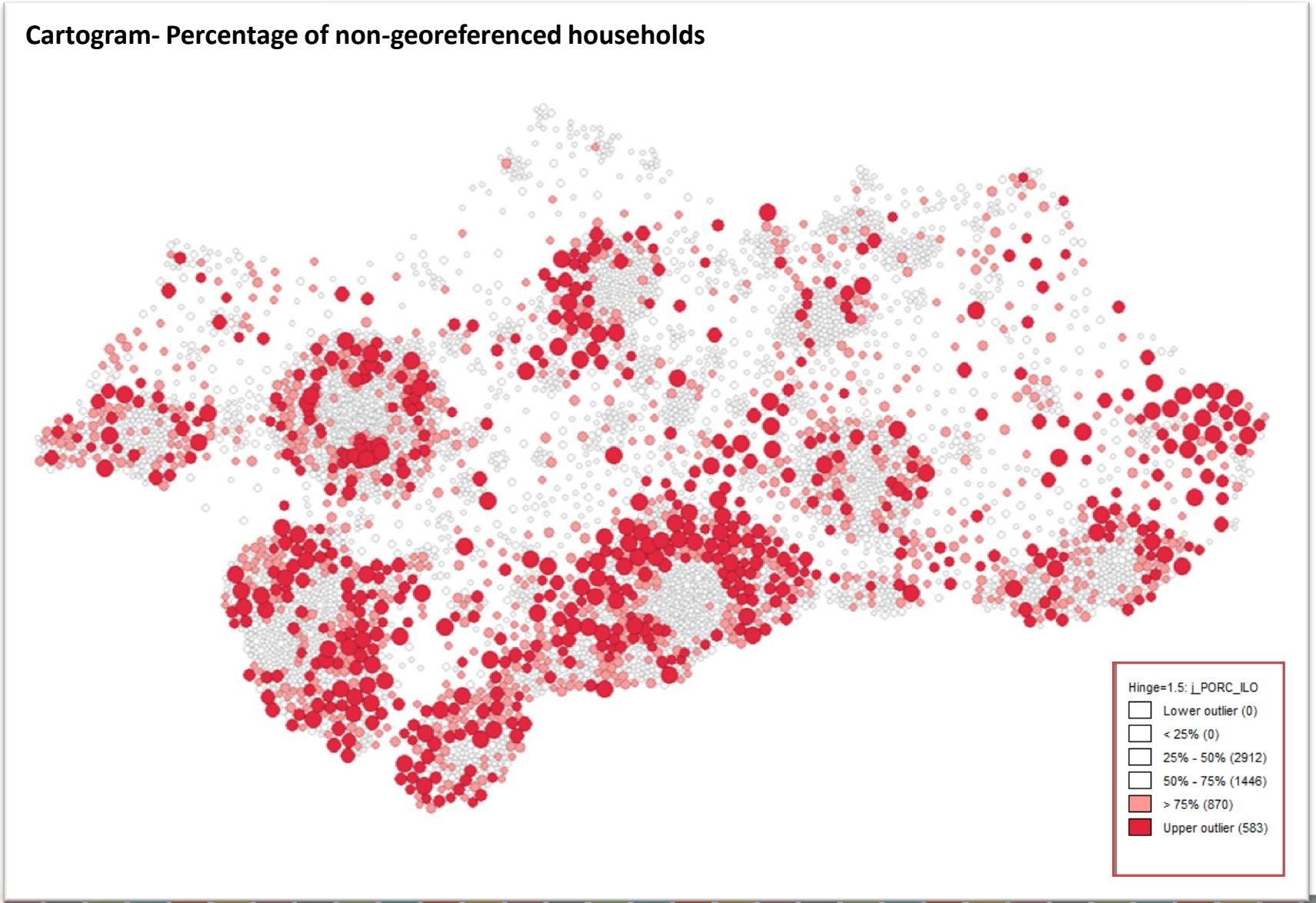
Alternatives…

To allocate spatially to a cell non georeferenced households from a census-tract

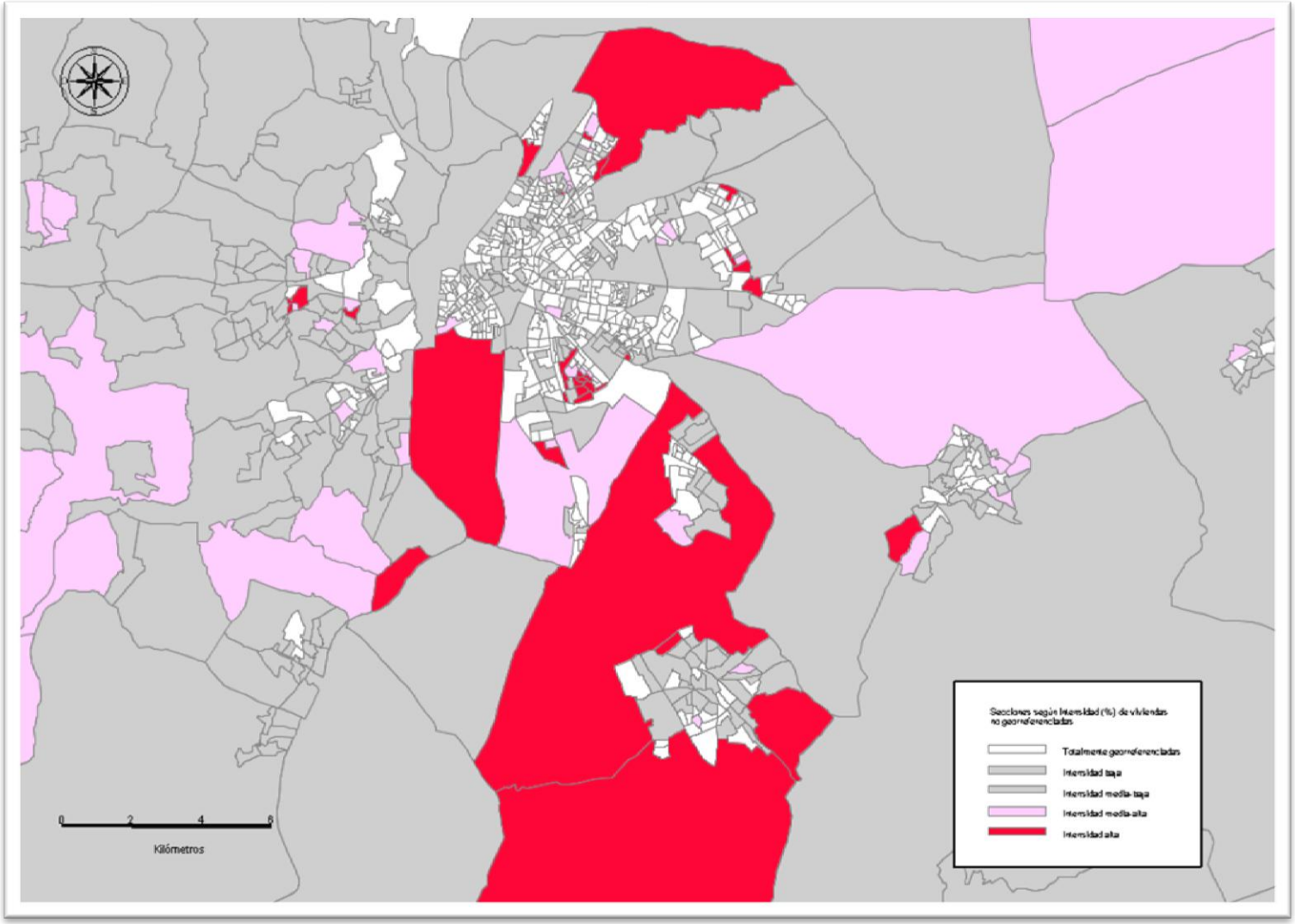
- Basic option: Assuming population distributes uniformly along the territory (choropleth)
- Alternative: Modelling settlement patterns for non georeferenced households based on ancilliary information. A descriptive analysis of georeferenced and non georeferenced households shows differences in settlement patterns and spatial correlation

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

Cartogram- Percentage of non-georeferenced households



2.2. Top-down approach: the spatial disaggregation of non georeferenced households



2.2. Top-down approach: the spatial disaggregation of non georeferenced households

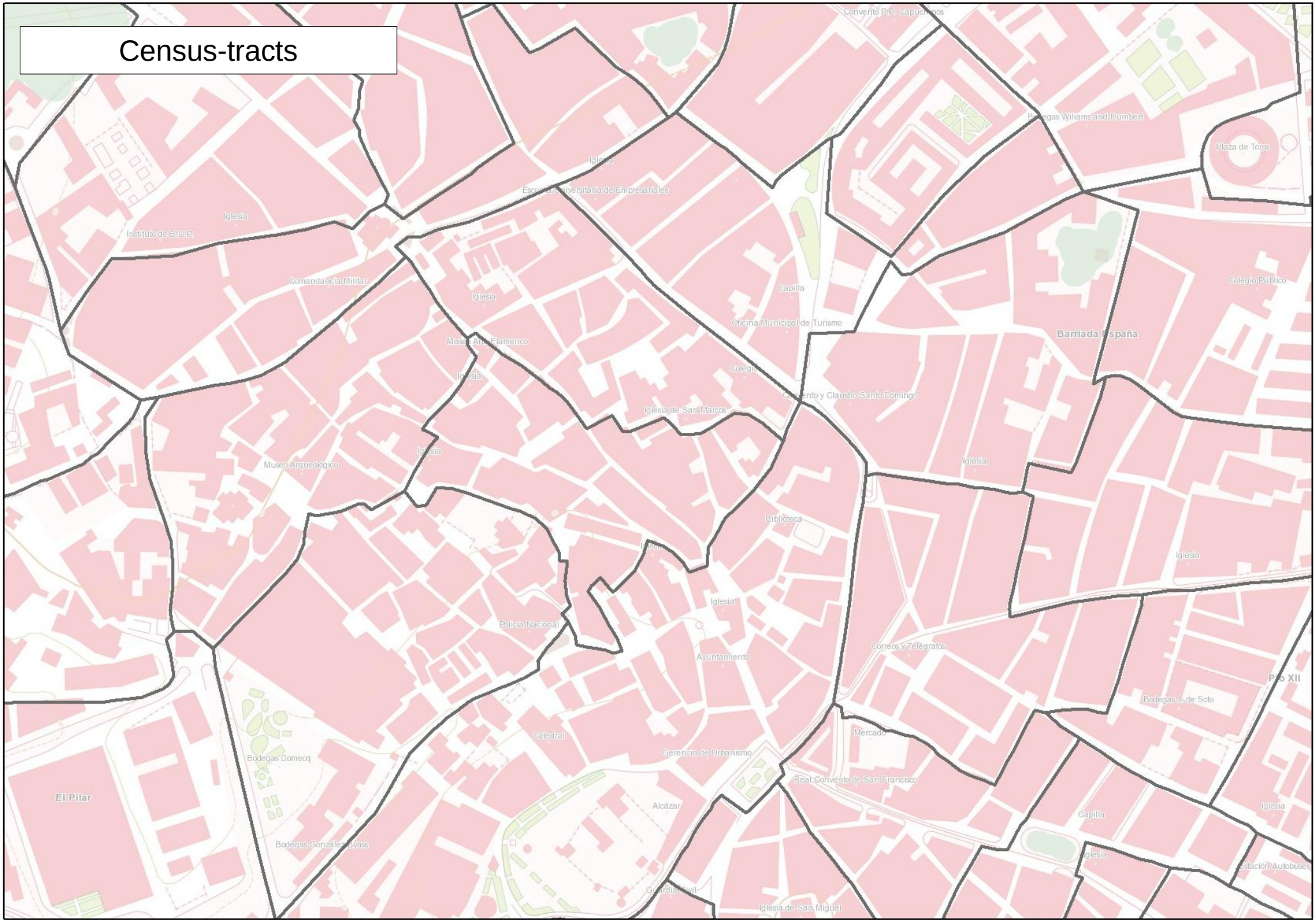
Information sources:

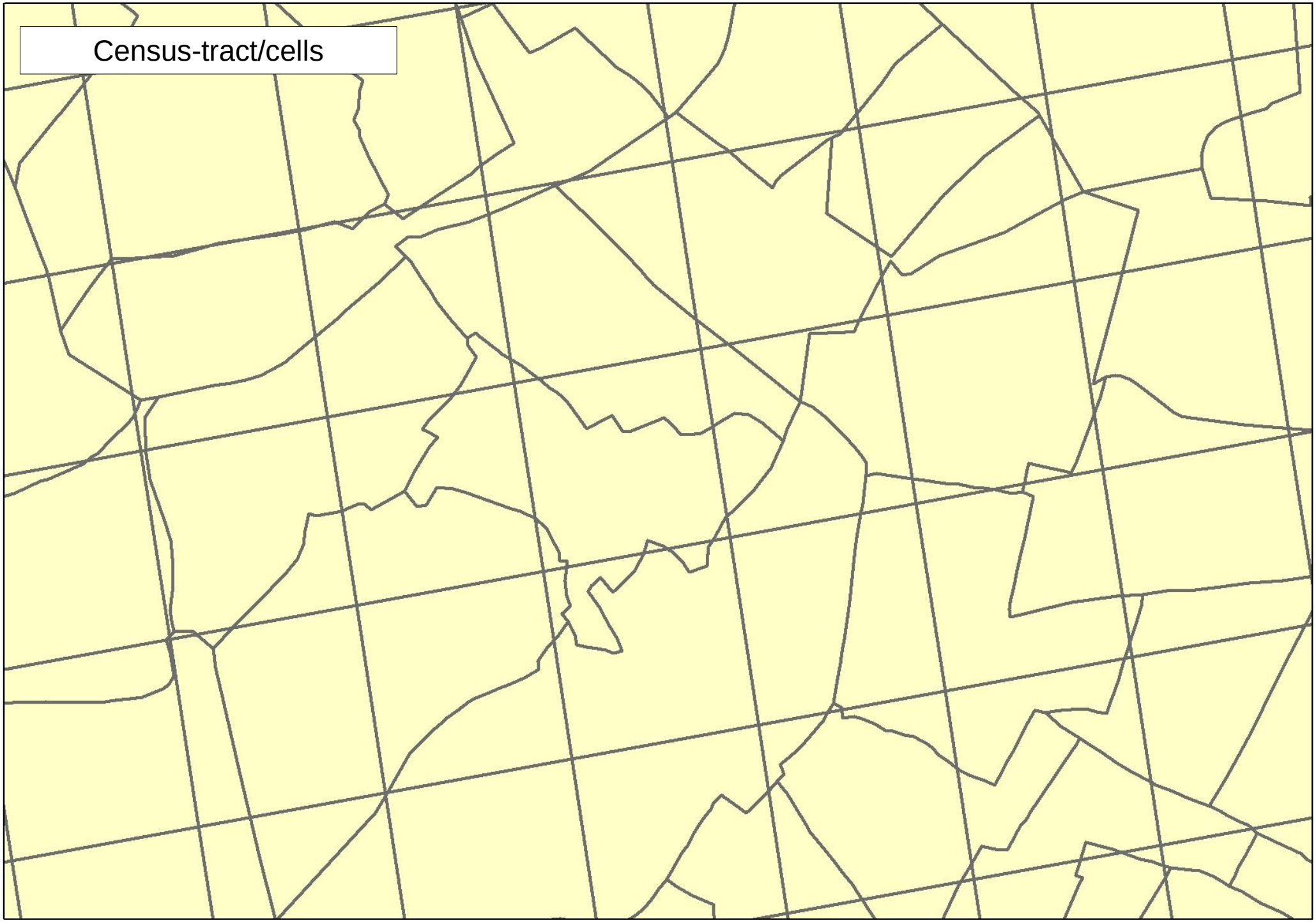
Census-tracts layer (National Statistical Institute, Spain)

And information processing…

census-tract/cells

- In order to transfer non-georeferenced households from the census-tracts zoning system to the grid zoning system a transferring spatial unit was defined: **Census-tract/cells**
- Census-tract/cells result from the intersection between the layer of census-tracts and the grid
- Census-tract/cells are the reference spatial unit for the analysis. The allocation model assigns population from census-tract to census-tract/cells
- Cells are built afterwards by adding all census-tract/cells composing a cell


Census-tracts
Convento P.P. Capuchinos
Bodegas Williams and Humbert
Plaza de Toros
Iglesia
Escuela Universitaria de Empresariales
Instituto de B.U.P.
Iglesia
Comandancia Militar
Iglesia
Capilla
Colegio Público
Oficina Municipal de Turismo
Barriada España
Museo Arte Flamenco
Iglesia
Colegio
Convento y Claustro Santo Domingo
Iglesia de San Marcos
Iglesia
Museo Arqueológico
Iglesia
Biblioteca
Iglesia
Iglesia
Policía Nacional
Correos y Telégrafos
Ayuntamiento
Pío XII
Bodegas J. de Soto
Mercado
Catedral
Gerencia de Urbanismo
Bodegas Domecq
Real Convento de San Francisco
El Pilar
Alcázar
Iglesia
Capilla
Iglesia
Bodegas González Byass
Estación Autobuses
Guardia Civil
Iglesia de San Miguel


Census-tract/cells

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

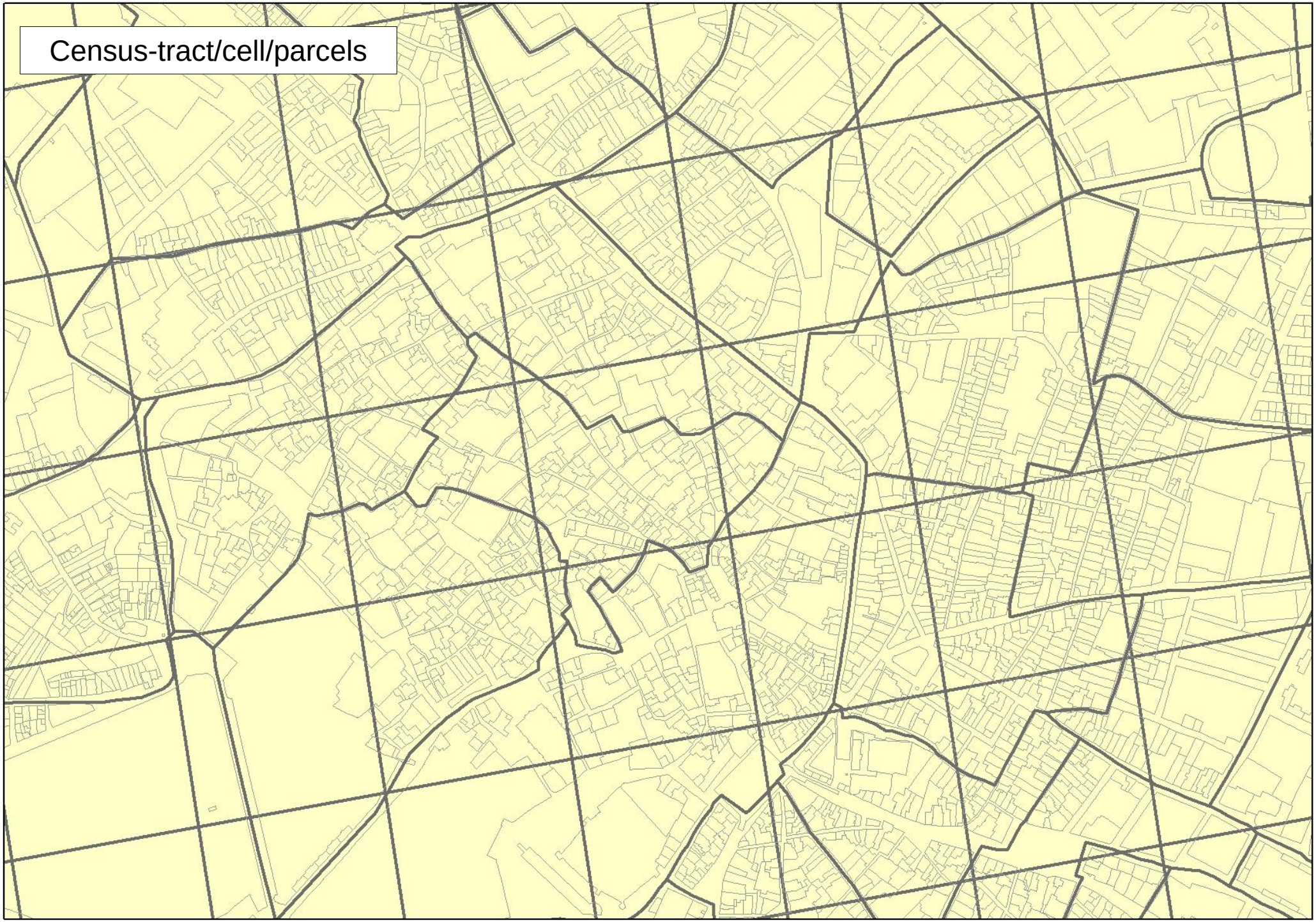
Information sources:

- Urban Cadastre layer (Cadastre Unit, Ministry of Finance and Public Administration, Spain)

And information processing…

census-tract/cell/parcels

- Cadastre territorial delimitation unit is parcel. In order to transfer parcel information to census-tract/cell unit a transferring unit was defined: census-tract/cell/parcel
- Census-tract/cell/parcels result from the intersection between the parcel layer with the census-tract/cell layer
- Cadastre analysis focused on urban parcels layers containing usage information, residential as well as other usages (residential, commercial, recreational areas versus industrial or construction works areas)


Census-tract/cell/parcels

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

Information sources:

- Urban Cadastre layer (Cadastre Unit, Ministry of Finance and Public Administration, Spain)

And information processing…

Alphanumeric data

- Residential usage parcels are key to the project; hence information available in cadastre regarding their attributes & characteristics was explored in detailed
- Properties information: age, type of road where a property is located, etc. considered relevant as non-georeferentiation is basically the result of a mismatching of a record linkage algorithm. Information referring to degree of urban consolidation (urban areas recently constructed, areas with high percentage of unusual types of roads, areas with high percentage of warehouses buildings, etc.) might shed light on the factors that hindered the linking process
- Properties' information aggregated by parcel. i.e: all attributes for the properties inside a parcel were added or averaged, depending on the cases

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

Variables appearing highly correlated with non georeferentiation incidence…

Four regression models are built by non georeferentiation intensity group

Population Register of Andalusia (RPA)

- Georeferenced households
- Georeferenced population by age
- Georeferenced population by nationality
- Georeferenced households and population in disseminated areas

Urban Cadastre

- Real estate/Immovable properties by type (industrial, residential…)
- Real estate/Immovable properties by type of road where it is located (avenue, street, square, lane, etc.)
- Real estate/Immovable properties year of construction
- Urban area and urban area containing residential real estate

2.2. Top-down approach: the spatial disaggregation of non georeferenced households

Our proposal…

Iterative algorithm based on linear regressions modelled by non georeferentiation intensity group

1. Modeling non-georeferenced households at census-tract level (administrative division)
2. Projection of non-georeferenced households at census-tract/cell level (dasymetric mapping)
3. Volume constraint, total non-georeferenced households at census-tract level (pycnophylactic constraint)
4. Models re-estimation at census-tract/cell level and process iteration until convergence criteria is attained

IECA project. Population grid cells sized 250 x 250m for Andalusia

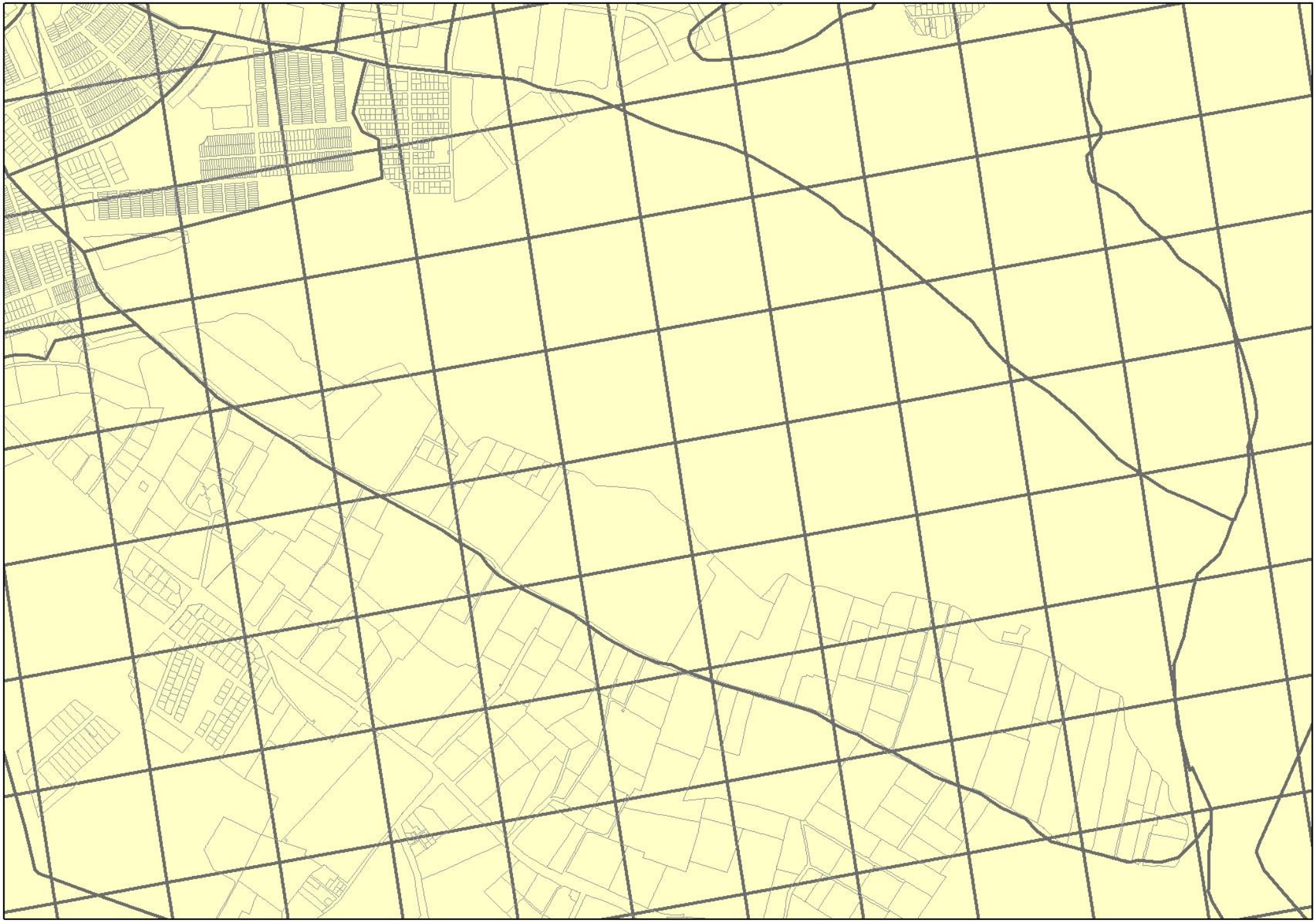
1. Highlights & main information sources
2. **Georeferentiation of the population**
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

IECA project. Population grid cells sized 250 x 250m for Andalusia

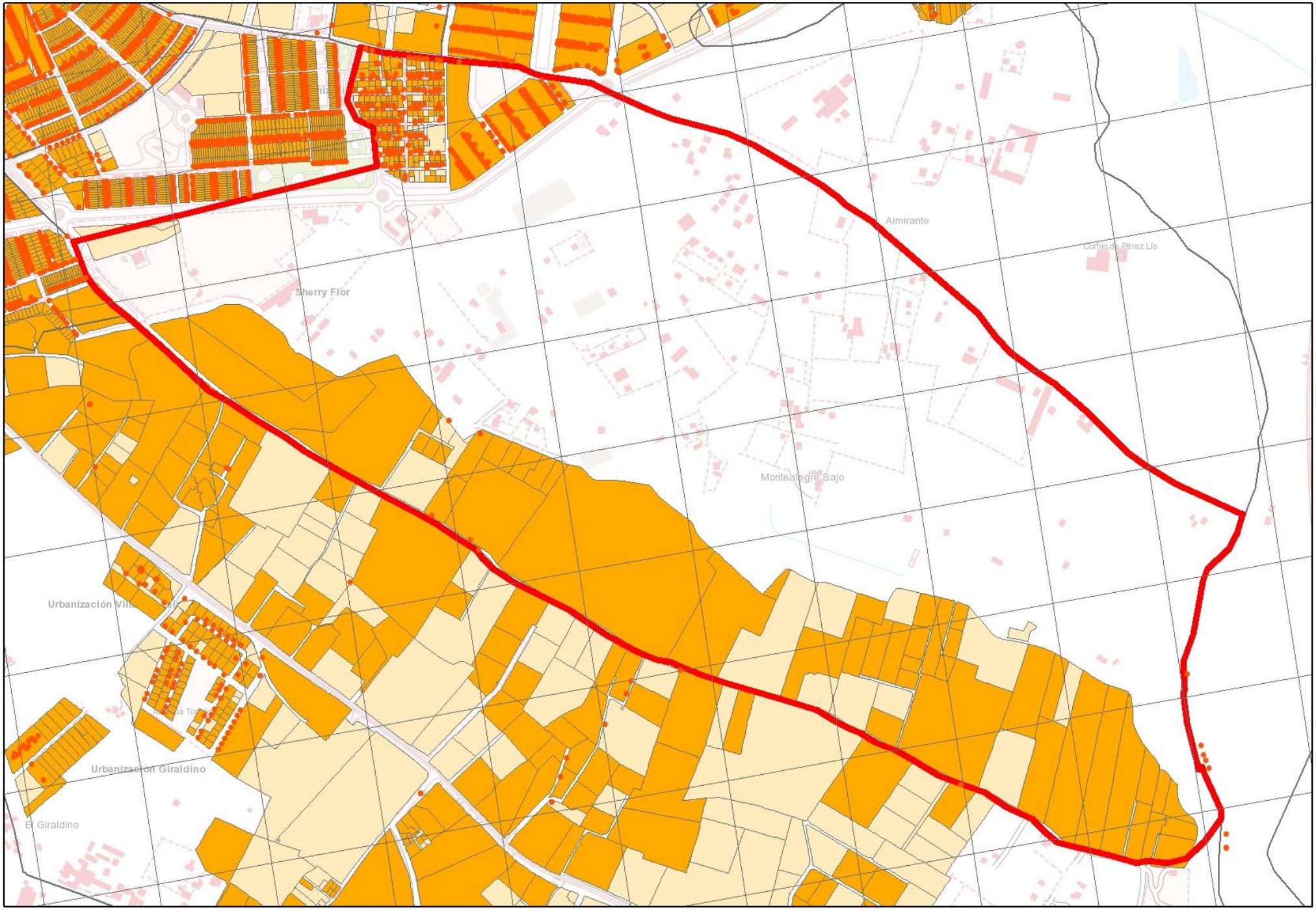
1. Highlights & main information sources
2. **Georeferentiation of the population**
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - **2.3 Hybrid results**
3. The dissemination of the results: visualization of the spatial distribution of the population in Andalusia

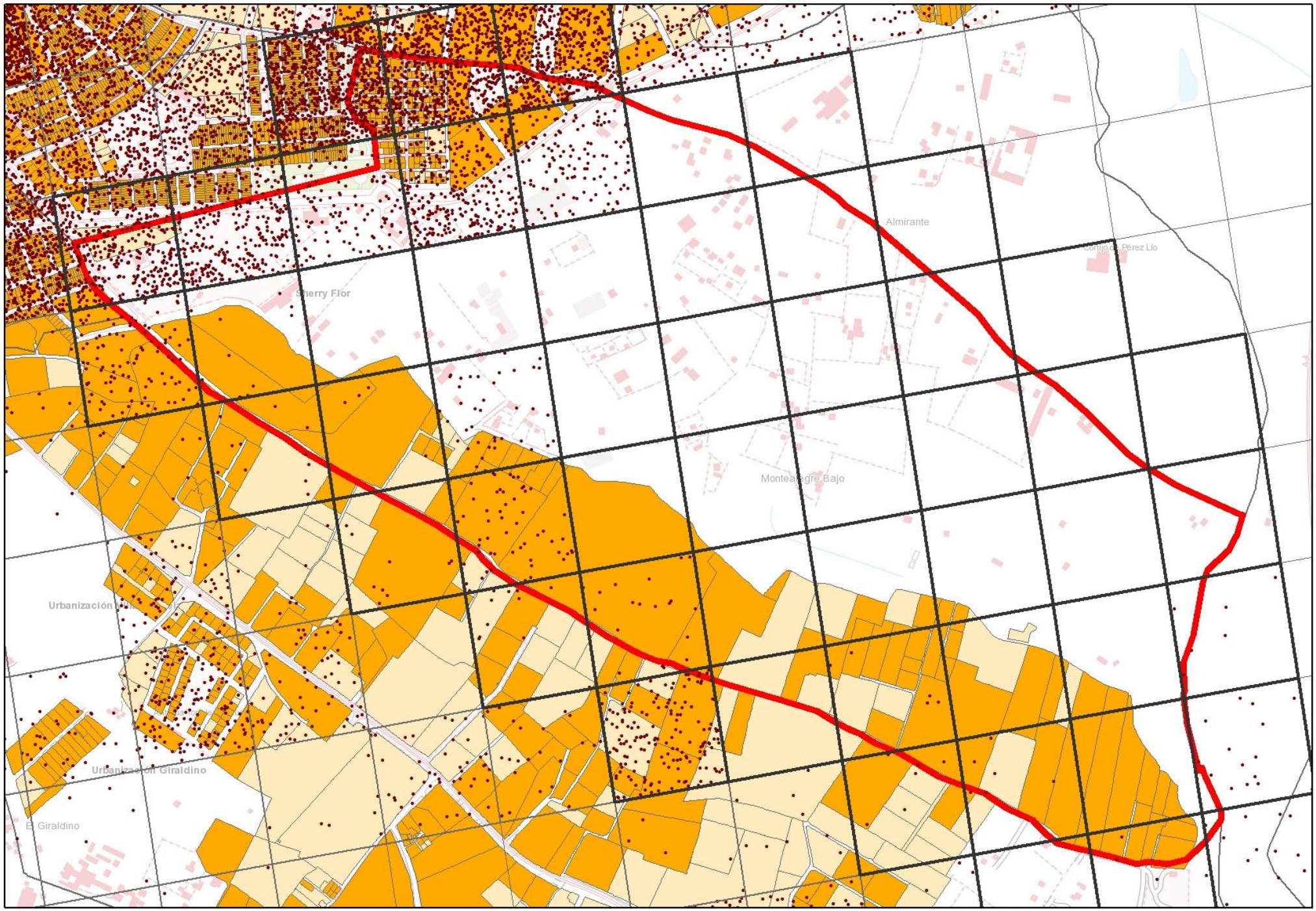
2.3. Hybrid results

- The final step is to add to each cell the geo-referenced population in the bottom-up approach (geocoding of the buildings) and the population assigned in the top-down approach (allocated non geo-referenced households)
- The end result can be seen in the following maps:

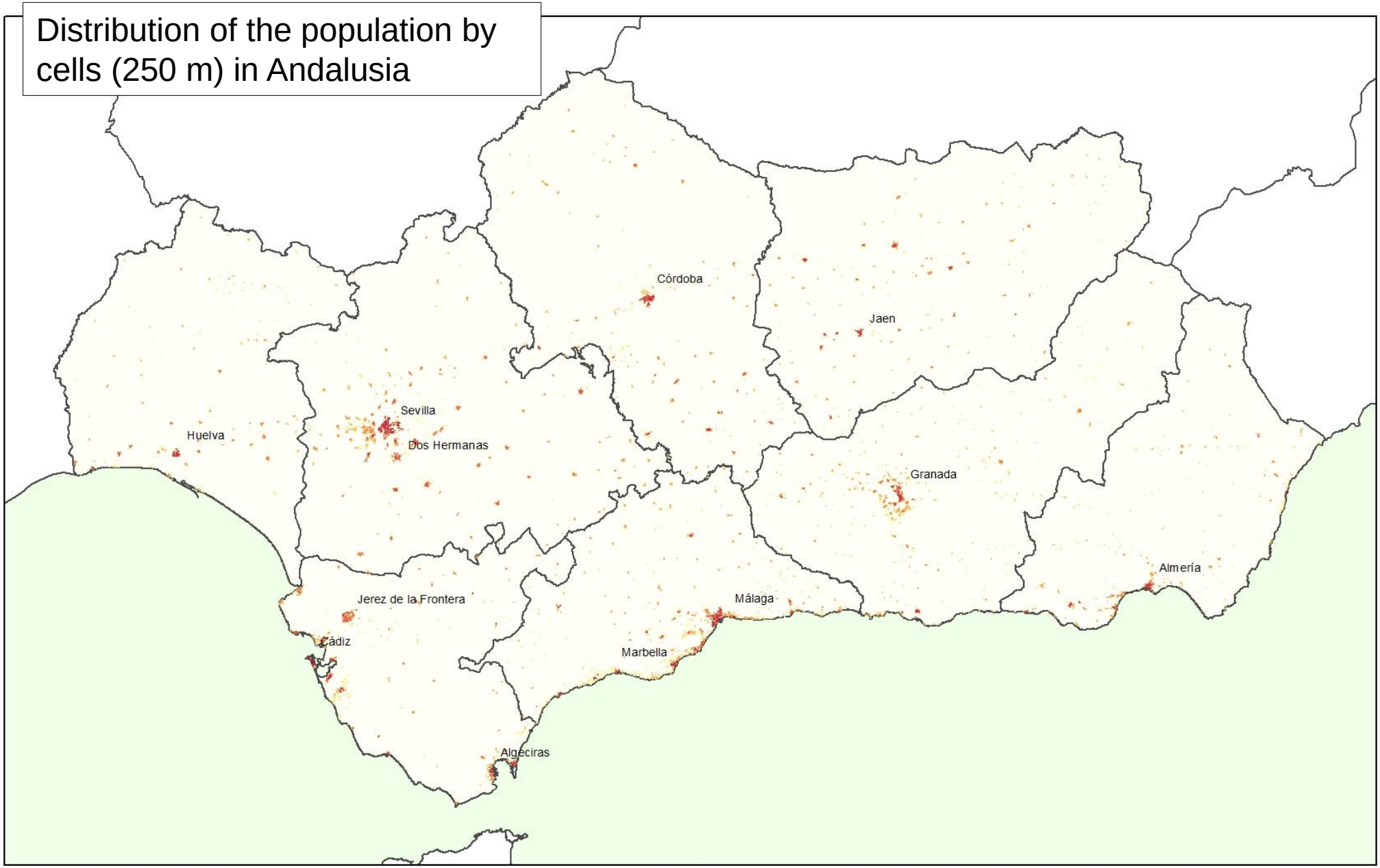



Almirante
Cortijo de Pérez Llo
Sherry Flor
Montealegre Bajo
Urbanización Villa Manuela
Montealegre
Montesalegre Bajo
Urbanización Montealegre
Urbanización Giraldino
El Giraldino


Almirante
Cortijo de Pérez Llo
Sherry Flor
Montealegre Bajo
Urbanización Giraldino
El Giraldino


Almirante
Cerro Pérez Llo
Sherry Flor
Monteagre Bajo
Urbanización
Urbanización Giraldino
Giraldino

Distribution of the population by cells (250 m) in Andalusia



IECA project. Population grid cells sized 250 x 250m for Andalusia

1. Highlights & main information sources
2. Georeferentiation of the population
   - 2.1. Bottom-up approach: the geocoding of buildings main entrances
   - 2.2. Top-down approach: the spatial disaggregation of non georeferenced households
   - 2.3 Hybrid results
3. **The dissemination of the results: visualization of the spatial distribution of the population in Andalusia**

3. The dissemination of the results

- **Visualization of the spatial distribution of the population in Andalusia**

http://www.juntadeandalucia.es/institutodeestadisticaycartografia/distribucionpob/index-en.htm

More information:

- Iria Enrique

  iria.enrique@juntadeandalucia.es
- Serafín Ojeda

  serafin.ojeda.ext@juntadeandalucia.es