The challenges of an annual update population grid including gender, age, country of origin and nationality.

Authors:
Manuel Illanes Mellid
Ignacio Duque de Arellano
Spanish Statistics Bureau
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
1. **Bottom-up methodology.**
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
The paradigm of a geostatistical mature system is founded on an exhaustive bottom-up approach:

- All units of information capture have identifiers with an associated geometry.
- All geographical units should be produced from the points address coordinates, using a bottom-up strategy.
Summary:

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. **Main issues of the approach.**
3. Previous projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Pre-Conditions:
✓ Legal framework with other agencies and Municipalities.
✓ Coordinate all statistics frameworks

In the meantime:
✓ We have to add the geo-identifiers or the coordinates to different registers or statistica final data.
✓ But it is not always possible to add the precise coordinates for all statistical records.
Summary.

1. Bottom-up methodology.
2. Main problems of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main problems of the approach.

3. *Previous projects.*
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Firm location of Business Register.
(Urban area near Bilbao)
Grid Cells by employees range (Business Register).
1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. **Next projects.**
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Related Project (already finished):
✓ Census 2.011 total population grid.
✓ First time
✓ Non exhaustive one.

Next Step:
✓ Continuous and enlarged population grid
✓ A grid with the information of all population, demographic characteristics and yearly updated
Continuous Population Register (Padron):  
✓ List of people.  
✓ Dwelling and building addresses are not in the conceptual model.  
✓ Location: Street code and name, number/Km, others (calificador, portal, escalera).

Addresses Gazetteer (Marco de direcciones GeoReferenciadas):  
✓ List of addresses (from Population register, Electoral register and other Statistical sources).  
✓ Dwelling and addresses are in the conceptual model with the information of residents.  
✓ Some of them without coordinates.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. **Challenges to achieve.**
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Main Challenges

Obtain for each year:
✓ Population Register Addresses not linked with the Gazetter
✓ Registers of the Gazetter without Coordinates.
✓ Location: Street code and name, number/Km, others (calificador, portal, escalera).

Between years:
✓ Estimations have to be the same for some registers.
✓ Random estimation should play no role.
✓ Some of them without coordinates.

Other issues:
✓ Migration of people between addresses.
✓ Changes in the Address Gazetter.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
6. **Global process.**
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Register Addresses → GeoReferenced Addresses

GeoCoding Process

Population Data Base ↔ GeoReferenced Addresses

Join of Population DB & Addresses

Group by INSPIRE Grid
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.

7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
GeoCoding Process

- Exact matching.
  - Address Code in the Gazetteer.
  - Street Code and Number.
  - Street Name and Number.
  - Street Code and probabilistic assignment to another numbers.

- Estimated assignment cell.
  - Probabilistic and weighted by population and district.

\[ \sum_{\forall \text{Cell } j} w_{ij} = 1 \]

\[ w_{ij} = \frac{p_{ij}}{\sum_{\forall \text{Cell } j} p_{ij}} \]
1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. **% of success by step of the process.**
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
% Population by assignment steps

% GeoCoding Steps excluding step 1

- Street Code and Number: 46.10%
- Street Name and Number: 28.36%
- Probabilistic to another numbers: 25.50%
- Weighted by district: 0.04%

% GeoCoding Steps

- Address Code: 91.52%
- Street Code and Number: 2.16%
- Street Name and Number: 2.40%
- Probabilistic to another numbers: 3.91%
- Weighted by district: 2.40%
% Adresses by assignment steps

% GeoCoding Steps
- Address Code: 87.04%
- Street Code and Number: 0.01%
- Street Name and Number: 4.20%
- Probabilistic to another numbers: 0.01%
- Weighted by district: 3.80%
- Other: 5.18%

% GeoCoding Steps excluding step 1
- Street Code and Number: 31.84%
- Street Name and Number: 39.31%
- Probabilistic to another numbers: 28.80%
- Weighted by district: 0.05%
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary:

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.

9. **Tools to GeoReference**.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
GeoCoding Tools

- ArcGis.
  - Weighted Population Tables.
  - View of Results.
- SAS.
  - Join of Data Bases.
  - Grouping of Variables.
- R.
  - GeoCoding Process.
  - Use of the libraries ArcGisBinding, Rodbc, SP, Rgeos, SF and dplyr.
  - All the layers are located in a remote Corporate GIS Server.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Testing and tuning the parameters.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Comparative with other sources.

Cells by population strata

Population Register 2014  Census 2011

Population Register 2014, Bottom-Up  Goerlich (SIOSE) 2011, Top-Down
Comparative with other sources.

Cells by population strata

Population Register 2014
Region of Andalucia 2013

Cells by population strata

Population Register 2014
Region of Madrid 2014
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Adaptable to other Grids.
Summary

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. Workflow.
Summary.

1. Bottom-up methodology.
2. Main issues of the approach.
3. Previous projects.
4. Next projects.
5. Challenges to achieve.
7. GeoCoding.
8. % of success by step of the process.
10. Test and tuning of the parameters.
11. Comparative with other sources.
12. Another Grids.
13. **Workflow.**
Pre-treatment of Data Base.

Sequential process of GeoReferencing. Iterative process by regions. Assignment of coordinates and INSPIRE cells.

Check Up and Testing.

Grouped Variables by Grid.

Saved Partial Results and Generation of Maps.
Thank you for your attention!

Your turn for questions.

Email: manuel.illanes.melli@ine.es