Statistical data for French cities
Between zones and grids
Man vs. wild geography

Part I (yesterday) – Beasts awaitings

The point of view of the geographer through several practical examples: how it is easy to misinterpret spatial data

Part II – Human needs, survival recipes

The point of view of the final users and the structure of the answer from the provider of statistical data

Answering doesn’t mean giving up with adequate answers
› Background
› User needs
› Living up to expectations
Preliminary issue: zonings in France
Did you say “gropingly”? 

› Above municipality level
  – INSEE’s morphological cities
  – Cooperative structures: administrative cities

› Inside municipalities
  – INSEE’s output areas, theoretically agreed by municipalities
  – Additional zonings used by some municipalities
  – Successive rounds of delineation of official “deprived neighborhoods”: all active, intersections allowed
    - ZUS – actions for population (1997)
    - ZFU – actions for the development of economic activity (1997)
    - CUCS – actions for population (2006)

➔ Some municipalities say “stop” (Hamel 2009)
Preliminary issue: data confidentiality
Braking not required!

- For statistics about firms: at least 3 observations, without anyone above 80% of the total. Unbreakable.
- For statistics about population:
  - Disclosure of individual data can occur occasionally
  - Essentially a geographical approach: some zonings or zones are formally agreed without consideration of size or possible intersections
    - Municipalities: minimum size is 0 inhabitants
    - IRIS (census standard dissemination areas): idem
    - Any zone officially used for public policies: idem
  - With some additional constraints set by the initial owner (e.g. 10 dwellings for fiscal data)
  - But population counts are not concerned!
Short history of infra-urban statistics

Before 1999: Only census data, available up to census data collection districts, no geometry.

1999: The beginning
Geometry available for municipalities (900) above 10,000 inhabitants.
A special production system is set to produce data from administrative sources, but restricted to a predefined set of zones, target for public policies and to a slightly larger set of municipalities (1200)

2003: More zones
Extension of administrative data production to the standard output areas used in the dissemination of the census (IRIS).
The set up of the new census collection process offers the first opportunities to localize data at address level (above 10,000 inh.); first experiments about the use of such data.

2006: Zones and grids together
First industrial data production of grided data (1999 census, administrative sources, above 10,000 inh.) in the context of the coming round of delineation of neighborhood for public policies.

2010 (17/7): First grided map of population for the whole country.
Controlling the changes
The virtuous cycles

- Statistical offices of ministries
- Users
- Commission for statistical information
- Partnerships with Local users
- INSEE

INSEE

Commission for statistical information

Partnerships with Local users

Statistical offices of ministries

Users
Following considerations come from...

› An unusual status for the French NSI
  – Data provider and disseminator, but also:
  – Making analyses by itself
  – Making analyses in close cooperation with local authorities

› The National Committee for Statistical Information (CNIS)
  – Task force about the spatial distribution of population, animated by the director of the town planning agency for Bordeaux. Report published September 2009.
From data to analysis: two potential paths

(1) All purposes geolocalized individual data

(2) Aggregated data for some geographical zones
Examples

Path 1

...but what is the real need?

Path 2
Both of them are a part of the solution

According to the report, users have to:

› Report their action for well defined defined territories
  – Build time series for policy evaluation purposes
  – Get precise counts of population, employment… for controlling the allocation of public resources
  => Fine tuned aggregated data = Take path 2

› Prepare future actions
  – Select target territories for public action without any consideration of already selected ones
  – Estimate potential populations for various scenarii
    - transportation, day care for children
  => Flexibility = Take path 1
### Path 2
**Unemployment in Marseille Deprived neighborhoods**

<table>
<thead>
<tr>
<th>Code région</th>
<th>Code unité urbaine (agglomération)</th>
<th>Code ZUS</th>
<th>Nom de la zone</th>
<th>Disponibilité (voir documentation)</th>
<th>DEFM toutes catégories</th>
<th>Catégorie 1</th>
<th>Catégo</th>
</tr>
</thead>
<tbody>
<tr>
<td>93</td>
<td>00753</td>
<td>9302010</td>
<td>BEISSON</td>
<td>0</td>
<td>164</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>00753</td>
<td>9302030</td>
<td>CORSY</td>
<td>0</td>
<td>181</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>00753</td>
<td>9302050</td>
<td>JAS DE BOUFFAN</td>
<td>0</td>
<td>1033</td>
<td>631</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>00753</td>
<td>9305010</td>
<td>CENTRE NORD</td>
<td>0</td>
<td>7598</td>
<td>5180</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>00753</td>
<td>9305020</td>
<td>LA ROSE, FRAIS VALLON, PETIT SEMINAIRE</td>
<td>0</td>
<td>1720</td>
<td>1098</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>00753</td>
<td>9305030</td>
<td>LES HAUTS DE MAZARGUES</td>
<td>0</td>
<td>538</td>
<td>366</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>00753</td>
<td>9305041</td>
<td>NORD LITTORAL : PLAN D'AOU-BRICARDE...</td>
<td>0</td>
<td>1055</td>
<td>1105</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>00753</td>
<td>9305042</td>
<td>L'ESTAQUE-SAUMATY</td>
<td>0</td>
<td>353</td>
<td>225</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>00753</td>
<td>9305051</td>
<td>GUINZIEME SUD : LA VISTE, CONSOLAT,...</td>
<td>0</td>
<td>2149</td>
<td>1375</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>00753</td>
<td>9305051</td>
<td>ST ANTOINE EST : LA SAVINE, ND LIMITE</td>
<td>0</td>
<td>1722</td>
<td>1159</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>00753</td>
<td>9305071</td>
<td>ST BARTHÉLEMY, LE CANET, DELORME...</td>
<td>0</td>
<td>4575</td>
<td>2973</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>00753</td>
<td>9305080</td>
<td>SAINT MAURONT, BELLEVUE, CABUCELLE</td>
<td>0</td>
<td>1690</td>
<td>1195</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>00753</td>
<td>9305090</td>
<td>AIRBEL</td>
<td>0</td>
<td>511</td>
<td>343</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>00753</td>
<td>9305101</td>
<td>VALLEE DE L'HOUEAUNE : ROUGUIERE,...</td>
<td>0</td>
<td>1390</td>
<td>893</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>00753</td>
<td>9305120</td>
<td>MALPASSE, ST JEROME</td>
<td>0</td>
<td>1344</td>
<td>1322</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>00753</td>
<td>9307020</td>
<td>LA LEQUE, AIGUES DOUCES</td>
<td>0</td>
<td>334</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>00753</td>
<td>9307030</td>
<td>AMARANTES</td>
<td>0</td>
<td>113</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>00753</td>
<td>9307040</td>
<td>LES COMTES</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Path 1
Low income zones
In the city of La Rochelle

Agglomération : La Rochelle
Commune : La Rochelle
Source : Cnam/TS 2006, Insee
None of them gives the final solution

› Path 1 is not fully available
  – Individual data cannot be disseminated (confidentiality)
  – Getting a precise location is not always possible (wrong location in the example); getting coordinates is expensive

› Path 2 is not fully available
  – Another aspect of the MAUP: everyone defines zones according to its own needs. One zoning is not enough but it is impossible to make a census of every existing zoning or to have some common basic block
  – Converting data from one zoning to another is possible but resources consuming
  – Zonings are a bad friend for people doing analyses (wrong interpretation in the example)
Spatial data dissemination about cities in France: a body with two heads

› For policy making: detailed statistical data for a small number of fixed zonings, with an aim of comparability within time
  – User defined zones (deprived neighborhoods, etc…)
    - (almost) no restriction of size and shape but stability is mandatory
  – INSEE’s own zoning as a default one
    - 2,000 inh. basic blocks

› For planning the future, no individual data dissemination but:
  – detailed statistical data on zones defined « on the fly ».
    - Data is mainly estimated, some conditions must be fulfilled (e.g. a minimum population size).
  – rough estimations for counts estimation purposes on zones defined « on the fly »
    - Without any condition of size but restricted to « population » counts
  – estimations of detailed statistical data for small fixed zones
    - for spatial structure analysis purpose only: data make sense only as a whole.
› For policy making: detailed statistical data for a small number of fixed zonings, with an aim of comparability within time
  – User defined zones (deprived neighborhoods, etc…)
    - Service is open; more current zones are freely disseminated on the web
  – INSEE’s own zoning as a default one
    - Disseminated on the web

› For planning the future:
  – detailed statistical data on zones defined « on the fly ».
    - Service will be opened at the end of this year for census data and extended to administrative sources later
  – rough estimations for counts estimation purposes on zones defined « on the fly »
    - Not open yet; dissemination for small grids will take place first (autumn)
  – estimations of detailed statistical data for small fixed zones
    - Data exists (1ha grids), dissemination should take place this year.
    - Data confidentiality will be reached by shuffling
Mixing statistics and urban planners view

INSEE-AUAT

INSEE-Préfecture

JL.LIPATZ

SCORUS meeting

August 2010
Thank you
Any questions?