

btb : Beyond the Border

Un package d'analyse Urbaine

Auriane Renaud, Arlindo Dos Santos et François Sémécurbe

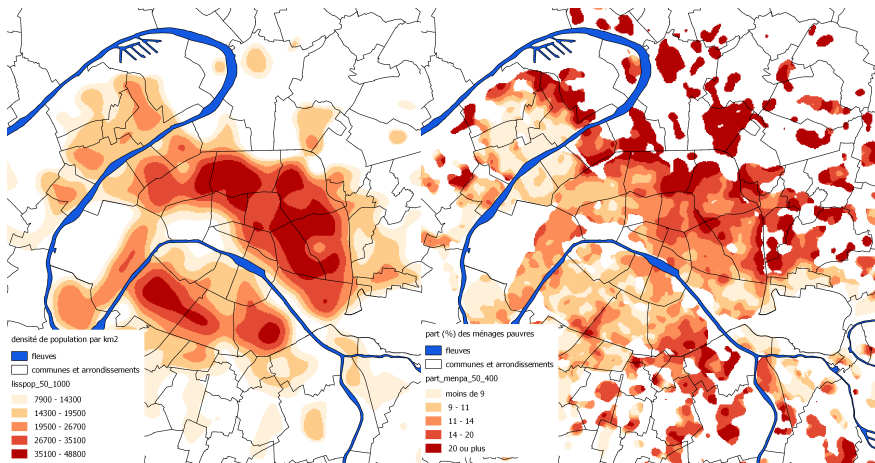
Pôle Analyses Urbaines - Insee

June 12, 2017

Plan

- 1 Introduction
 - Analyses urbaines?
 - Les outils
 - Étude de l'offre existante
- 2 Le package R
 - Les choix techniques
 - Les choix politiques
- 3 Le futur
 - Rajouter des fonctionnalités

Exemples de cartes



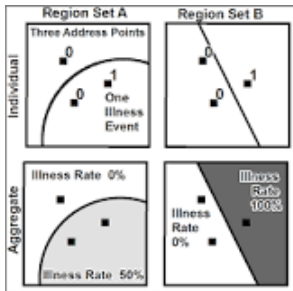
Source : Filosofi 2012

L'analyses urbaines en quelques mots

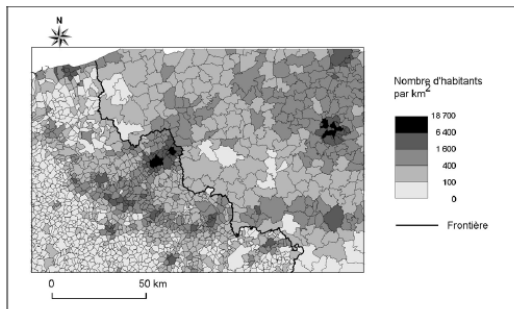
L'Insee a vocation à éclairer le débat public («Mesurer pour comprendre») en proposant à la fois des données de qualité et des études pertinentes. Son champ d'action est multiscale: de l'échelle nationale aux Quartiers Prioritaires.

- Au sein de l'Insee, nous disposons de nombreuses bases de données de grandes tailles géolocalisées: la localisation est précise (longitude et latitude);
- Nous travaillons en «Partenariat» avec des partenaires institutionnels pour répondre à leurs besoins de connaissance;
- Pour les acteurs territoriaux usuels, nous pouvons mobiliser nos données sur leurs territoires de souveraineté mais quid des acteurs dont la vocation est foncièrement infraterritoriale... les acteurs urbains...

Pourquoi Lisser... Le Modifiable Areal Unit Problem



source: « Andresswift – CC
By SA »



Sources : Eurostat (1990), Megrim (1991), d'après Dumas, Guérois (2001)

Source : Mariane Guerrois (2003)

Les outils de l'analyse urbaine

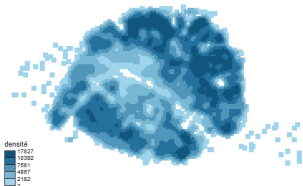
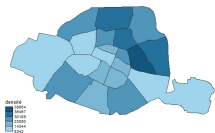
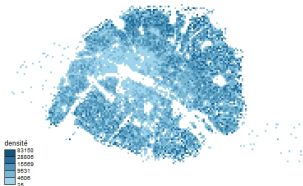
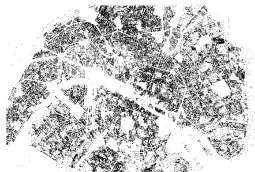
Trois supports géographiques pour restituer les données statistiques:

Le territoire est un découpage non régulier de l'espace. Plusieurs difficultés: le MAUP, comment associer une information à un territoire...

Le carroyage est un découpage régulier de l'espace sous la forme d'une grille de cellules carrées. Par construction, les données carroyées peuvent être très chahutées

Le lissage peut être considéré comme une extension du carroyage et consiste à décrire non pas la population du carreau mais son environnement dans un rayon donné.

Exemple : densité de population



source: Filosofi 2012

Les packages existants

- Plusieurs packages contiennent des fonctions de lissage: KernSmooth, spatstat... Ces fonctions reposent sur une Transformée de Fourier Rapide pour calculer la convolution (dénomination mathématique du lissage).
- Par construction, ce procédé n'est pas adapté aux situations où les effets de bords sont importants...
- Situation courante en statistique: ne pas envoyer de la population à la mer ou à l'étranger
- Proposer des indicateurs qui explorent en profondeur la distribution locale d'une variable... les quantiles mobiles

Plan

- 1 Introduction
 - Analyses urbaines?
 - Les outils
 - Étude de l'offre existante
- 2 Le package R
 - Les choix techniques
 - Les choix politiques
- 3 Le futur
 - Rajouter des fonctionnalités

Exploiter la capacité d'interfacer R avec C++

Le logiciel R est ultra polyvalent et à une grammaire très flexible mais en contrepartie il est lent:

	Fortran	Julia	Python	R	Matlab	Octave	Mathe- matica
	gcc 5.1.1	0.4.0	3.4.3	3.2.2	R2015b	4.0.0	10.2.0
fib	0.70	2.11	77.76	533.52	26.89	9324.35	118.53
parse_int	5.05	1.45	17.02	45.73	802.52	9581.44	15.02
quicksort	1.31	1.15	32.89	264.54	4.92	1866.01	43.23
mandel	0.81	0.79	15.32	53.16	7.58	451.81	5.13
pi_sum	1.00	1.00	21.99	9.56	1.00	299.31	1.69
rand_mat_stat	1.45	1.66	17.93	14.56	14.52	30.93	5.95
rand_mat_mul	3.48	1.02	1.14	1.57	1.12	1.12	1.30

source: julialang.org , le C++ est la référence.

- Pour contourner cette limite, nous avons développé un package basé sur un développement en C++ à l'aide du package Rcpp .
- Nous bénéficions ainsi de la Syntaxe R et de la puissance du C++ avec un coût de développement relativement modeste.



Un package réalisé par des membres de l'Insee

L'Insee ne produit pas de packages...

- btb a été réalisé par des agents de l'Insee mais n'est pas un produit de l'Insee...
- pour des questions juridiques...
- et de circuit de validation.

Néanmoins, il est ou va être utilisé par de nombreux acteurs institutionnels.

Plan

- 1 Introduction
 - Analyses urbaines?
 - Les outils
 - Étude de l'offre existante
- 2 Le package R
 - Les choix techniques
 - Les choix politiques
- 3 Le futur
 - Rajouter des fonctionnalités

Fonctions en cours de développement

- Intégrer des outils pour proposer des analyses du type fonction de Ripley;
- Paralléliser les calculs pour gagner en vitesse;
- Définir des pôles d'activités pour l'étude du commerce de centre-ville...