



# Méthodologie des données carroyées

## CNIS : Commission Territoires

Vincent Loonis

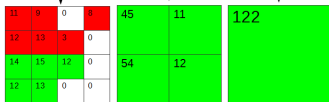
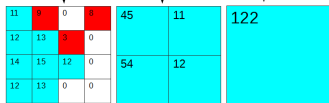
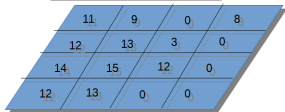
Division des Méthodes et Référentiels Géographiques

26 octobre 2017

- En 2013, la méthodologie de diffusion données de RFL 2010 dans des rectangles était très adhérente à la source (méthode des rectangles).
- Celle de Filosofi 2015 devrait préfigurer une méthode transposable à toutes les sources géoréférencées finement tout en gérant la confidentialité.
  - **Gestion post-tabulée** : diffusion dans un système fixé a priori de grilles emboîtées.
  - **Gestion pré-tabulée** : identification des observations à risque dans une grille donnée, puis perturbation des données.
  - **Mixte des deux méthodes.**

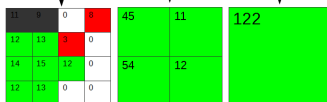
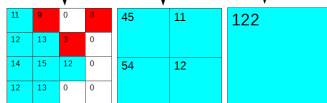
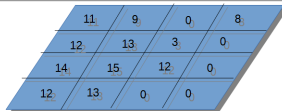
## Pour un seuil de diffusion de 11 observations

Exemple de grille superposées



■ Diffusé
 ■ Non diffusé
  Case vide

Exemple de grille superposées avec diffusion au plus fin



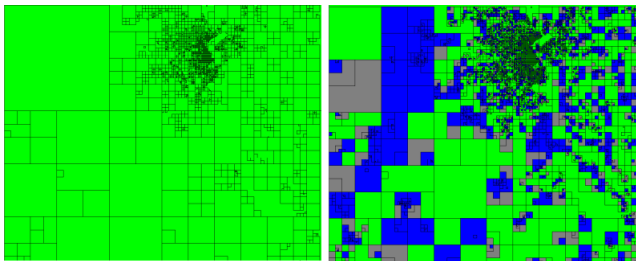
■ Diffusé
 ■ Non diffusé
  Case vide

Non diffusé pour pouvoir descendre au plus fin

# Exemples de gestion post-tabulée

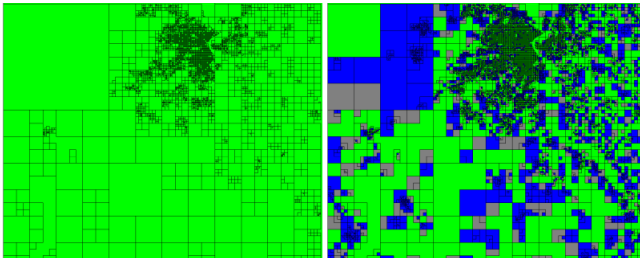
Découpage autour de Bordeaux, pour la diffusion des données de FiLoSoFi 2014, en grilles superposées de maille 125m, 250m, 500m, 1000m.

À gauche, arrêt dès qu'un carreau a moins de 11 carreaux et diffusion au niveau supérieur.



À droite, niveau de diffusion le plus fin possible, sans rompre la confidentialité : les carrés gris correspondent au carreaux non diffusés et les bleus à ceux qu'on ajoute par rapport à la figure de gauche.

La même chose avec les données de la CNAF et un seuil de 5 observations



- Choix de la taille de la grille
- Cohérence avec une diffusion sur d'autres niveaux géographiques (Iris, QPV, Communes...)
- Apport et acceptabilité des méthodes de swapping
- Définition d'une liste de variables sensibles
- Comment gérer les données du recensement qui sont issues d'un sondage ?