

# **Analyse exploratoire des données géo-référencées**

# SIG et Analyse Spatiale

- ▶ **Systemes d'Information Géographique (SIG)**
- ▶ **Analyse spatiale**
- ▶ **Données spatiales ou géo-référencées**

# Systemes d'Information Géographique (SIG)

## ▶ SIG en tant qu'ensemble d'outils

- Burrough : « set of tools for collecting, storing, retrieving at will, transforming and displaying spatial data from the real world for a particular set of purposes »

## ▶ SIG en tant que science (la « nouvelle » géographie)

- Goodchild : Science de l'information géographique

⇒ questions scientifiques génériques relatives aux données géographiques

⇒ rôle central de l'analyse spatiale

## ▶ Fonctions SIG (Anselin et Getis , 1992)

1. Input
2. Archivage
3. Analyse
4. Output

# Qu'est-ce que l'analyse spatiale ?

- ▶ Des données à l'information...
  - valeur ajoutée au-delà de la cartographie des données
  - transformation, manipulation et application de méthodes analytiques aux données spatiales (géographiques)
- ▶ Absence d'invariance par rapport à la localisation spatiale
  - analyser **où** est-ce que les résultats changent quand la localisation de l'objet d'étude change ?
    - ▶ Concentration géographique (clusters) – Autocorrélation spatiale
  - l'espace géographique joue un rôle important



Analyse spatiale *avant la lettre* : carte des décès dus à l'épidémie de choléra à Londres en 1854 établie par le Dr. John Snow (1813-1858), épidémiologiste (14000 morts, taux de mortalité de 1.3%)



La majorité des décès dus à l'épidémie de choléra de Soho vivait plus près de la pompe infectée de Broad Street que de n'importe quelle autre pompe (cf. diagramme de Voronoï)<sup>37</sup>

# Les différentes étapes de l'analyse spatiale

- ▶ **Trouver** des schémas spatiaux pertinents
  - Analyse Exploratoire de Données Spatiales ou Géo-référencées
- ▶ **Visualiser** ces schémas spatiaux
  - Analyse cartographique
- ▶ **Expliquer** ces schémas spatiaux
  - Modélisation spatiale, régressions spatiales
- ▶ L'analyse spatiale mobilise
  - une approche théorique spécifique
  - des outils informatiques et cartographiques spécifiques

# Mise en œuvre de l'analyse spatiale

- ▶ Au-delà du SIG
  - Fonctionnalités analytiques non intégrées dans les SIG commerciaux standard
  - L'exploration requiert une approche interactive
  - La modélisation spatiale requiert des méthodes statistiques spécifiques
    - ⇒ Traitement explicite de l'autocorrélation spatiale
    - ⇒ L'espace-temps n'est pas l'espace + le temps
- ▶ ESDA et Econométrie Spatiale

# Données spatiales ou géo-référencées

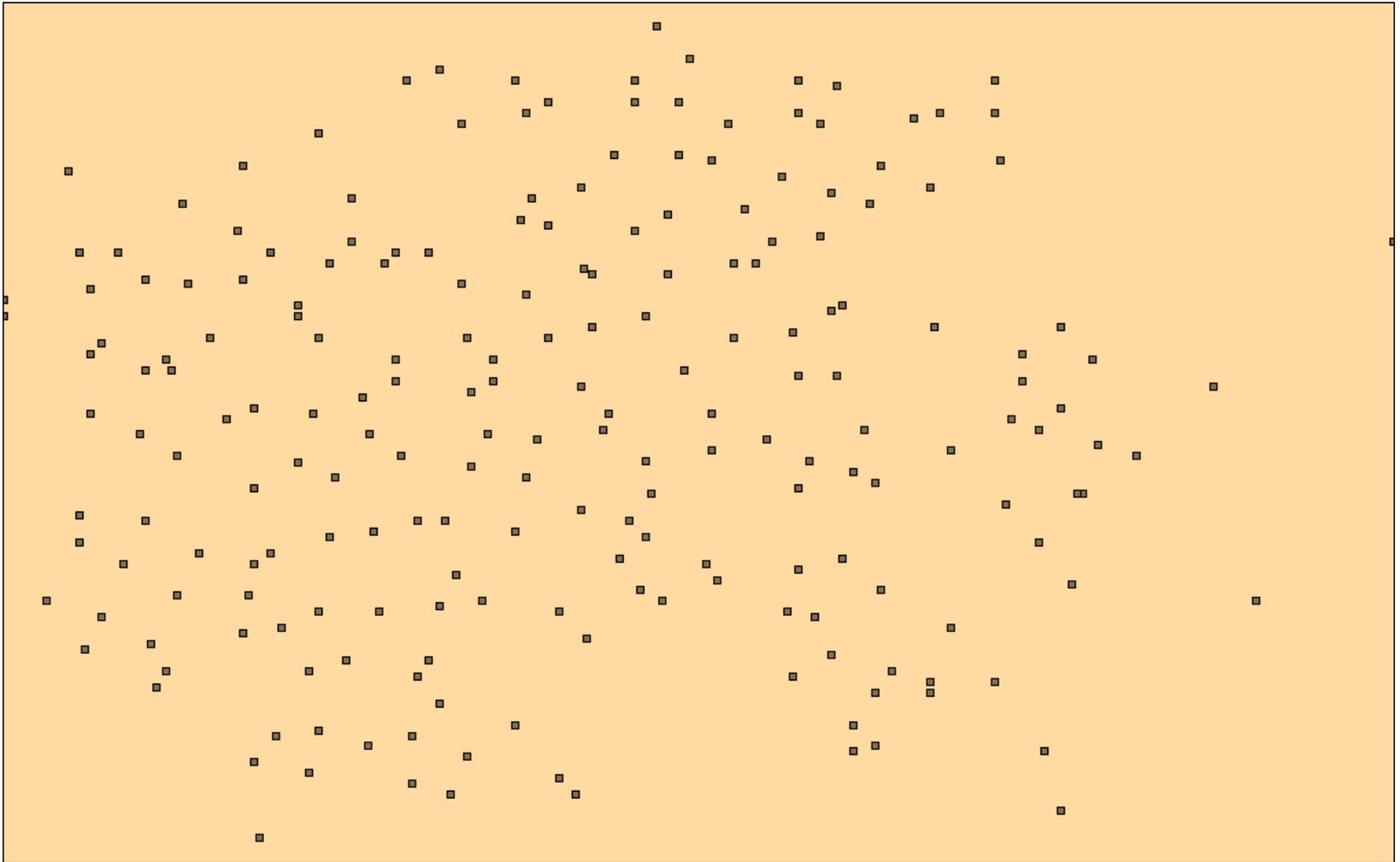
« What is Special about Spatial Data? » (Anselin)

- ▶ La localisation relative des données joue un rôle dans la compréhension et l'analyse des phénomènes économiques et sociaux
  - ⇒ **Elle doit être prise en compte dans l'analyse**
- ▶ La dépendance en coupe transversale / spatiale est très plausible
  - Interactions spatiales, contagion, externalités spatiales, effets de débordement géographique (spillovers), effets d'imitation (copy-cattng)
- ▶ Première loi de la Géographie (Tobler, 1979)
  - « Everything depends on everything else, but closer things more so »

# Nature des données spatiales

Deux type de données spatiales :

- ▶ **Données géo-référencées** : observations localisées dans l'espace
  - Informations sur la valeur de la variable + informations sur la localisation
  - Cas particulier de données en coupe transversale
- ▶ **Objets spatiaux**
  - **Points : coordonnées x,y**
    - ▶ Localisations d'objets : villes, usines, entreprises, magasins, maisons...
    - ▶ Localisation d'événements : transactions, accidents, crimes...
  - **Lignes : arcs**
    - ▶ Réseaux de transport, réseaux de communication
  - **Aires : polygones** (ensemble d'arcs reliés entre eux  $\Rightarrow$  unités territoriales)
    - ▶ Pays (NUTS0), Régions (NUTS2), Départements (NUTS3) (UE)
    - ▶ Regions, States, Counties, Census tracts, Census Blocks... (USA)



**Points** : 211 transactions immobilières dans le comté de Baltimore (1980)



**lignes** : un réseau routier



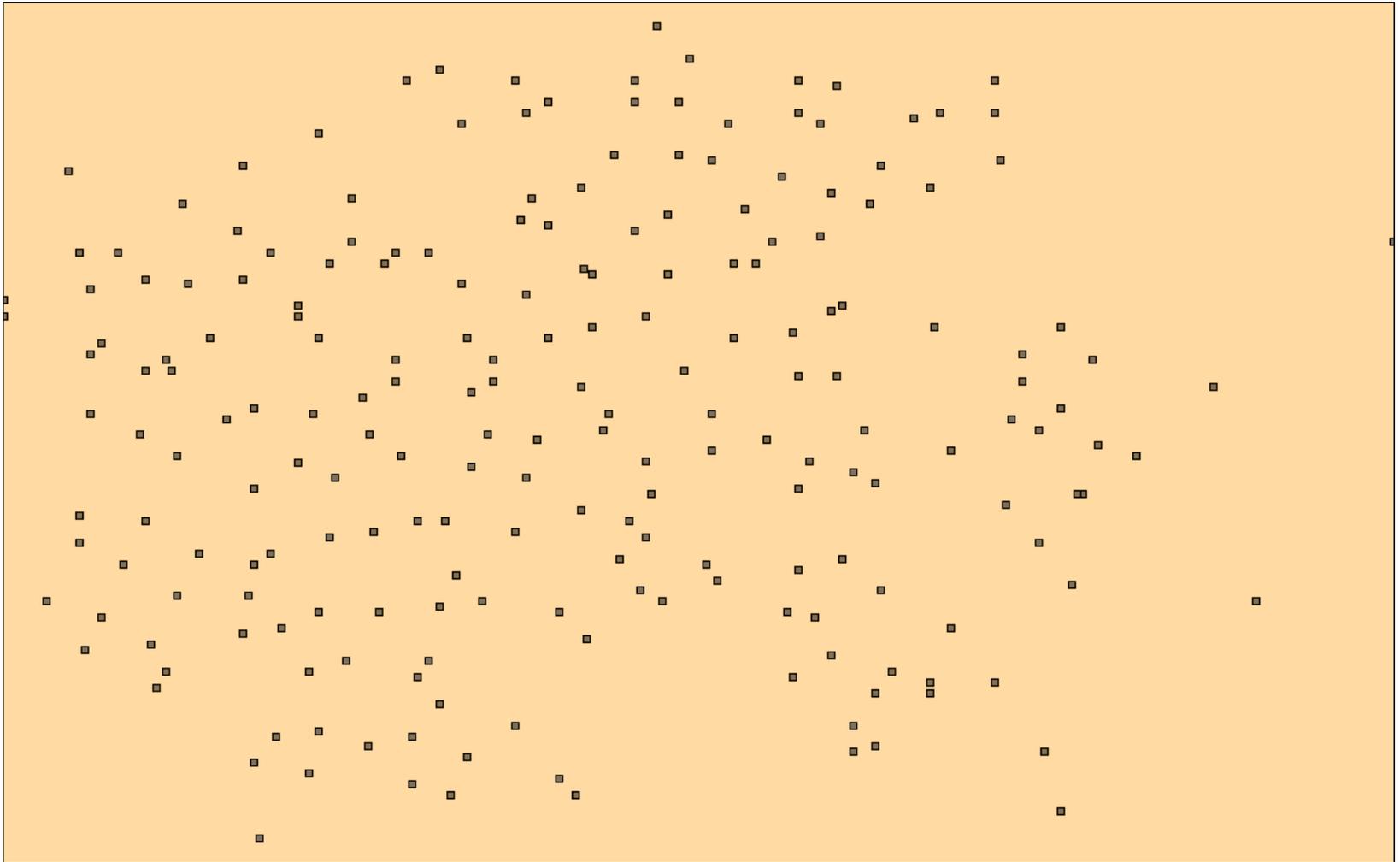
Objets spatiaux : Centre ville de Dallas (Texas)

- **Points** : écoles publiques
- **Lignes** : rues
- **Polygones** : unités territoriales (census block group boundary)

# Analyse de Données Spatiales

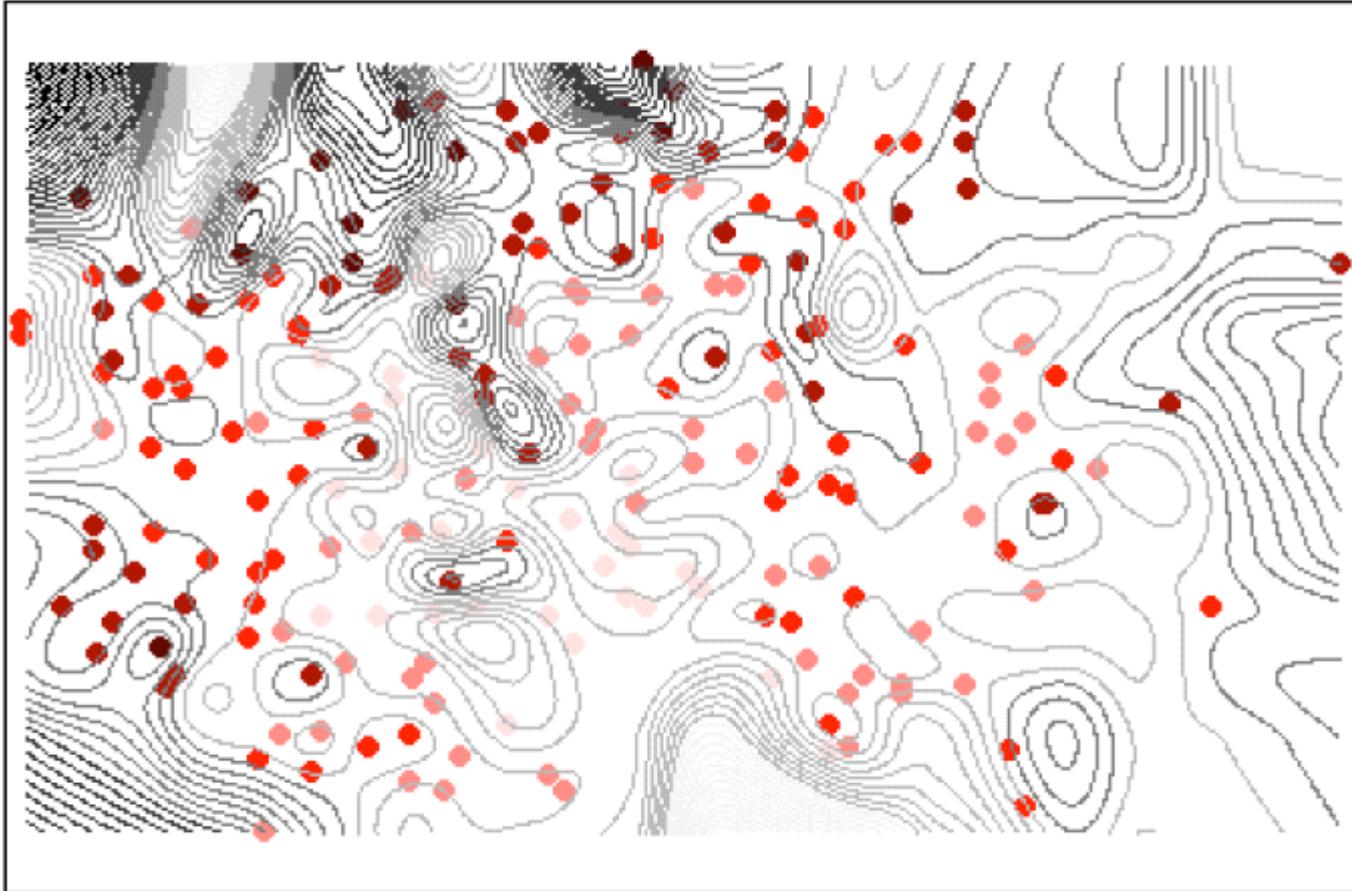
- ▶ Analyse de processus ponctuels :
  - points comme évènements : accidents, crimes (localisation)
  - schémas de points sur une carte : on s'intéresse uniquement à la localisation
- ▶ Analyse géostatistique :
  - les points représentent un échantillon aléatoire de localisations dans une surface continue
  - moniteurs de pollution, ventes de maison
  - modèles pour surfaces continues : on s'intéresse à l'interpolation, on utilise les observations pour modéliser toute la surface
- ▶ **Analyse d'unités territoriales :**
  - **polygones et points (centroïdes), on observe un échantillon exhaustif, le nombre d'objets est fini**
  - **on s'intéresse à l'autocorrélation et l'hétérogénéité spatiale**

# Analyse de processus ponctuels



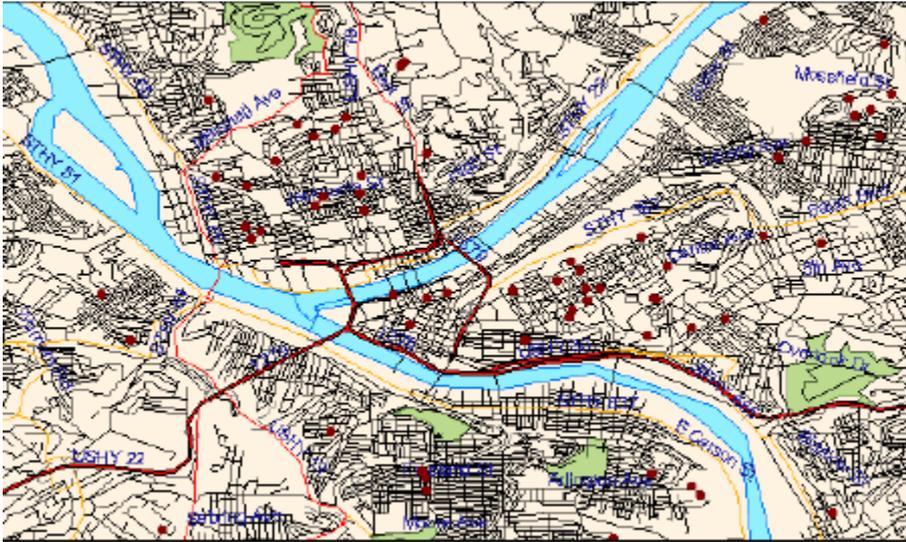
**Points** : 211 transactions immobilières de maisons individuelles dans le comté de Baltimore (1980)

# Analyse géostatistique



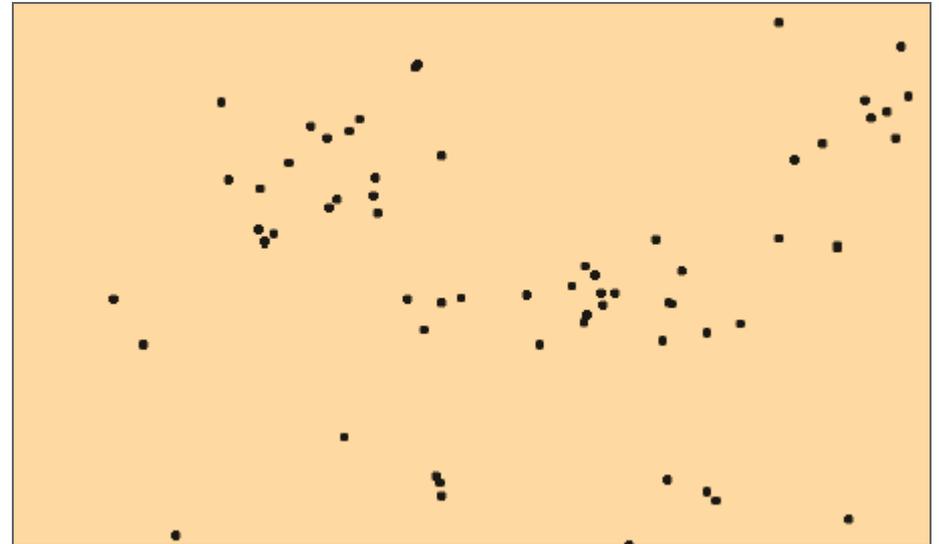
Interpolation géostatistique du prix de vente des maisons individuelles dans le comté de Baltimore (1980)

# Analyse de processus ponctuels

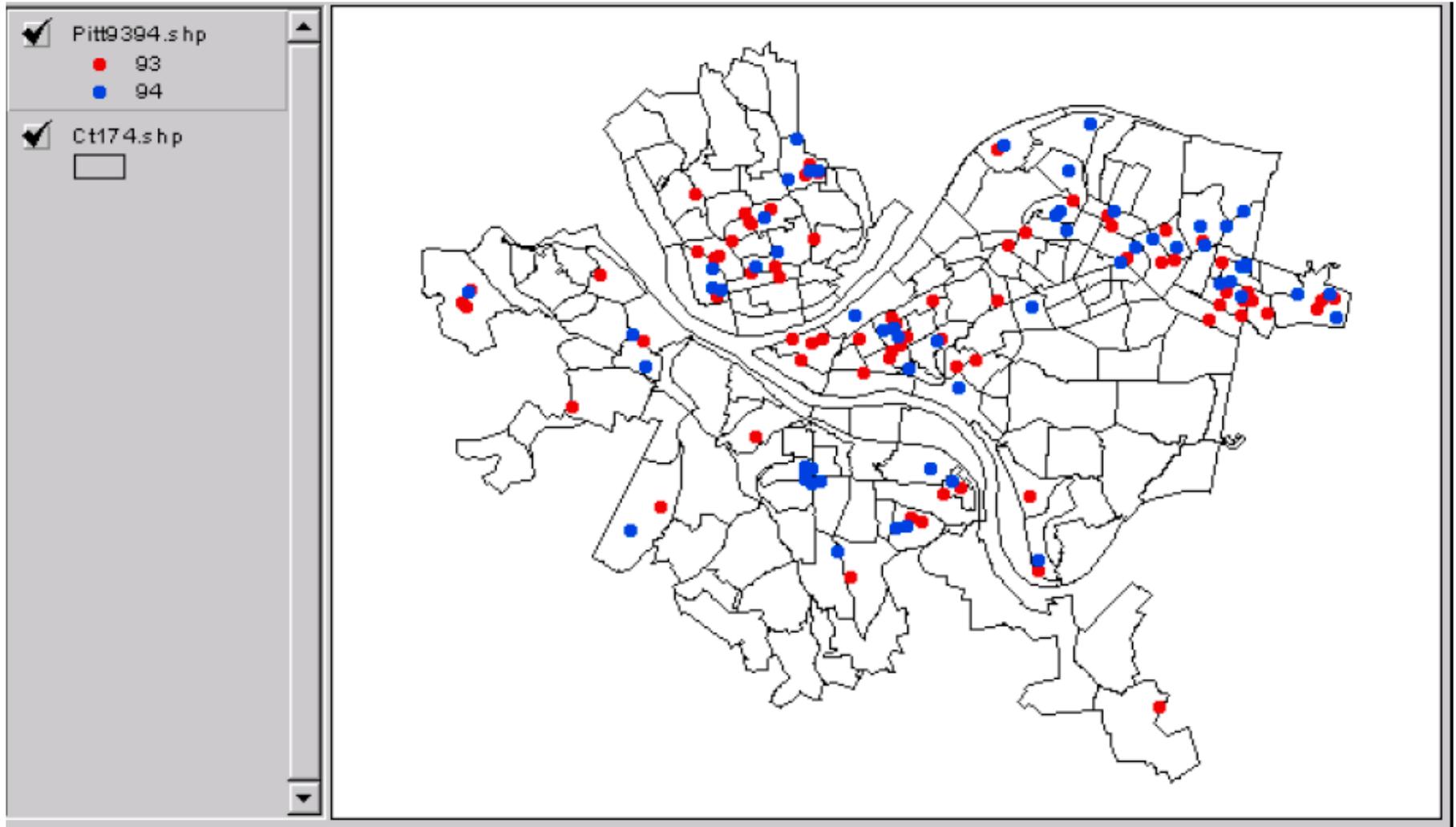


Points dans un contexte géographique  
Ville de Pittsburgh  
(Pennsylvania, USA)

Schéma de points dans le plan

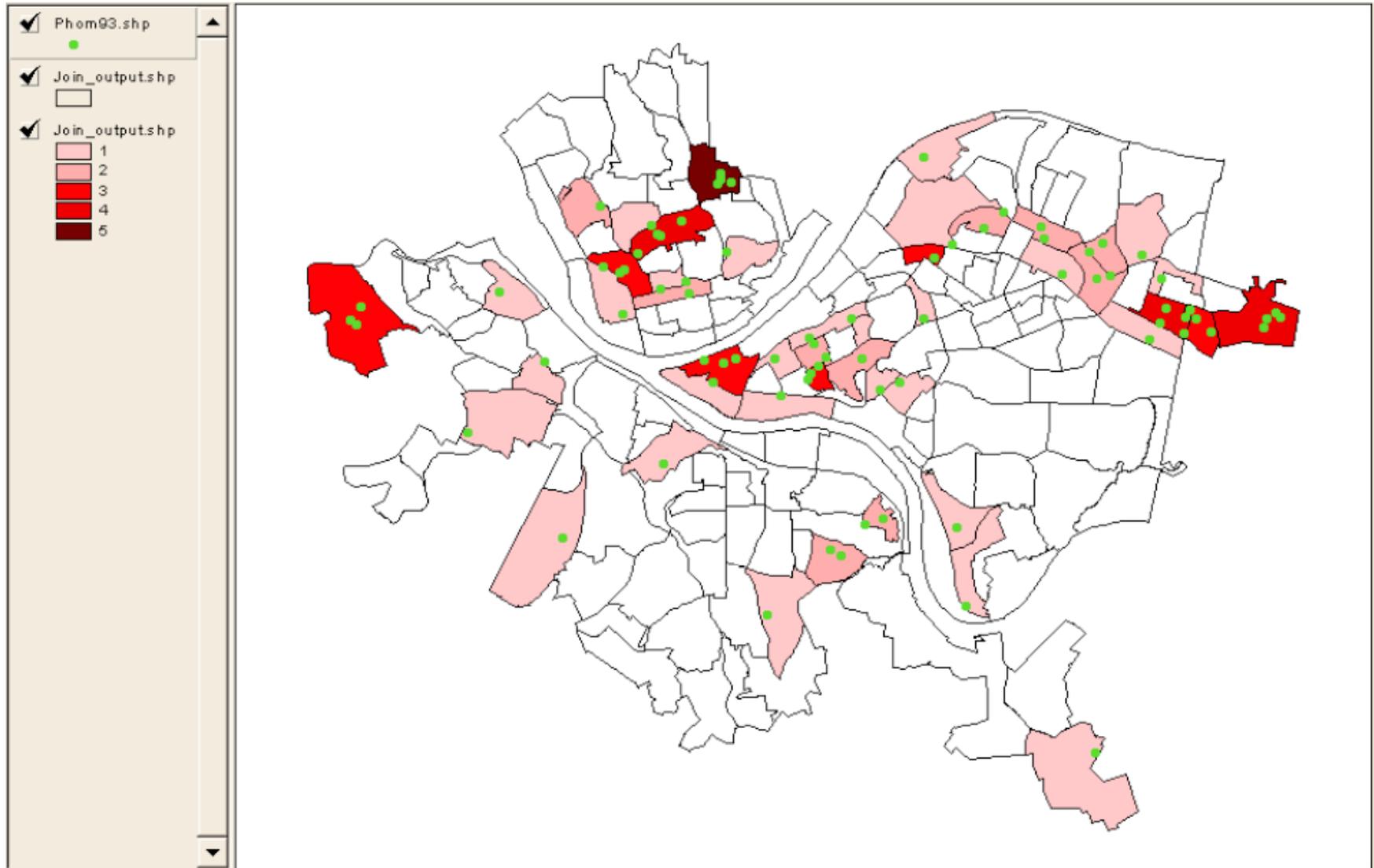


# Analyse d'unités territoriales (1)



Localisations des homicides en 1993-1994 à Pittsburgh (Pennsylvania, USA) : points dans les unités territoriales (source : Anselin, 2003)

# Analyse d'unités territoriales (2)



Localisations des homicides en 1993-1994 à Pittsburgh (Pennsylvania, USA) : affectations des effectifs aux unités territoriales (source : Anselin, 2003)

# Représentation des objets spatiaux discrets

## ► Objets spatiaux discrets

### ■ Polygones

- unités territoriales représentées par leur frontière
- polygones  $\Rightarrow$  points : centroïdes

### ■ Points : coordonnées x,y

- localisations représentées par leurs coordonnées
- points  $\Rightarrow$  polygones : tessellation ou pavage
- Définition : diagramme ou polygones de Voronoï / polygones de Thiessen  
Soit E un espace euclidien, soit S un ensemble fini de n points de E, les éléments de S sont appelés points centraux

Polygone de Voronoï pour un point central p :

$$Vor_s(p) = \{x \in E / \forall q \in S, d(x, p) \leq d(x, q)\}$$

*On appelle polygone de Voronoï associée à un élément p de S l'ensemble des points qui sont plus proches de p que de tout autre point de S*

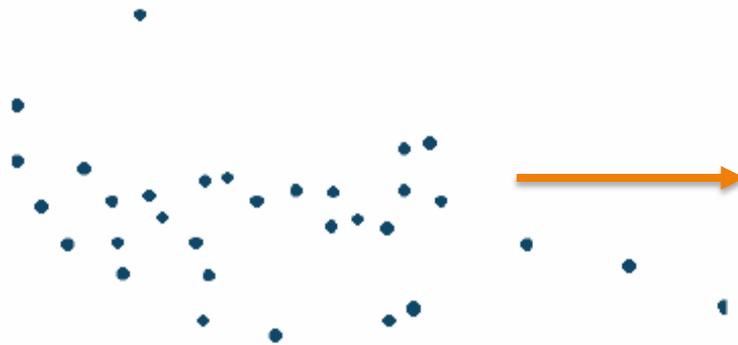
- On affecte à tous les points du polygone la même valeur que la valeur prise par le point central
- zone d'influence, zone de marché, zone d'achalandage à coût de transport minimal
- [http://interstices.info/jcms/c\\_24839/jouez-avec-les-diagrammes-de-voronoi](http://interstices.info/jcms/c_24839/jouez-avec-les-diagrammes-de-voronoi)



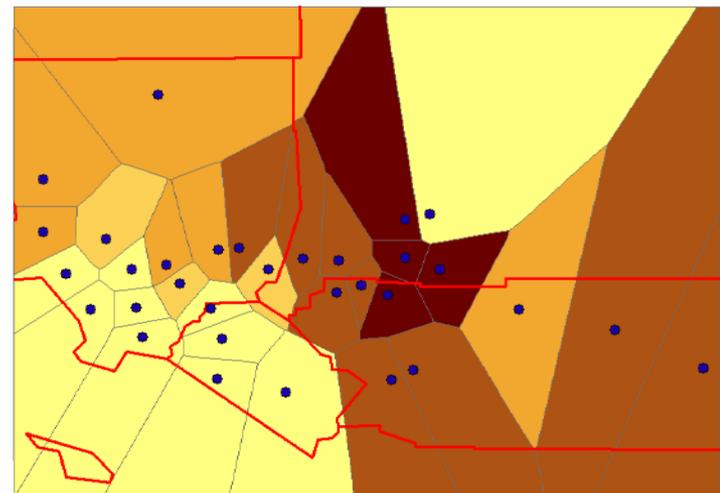
# Polygones vers points : Etats américains vers centroïdes



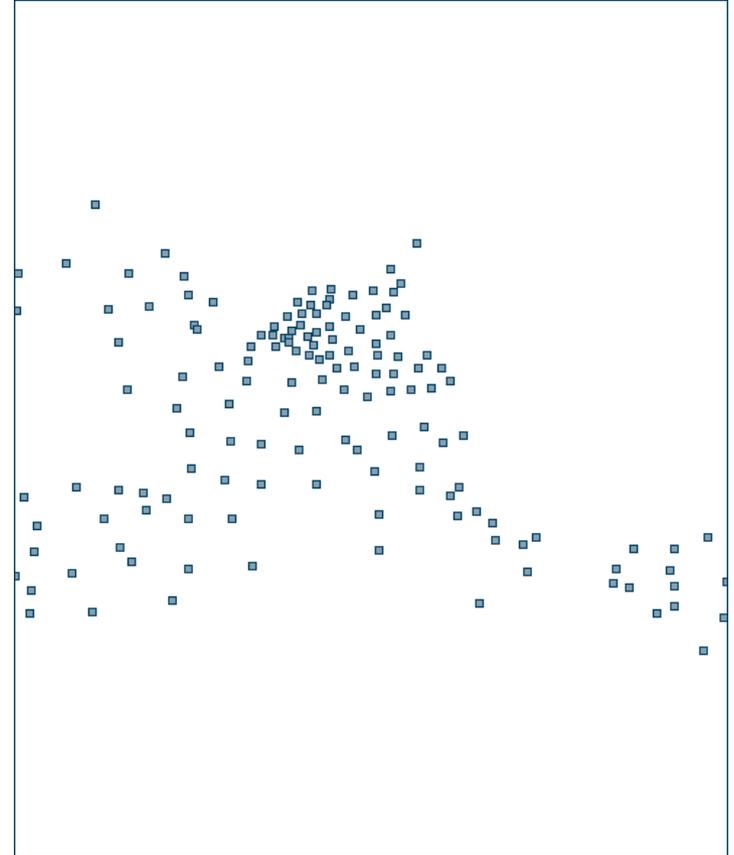
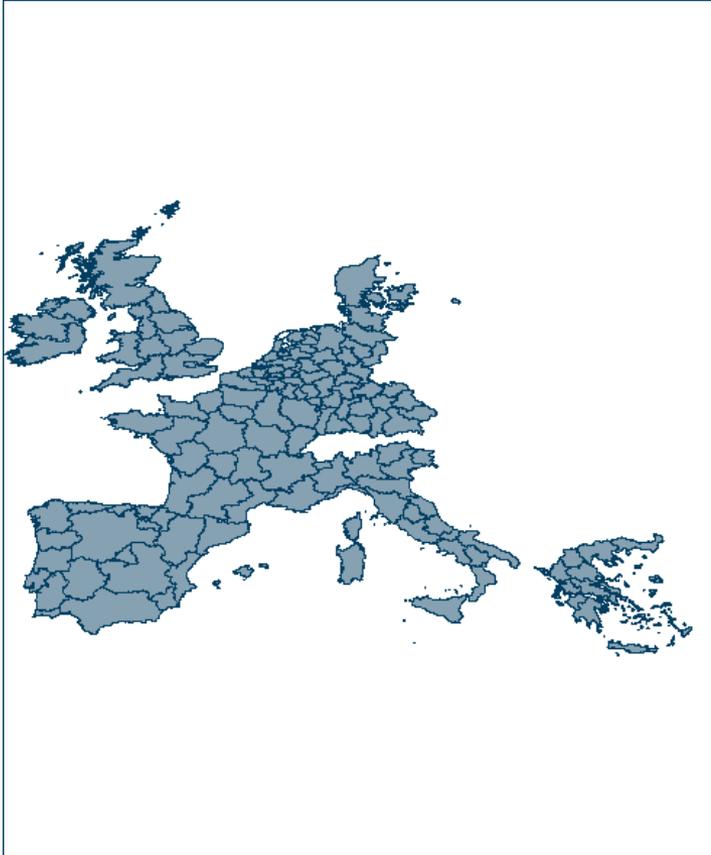
# Points vers polygones : polygones de Thiessen pour le bassin de pollution de Los Angeles



Localisation des capteurs de pollution



# Polygones vers points : 138 régions de l'UE vers centroïdes



# Points vers polygones : polygones de Thiessen pour les 138 régions de L'UE

