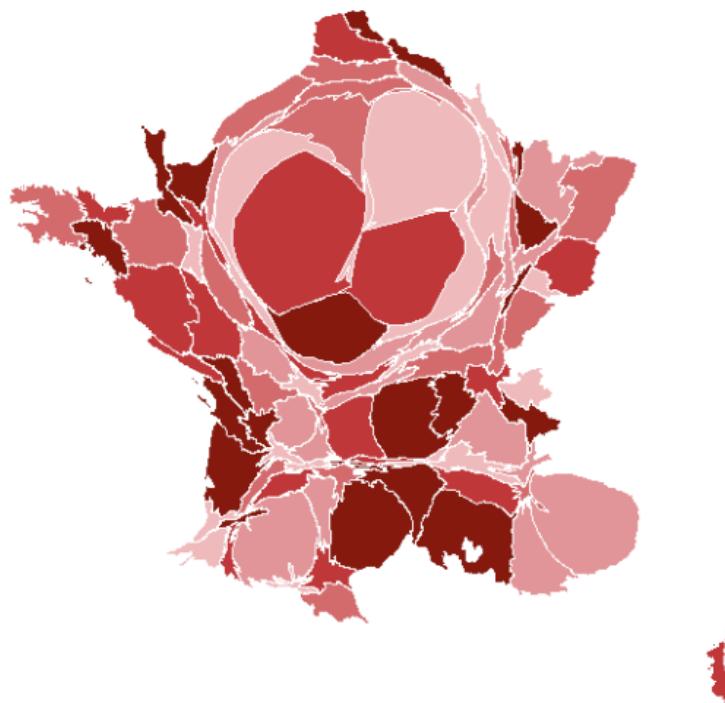


Analyse Spatiale - DeSIGeo

ANALYSE STATISTIQUE DES CRIMES EN FRANCE



Groupe :

- David Delord
- Antoine Duruflé
- Sébastien Quinot
- Pascal Vuylsteke

Introduction

Travaux Pratiques d'analyses spatiales réalisées dans le cadre du cours homonyme.

L'ensemble des données et des résultats sont mis à disposition sur notre [repository Git](#):

<https://pascalpvk.github.io/AnalyseSpatialeENSG/>

Les données présentées ont été préparées et utilisées pour un précédent cours DeSIGéo, celui de Statistiques. Elles sont utilisées dans ce présent rapport pour générer de nouveaux résultats. Il est possible de [consulter le rapport précédent pour obtenir un contextualisation des sources](#).

Introduction	1
Données pour l'étude	2
Données ponctuelles	2
Données surfaciques	3
AutoCorrélation	3
Total de crimes et délits par habitant :	3
Auto-corrélation spatiale du total de crimes et délits par habitant:	4
Total de crimes et délits par magistrat	5
Auto-corrélation spatiale du total de crimes et délits par magistrat :	5
Résultats obtenus dans r-studio avec la bibliothèque « cartography »	6
Carte des délits constatés en France par département	6
Délits et revenu médian par département en France, 2018	7
Délits et revenu médian par département en France, 2018	7
Délits en France par département	7
Cartogramme de Dorling	8
Délits en France - Cartogramme de Dougenik	8
Les outils natifs de QGis	9
Premières analyses spatiales avec QGIS	9

Données pour l'étude

Les jeux de données exploitées lors de ce TP sont disponibles sous divers formats. Pour gagner en lisibilité, ils ont été classés suivant leur géométrie (surfique ou ponctuelle)

Données ponctuelles

- [Service](#)
- [Tribunal](#)

Données surfaciques

- [Département 1](#)
- [Département 2](#)
- [Tribunaux](#)

Etude des Auto-Corrélations Spatiales

Pour étudier la répartition des crimes et délits sur le territoire, il est intéressant de se demander s'il existe des « pôles » de criminalité avec une certaine cohérence spatiale, ou si la répartition est uniquement locale. Pour cela, on peut chercher à calculer le coefficient d'auto-corrélation spatiale entre différentes variables, et établir une cartographie de l'auto-corrélation (les zones voisines se ressemblent-elles ?).

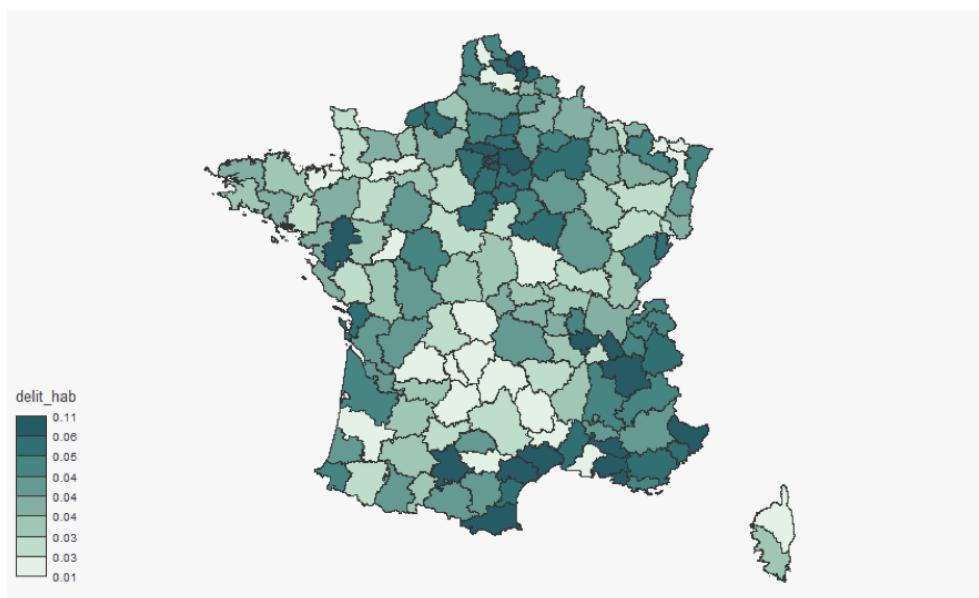
Nous avons réalisé cette étude grâce à R en nous inspirant de ce qui a été fait en cours, grâce

notamment à la fonction « localmoran » (librairie « spdep ») qui permet de définir le coefficient d'auto-corrélation spatiale de chaque entité avec ses voisins.

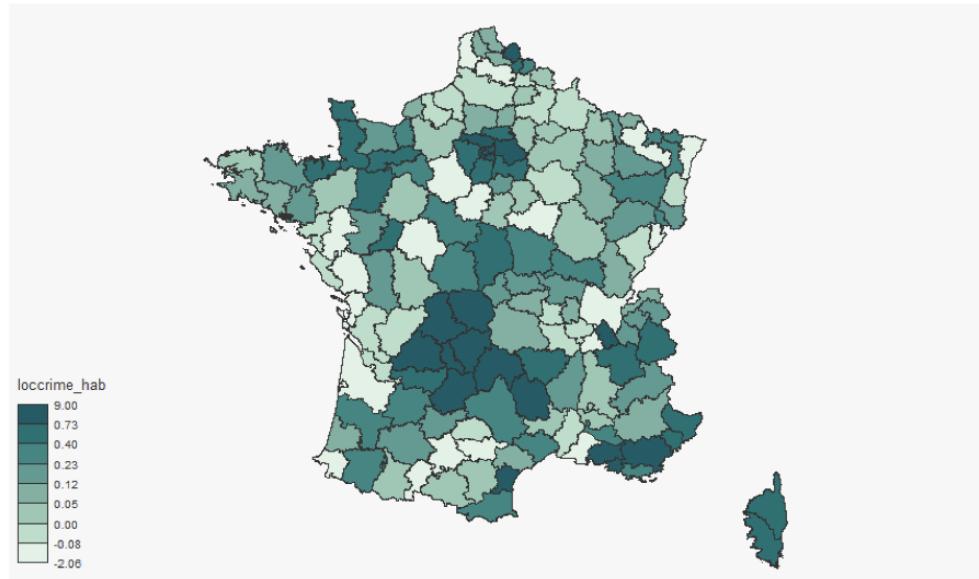
Cette étude a été uniquement faite au niveau des tribunaux car c'est le seul niveau de granularité permettant de travailler avec le nombre de magistrats (on peut par contre faire les crimes par habitants sur les autres niveaux de granularité). Nous avons étudié en particulier 2 variables qui nous ont paru intéressantes : le total de crimes et délits par habitant et le total de crimes et délits par magistrat.

Voici les cartes réalisées : en premier la carte choroplète brute de la variable, et en second la carte choroplète de l'auto-corrélation spatiale pour cette même variable :

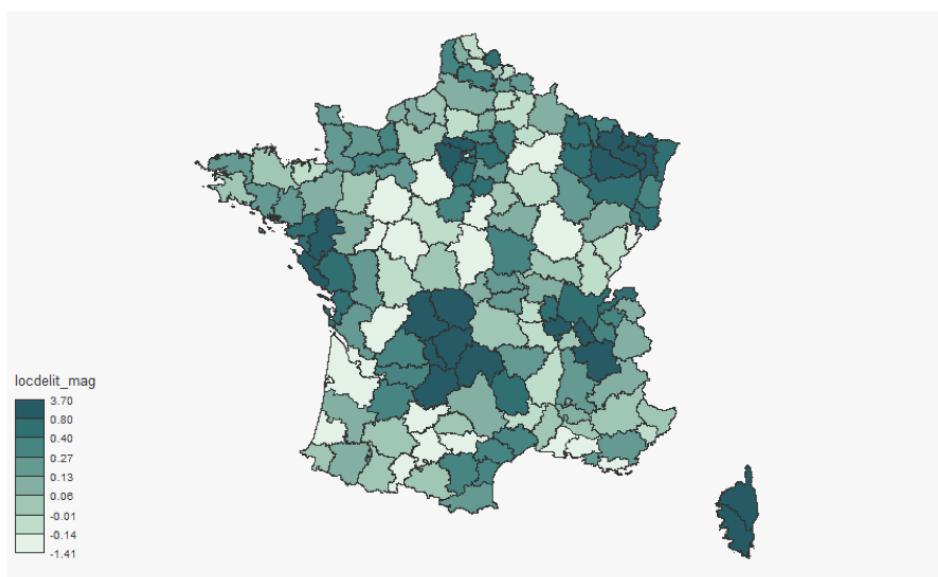
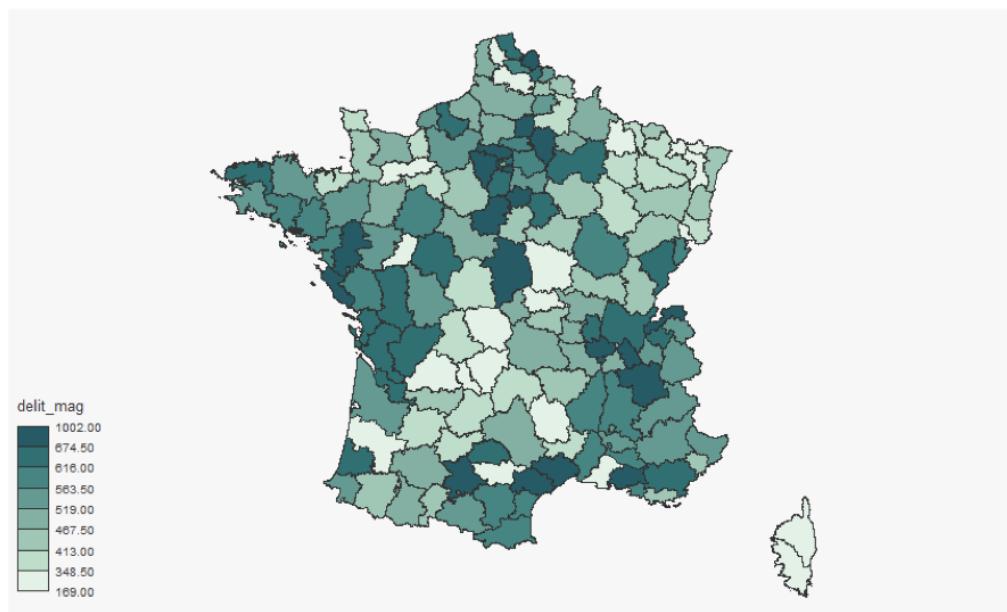
Total de crimes et délits par habitant :



Auto-corrélation spatiale du total de crimes et délits par habitant:



Total de crimes et délits par magistrat



Auto-corrélation spatiale du total de crimes et délits par magistrat :

Les résultats sont intéressants : il existe bien des zones proches entre elles, mais avec des ruptures assez brutales. On voit ici bien la différence entre zones fortement urbanisées et zones fortement rurales. De plus, l'auto-corrélation spatiale des 2 variables est assez différentes : des pôles de stabilité spatiales sont certes en commun (région parisienne, massif central, Normandie), mais les autres pôles

de stabilité (PACA, Loire-Atlantique, Alsace-Lorraine, Lyonnais) ne sont pas partagés.

Resultats obtenus dans r-studio avec la bibliothèque « cartography »

R-studio est un environnement de développement gratuit qui offre de nombreuses possibilités quant aux analyses statistiques et plus généralement mathématiques.

De nombreuses bibliothèques ont ainsi été créées pour compléter ces outils de calculs, dont des outils de représentation des données.

En particulier, la bibliothèque « cartography » permet de créer des représentations cartographiques très simplement et intuitivement. Plus loin, cette outil permet de réaliser des analyses spatiales.

L'objectif de ce document est de présenter brièvement quelques résultats obtenus avec son aide dans le cadre d'une analyse de données concernant les crimes recensés en 2018, en France métropolitaine, par département. Les données utilisées sont disponibles sur opendata.gouv et sur notre GitHub

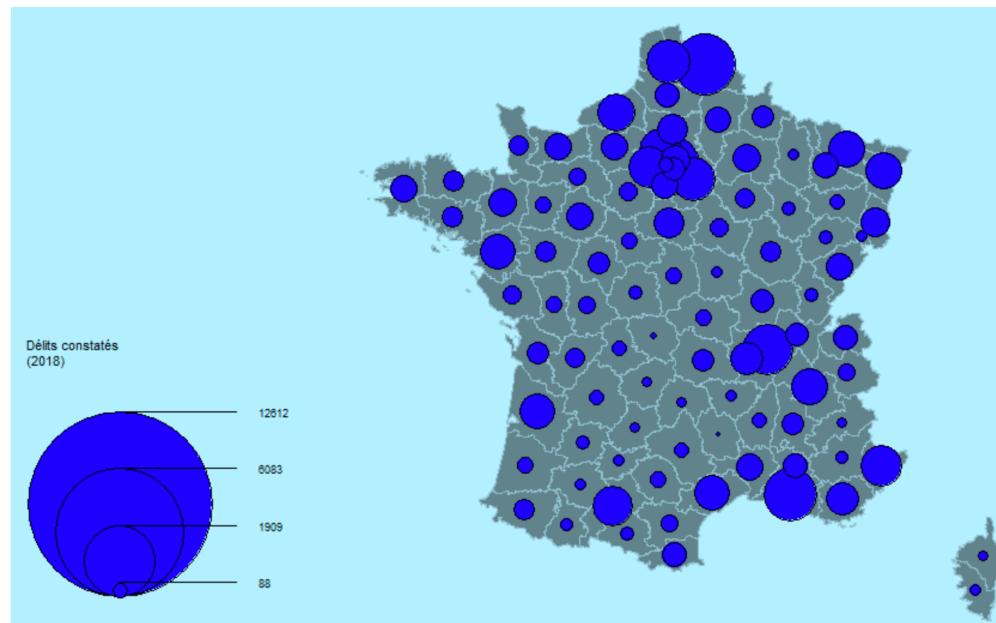


Carte des délits constatés en France par département

En premier lieu, la bibliothèque « cartography » permet de réaliser simplement une carte avec des symboles proportionnels.

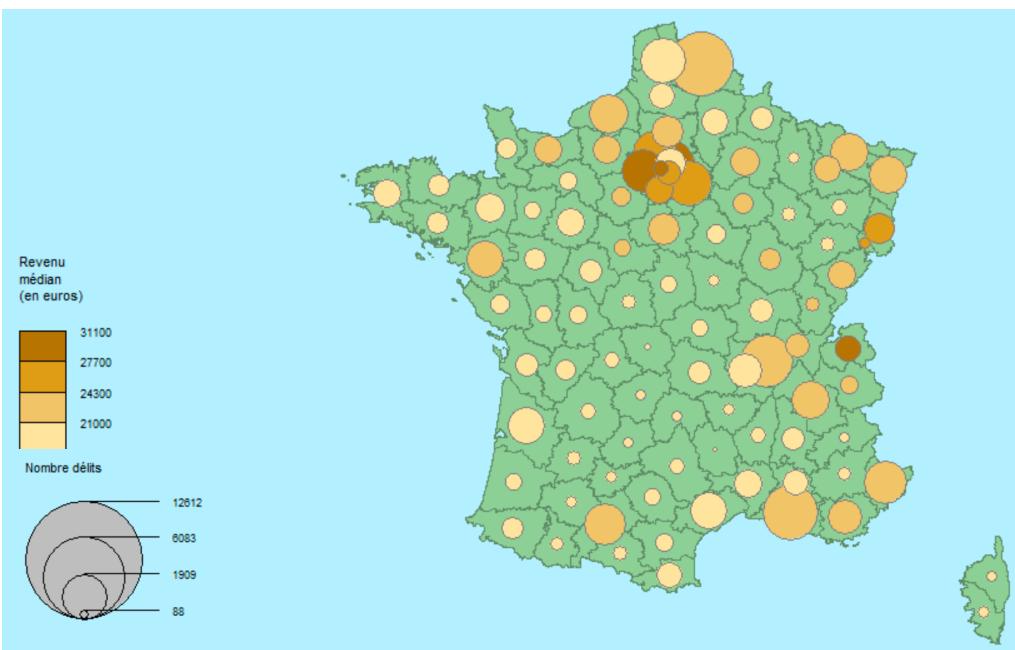
Nous pouvons ainsi obtenir une carte du nombre de délits constatés par département, avec un ponctuel dont la taille est proportionnelle à ce nombre.

Dans notre exemple, nous constatons que les départements incluant les plus grandes villes de France et les plus peuplés sont également ceux qui recensent le plus grand nombre de délits.



On peut dès lors lier le nombre de crimes à la population de la zone étudiée.

Délits et revenu médian par département en France, 2018



L'outil « cartography » permet également de réaliser des cartes comportant des symboles ponctuels de tailles et couleurs variables. Les seuils sont choisis par l'utilisateur.

Ainsi, dans l'exemple proposé ici nous avons choisi de lier la taille du symbole au nombre de crimes par département et la couleur représente la catégorie du revenu médian moyen de ceux-ci.

On peut ainsi constater que le nombre de crimes est plus élevé dans les départements où le revenu médian est, lui aussi, parmi les plus élevés.

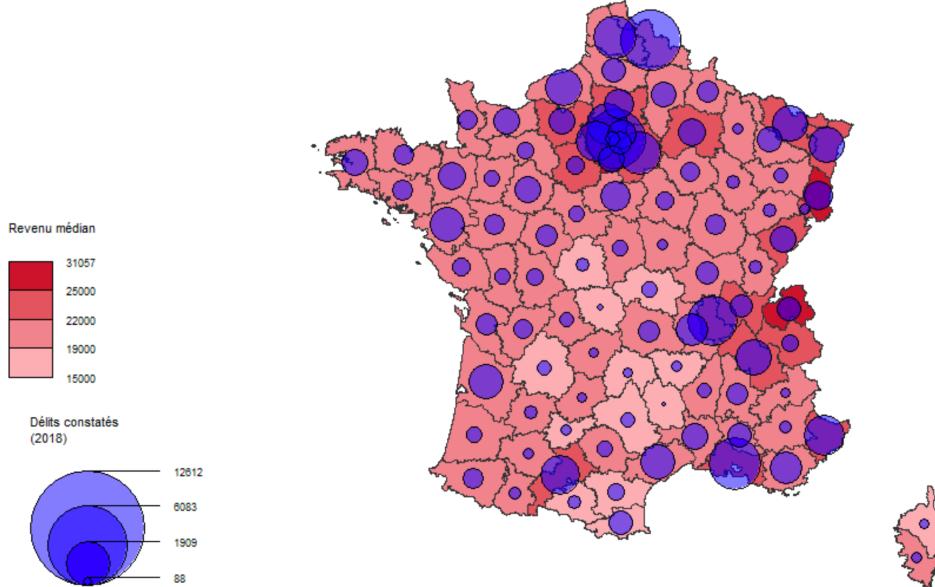
Délits et revenu médian par département en France, 2018

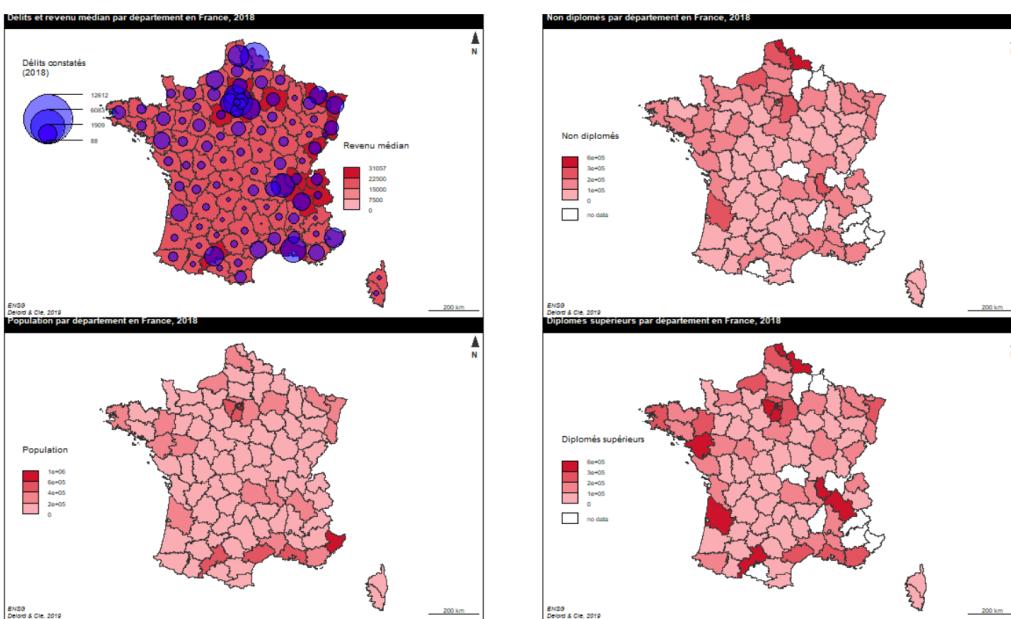
Une autre possibilité offerte par cette bibliothèque est de lier une représentation choroplète et un symbole ponctuel, avec taille proportionnelle, sur une même carte.

Nous avons ici superposer les données liées au nombre de délits (ponctuels bleus, avec niveau de transparence) et les données liées au revenu médian (carte choroplète rouge).

Le résultat obtenu est ici identique à celui vu le paragraphe plus haut (ce sont les mêmes données, représentées différemment).

Nota : cette représentation peut toutefois souffrir de la superposition des variables qui en rend l'exploitation parfois plus difficile.





semblent être liés spatialement.

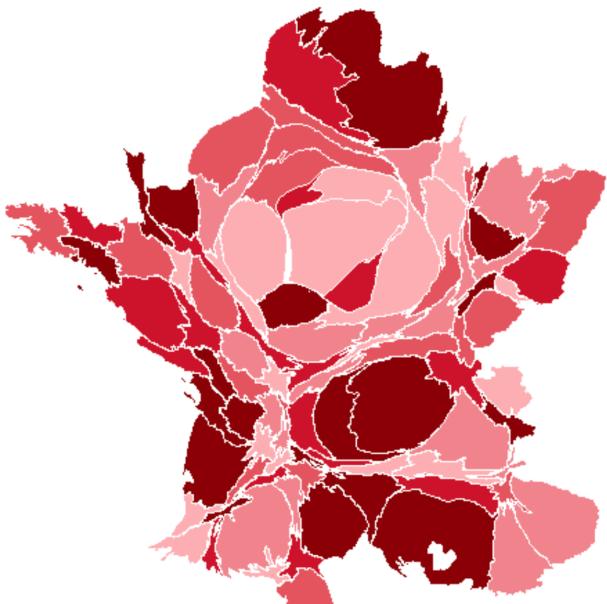
Un intérêt particulier de l'outil peut être la mise en regard de plusieurs cartes, comme illustré ci-dessus.

En effet, cette représentation permet d'illustrer d'éventuelles relations entre les différentes données disponibles.

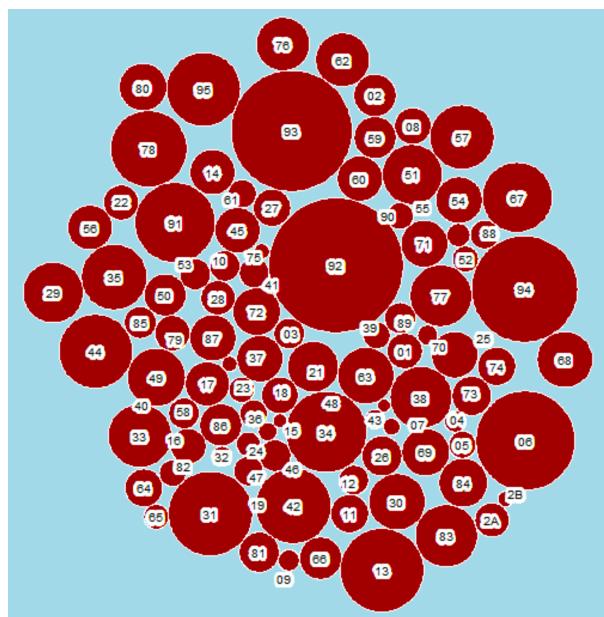
Dans l'exemple proposé, nous pouvons voir que le nombre de délits par département peut être rapproché du revenu médian. De même, le nombre de diplômés et de non diplômés par département

La bibliothèque « cartography » permet également d'effectuer des cartes d'analyses plus extravagantes, comme les cartogrammes de Dorling ou de Dougenik.

Ces représentations à la croisée de la carte et du diagramme permettent de mettre l'accent sur l'analyse mais en plaçant la spatialité à un second plan, et vis-à-vis duquel on se permet une certaine latitude.



**Délits en France par département
Cartogramme de Dorling
(à droite)**



Les exemples proposés ici permettent d'illustrer la répartition des crimes en France, par département. On pourra ainsi noter une surreprésentation de la région parisienne et des grandes villes de province

**Délits en France - Cartogramme de Dougenik
(à gauche)**

En conclusion, nous pouvons noter l'intérêt de la bibliothèque qui permet de visualiser simplement et efficacement des données et effectuer leur analyse.

En particulier et au travers des quelques exemples de réalisation effectués, nous avons pu obtenir un aperçu assez précis de la répartition géographique, à l'échelle départementale, des délits en France métropolitaine. Toutefois, les autres données ne présentaient que parfois quelques similitudes spatiales, et ainsi aucune corrélation franche n'a pu être établie entre ces données et les délits, à l'échelle départementale. On notera toutefois que les données exploitées sont représentatives de l'organisation administrative et centralisée de la France, avec un bassin majeur en région parisienne et plusieurs satellites en province.

Les outils natifs de QGis

Les résultats présentés ont principalement vocation à illustrer quelques-uns des outils natifs de QGIS utiles dans le cadre d'une analyse spatiale.

Premières analyses spatiales avec QGIS

Il existe de nombreux outils dédiés à l'analyse spatiale et de nombreux autres permettant d'en réaliser.

Le SIG libre QGis permet ainsi de réaliser de premières analyses spatiales, en s'appuyant notamment sur la représentation des données détenues.

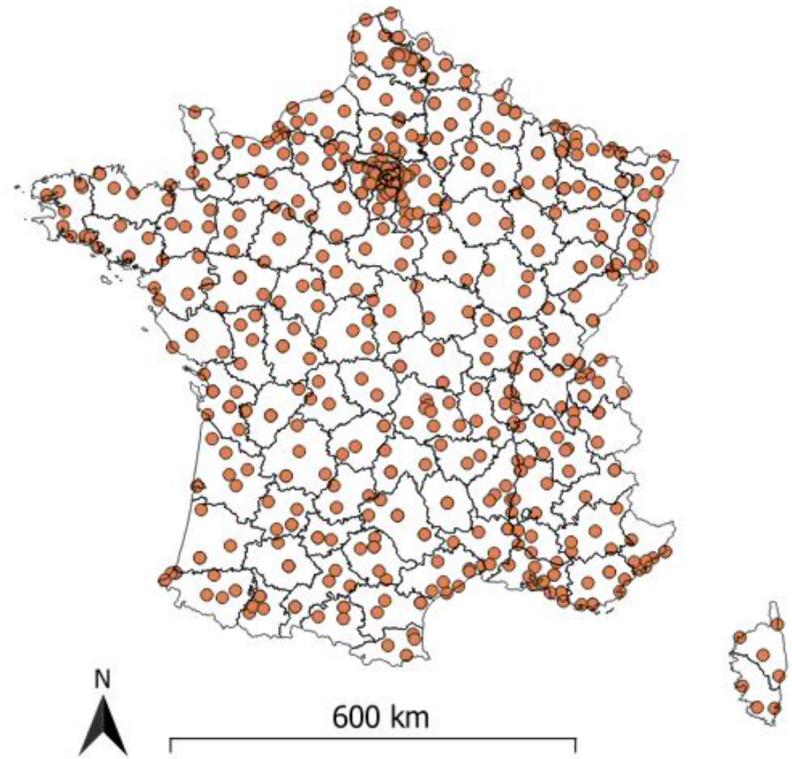
Les résultats obtenus permettent ainsi de déceler certains phénomènes spatiaux et conduire ainsi à une exploitation pertinente des données étudiées. Nous nous proposons, ici, de présenter quelques résultats obtenus avec QGis sur un set de données concernant les données sur les crimes de 2017 et des données de l'INSEE, toutes disponibles librement sur Internet (et en particulier sur notre GitHub).

Analyse spatiale : carte des services

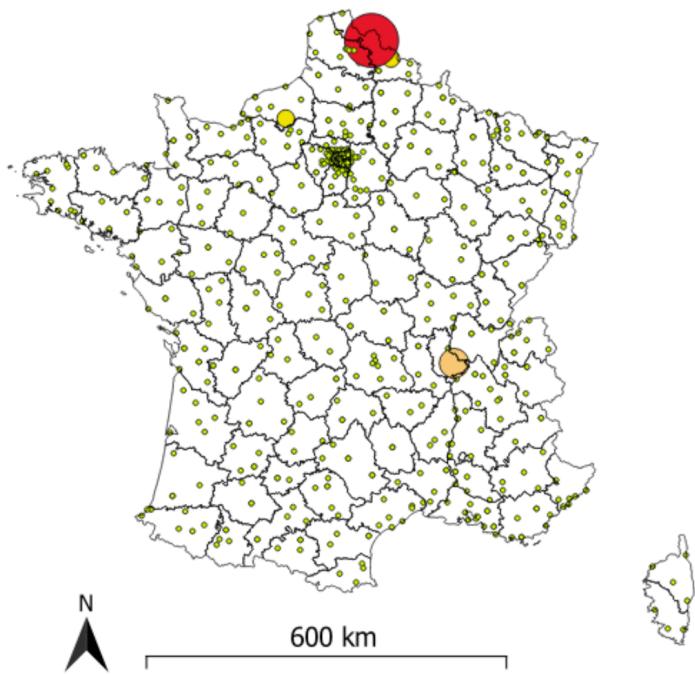
Utilisation d'une simple représentation des ponctuels.

Cette méthode permet d'établir en première approche la répartition géographique de données, qui constitue en soi une première analyse des données.

Dans notre exemple, il s'agit ici de représenter la localisation des services de police et de gendarmeries, pour lesquels des données concernant les crimes sont disponibles. On notera la concentration significative à proximité de les plus grandes villes de France métropolitaine.



Analyse spatiale : carte de caractérisation des crimes par services (gradué)



Utilisation d'une représentation de symboles classés et proportionnés (ici avec des seuils égaux, générés automatiquement).

Cette représentation, encore simple permet de réaliser des analyses de la répartition d'un phénomène.

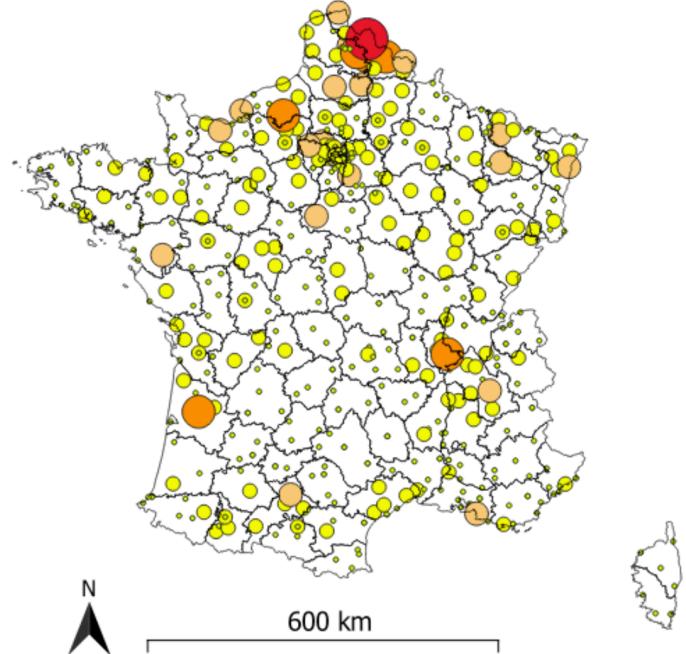
Dans l'exemple, il s'agit de la représentation du total des crimes enregistrés par service (police ou gendarmerie) en 2017. On note que cette méthode souffre des valeurs extrêmes et qu'il faudrait définir des seuils personnels pour s'en affranchir. On notera un nombre très élevé de crimes à Lille, Lyon et Rouen. On peut aussi identifier une concentration élevée de crimes en région parisienne, mais répartie sur de nombreux services.

Utilisation d'une représentation de symboles classés et proportionnés (les classes sont ici générées automatiquement avec la méthode Jenks).

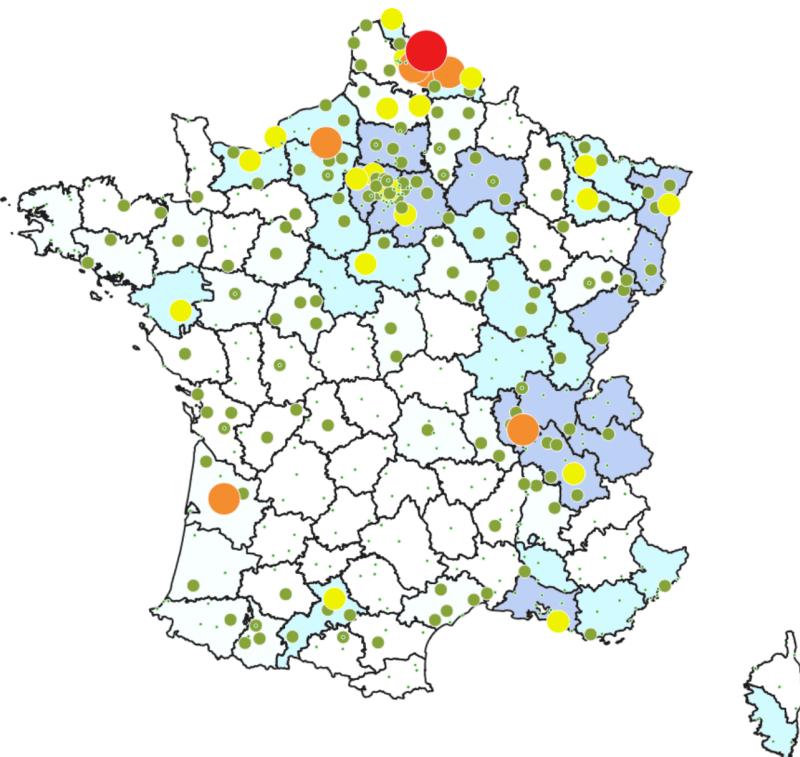
Cette méthode permet d'uniformiser la représentation des données. Le résultat apparaît plus pertinent mais il faut être attentif aux seuils identifiés par l'outil : en effet, la comparaison avec la carte précédente (pourtant les mêmes données) montre que les résultats sont très différents. L'interprétation pourrait alors prêter à confusion. Il est nécessaire d'être attentif aux seuils automatiquement créés ici.

On notera ici la concentration des crimes dans les principales villes françaises, assez logiquement puisque recensant plus de population.

Analyse spatiale : carte de caractérisation des crimes par services (Jenks)



Analyse spatiale : carte de chaleur des crimes par services avec revenu médian par département



Utilisation d'une représentation de symboles classés et proportionnés (les classes sont ici générées automatiquement avec la méthode Jenks) et superposition avec une carte choroplète représentant le revenu médian par département.

L'analyse permise par la superposition de couches permettrait, visuellement, de montrer l'absence de lien entre le revenu médian et le nombres de crimes.

Utilisation de carte de chaleur.

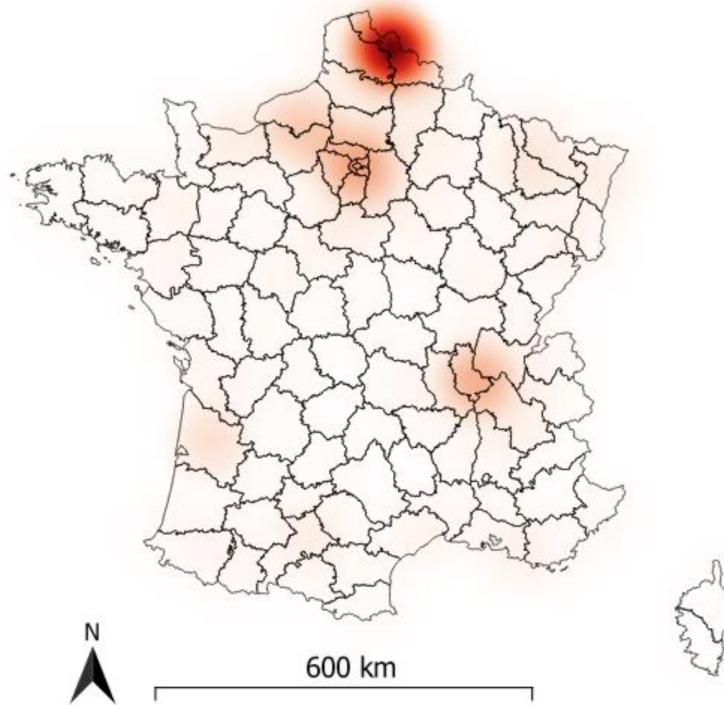
Cette représentation permet de dégager les tendances d'un phénomène, représenté par des données chiffrées, sous la forme de Kernels. Dans notre exemple, nous identifions ainsi les zones les plus « chaudes » concernant les crimes : les lieux où il y en a la plus grande densité : ici, Lille.

Utilisation de carte de chaleur et carte choroplète.

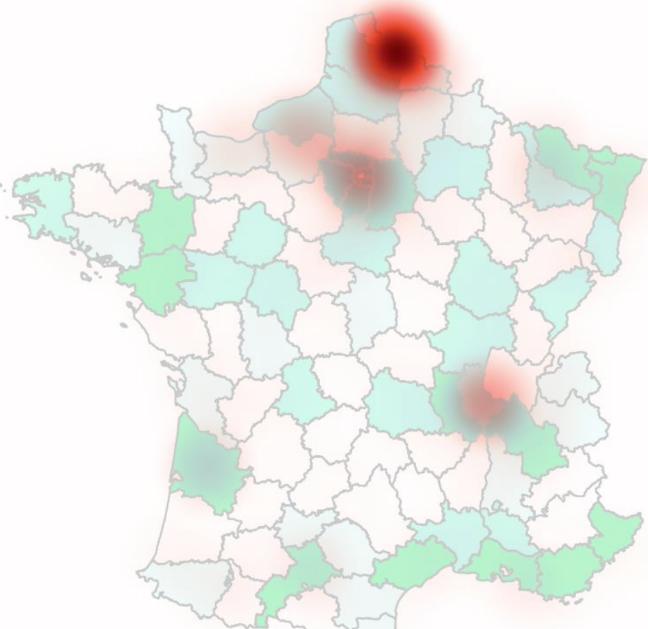
Cette représentation permet de mettre en évidence (ou pas) d'éventuelles corrélations entre les données à disposition.

Dans cet exemple, on pourra constater que la population d'une zone (carte choroplète) ne semble pas être un facteur clé du nombre de crimes recensés.

Analyse spatiale : carte de chaleur des crimes par services



Analyse spatiale : carte de chaleur des crimes par services avec population totale départementale



favoriser la compréhension et l'analyse de données géo-référencées.

Ces quelques représentations nous ont permis de mesurer le potentiel de QGis dans le cadre d'étude spatiale, et ceci sans faire appel à des plug-in supplémentaire. En effet, la richesse de l'outil de représentation, allié aux outils de traitement (non vus ici), permettent d'imaginer de nombreuses possibilités d'analyse des données vectorielles ou raster. De plus, nous avons pu prendre conscience de certaines difficultés liées à l'analyse spatiale, comme la qualité de la représentation choisie qui impacte la compréhension de celle-ci.

Dans le cadre de l'exploitation des données liées au crime, nous n'avons pas pu obtenir de conclusions très pertinentes. La principale a été que les données semblent être incomplète, au vu de la centralité de Lille dans l'ensemble des analyses réalisée et la discrépance d'autres villes pourtant majeures. Toutefois, nous avons pu confirmer la sagacité de l'emploi de l'outil pour