

24 DÉCEMBRE 2013 PAR BRICE

Carte des codes postaux reconstituée grâce à Service-Public.fr

30/12/13 16h30 : Ajout d'un calque représentant les codes postaux dans OpenStreetMap pour comparer.

Un jeu de données libre, complet et à jour des codes postaux n'existe pas vraiment actuellement. L'utilité d'un tel fichier ne saute pas aux yeux, et pourtant, ce besoin se fait sentir régulièrement et particulièrement pour la cartographie. À chaque fois que j'ai besoin de géolocaliser des informations au niveau d'une ville, j'utilise un jeu de données qui contient les noms de villes, les codes INSEE, les numéros et les noms de départements que j'indexe dans un Solr (moteur de recherche libre) d'une manière un peu particulière afin de limiter le taux d'erreur, pour l'interroger ensuite avec les informations déterminantes de chacune des lignes du jeu de données à traiter.

Le préalable pour utiliser cette technique est que les noms de villes et les numéros ou noms de départements soient présents dans le jeu de données à géolocaliser (cela m'a servi pour la réserve parlementaire et certains décrets par exemple), le résultat est que les informations ressortent ensuite associées à leur code INSEE. S'il y a plusieurs résultats, l'outil tente une égalité sur les noms de villes en les normalisant et indique la liste complète des couples codes INSEE/noms de villes possibles en cas d'échec. Même si cela fonctionne plutôt bien, cela peut donc demander pas mal de vérifications manuelles sur des jeux de données volumineux et/ou anciens.

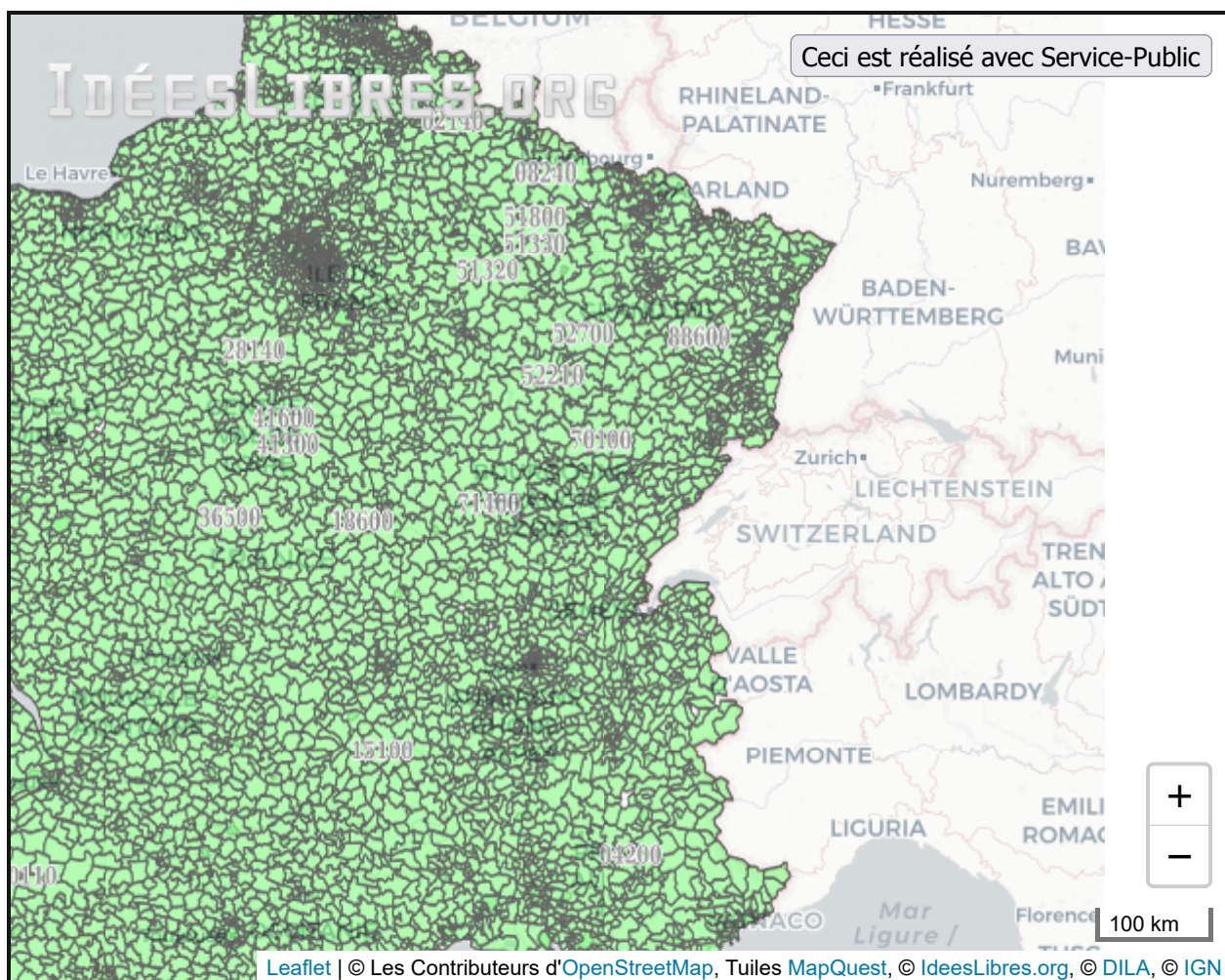
Pour les marchés publics (voir section « L'information à posteriori »), l'information code postal des prestataires est présente mais pas toujours avec le nom de leur ville.

Pour réaliser la même opération, il faut donc disposer des correspondances codes INSEE/codes postaux. Ce besoin s'étant fait sentir également (à des fins de vérification) sur la liste de discussion française d'OpenStreetMap, je me suis donc amusé à regarder ce qu'on pouvait reconstituer via les codes postaux des adresses des organismes présents sur Service-Public.fr (fichier mis à jour pour l'occasion) et les codes INSEE liés en excluant les CEDEX.

On y trouve 6 383 codes postaux concernant 37 356 communes (France métropolitaine et outre mer). Ce premier résultat n'est pas mauvais mais il est incomplet et comporte quelques erreurs provenant des codes INSEE des villes sièges associés à certains organismes (exemple: une adresse utilise 21800 en code postal et est référencée par le code INSEE de Dijon), il paraît néanmoins possible de détecter ces cas avec le fichier des correspondances ci dessous et ce sera ma prochaine étape. Ce qu'il y a de bien avec l'Open Data c'est que chacun peut améliorer, ne vous gênez donc pas si cela vous intéresse (histoire de ne pas faire le même travail chacun dans son coin, je veux bien être tenu au courant par mail ou twitter, les commentaires étant désactivés sur ce blog pour cause de spam et de changement de moteur de blog imminent).

La source officielle pour ce genre d'information est la Poste (avouez que vous vous en doutiez), mais ils ne fournissent pas encore ce fichier librement. Le seul moyen de vérification officiel fourni est donc un formulaire sur leur site, avec des CGU peu propices au scraping. Vivement que la Poste nous permette d'arrêter ces acrobaties en libérant plus de choses...

Voir la carte en plein écran



Les zones entourées par un trait noir sur fond vert représentent l'agrégation des communes du Geofla de l'IGN utilisant le même code postal. Les zones plus vertes sont celles où plusieurs codes postaux sont utilisés (et contiennent certains des cas cités plus haut). Des traits fins et verts représentent les limites communales du Geofla. Les « trous » sont les communes où l'information du code postal est manquante. Il est possible de combler ces manques avec [ce fichier](#) (dont les codes postaux sont issus de Wikipedia). À noter qu'un petit bug d'affichage des labels sur la carte me fait penser qu'il est temps que je mette à jour ma version de Geoserver.

Télécharger les données (licence ODbL | © IdéesLibres.org, DILA 2013, IGN Geofla 2012) :

- [Correspondance codes INSEE, codes postaux, nom, geofla](#)
- Agrégation des codes postaux et des communes du Geofla.

(Tous les polygones sont en EPSG 4326).

Vous souhaitant de bonnes fêtes de fin d'année !

 CSV, JEU DE DONNÉES, OPENDATA

Une pensée sur “Carte des codes postaux reconstituée grâce à Service-Public.fr”

Nicolas Dumoulin

7 JANVIER 2014 À 12 H 52 MIN

J’ai repéré une erreur à Beaumont (limitrophe au sud de Clermont-Ferrand) qui affiche 63541 alors que le CP est 63110.

En regardant dans l’annuaire service-public, j’ai trouvé l’origine de ce code, c’est l’AFPA qui donne une adresse en Cedex, mais le site web de l’AFPA affiche bien des adresses en 63110. Par ailleurs, d’autres organismes, comme la mairie, sont en 63110.

Les commentaires sont fermés.