

# Législatives: faiblesses de l'*Open Data*

Éric Guichard

5 novembre 2022

## 1 Des données pour le citoyen $\lambda$ ?

En travaillant le fichier des coordonnées des circonscriptions législatives fourni par l'Insee<sup>1</sup>, nous découvrons de minuscules polygones dont les abscisses (longitudes) et ordonnées (latitudes) ne dépassent pas les valeurs

$x$  : Max= 4.52571092650666 Min= 4.52571092565532

$y$  : Max= 43.459711691444184 Min= 43.459711691106280

Ce qui fait, pour la 8<sup>e</sup> circonscription du Var, un polygone inscrits en des rectangles de largeur  $85134 * 10^{-14}$  et de hauteur  $337904 * 10^{-15}$ .

Considérons la **hauteur**, facile à calculer : la terre ayant une circonférence de 40 000 km, 90 degrés de latitude correspondent à 10 000/90 km, soit (en considérant que  $\frac{4}{9} \sim 0.5$ ), 11 000 mètres.

La précision de  $10^{-15}$  correspond donc  $1.1 * 10^{-10}$ , soit 1 Ångström (0,1 nanomètre, moins d'un millionième de millimètre).

Notre hauteur vaut donc  $1.1 * 337904$  Ångströms, soit 37 millimètres : **4 cm**.

Pour la **largeur**, il faut prendre en compte le fait que les parallèles de la terre se réduisent en fonction de la latitude : celui du pôle Nord est de longueur nulle, celui de l'équateur vaut 40 000 km.

Arrondissons à 0.7 le coefficient multiplicateur de celui de la France (Valence :  $\frac{1}{\sqrt{2}}$ ). Notre parallèle vaut donc 28 000 km, à diviser par 2 car on fait la distinction entre longitudes positives et négatives et notre précision est de 1.7 Ångströms.

Notre largeur vaut donc  $1.7 * 851340$  Ångströms, soit 148 millimètres : **15 cm**.

## 2 Premier bilan

Nous concevons qu'une circonscription, une ville, un département aient plusieurs composantes connexes. Mais est-il indispensable d'en intégrer certaines sur lesquelles un électeur ne puisse se tenir ? Pour rappel, notre polygone est inférieur à **4 x 15 cm**.

Nous ne comprenons pas à quoi peut servir une telle précision. 1 Ångström, c'est ridiculement petit face à un grain de sable au bord de la mer.

---

1. <https://www.insee.fr/fr/statistiques/6441661?sommaire=6436478&q=circonscriptions>

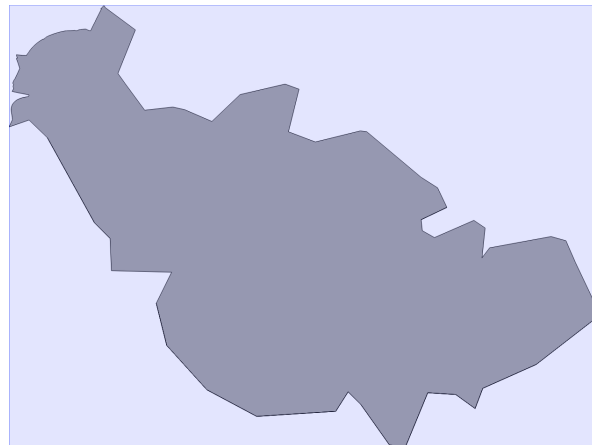
Cette précision *alourdit* aussi le fichier `contours.csv`, qui pèse 20 Mo (et qui contient des structures héritées du format shp, non accessible au premier venu).

### 3 Autres soucis

Nous l'avons compris, pour chaque polygone obtenu, nous devons mesurer la surface du rectangle qui l'inscrit pour savoir s'il mérite ou non d'être conservé.

En pratique, cela ne suffit pas. Certains « polygones » semblent suffisamment conséquents pour être conservés ( $1/5^e$  de la superficie d'une circonscription). Or, après étude, ils semblent constitués de simples lignes qui reviennent sur elles-mêmes, ce qui explique le fait qu'ils ne soient pas coloriables : si A, B, C sont 3 points non alignés du plan, un tel polygone est du type :  $A \rightarrow B \rightarrow C \rightarrow B \rightarrow A$ .

Exemple : considérons le « Multipolygone » 3801, de la première circonscription de l'Isère. Il a deux composantes connexes, même si on n'en voit qu'une. Cf. figure 1.



**Figure 1** – Isère. Les **deux** polygones de la première circonscription.

Supprimons le premier (et gros : sa surface vaut 0.023) polygone. Apparaît alors le second, de surface 0.0056. Cf. figure 2.

#### 3.1 Un étrange polygone

Voici son descriptif en svg :

```
<path id="3801-2" fill="slategray" stroke="black" stroke-width="0.0001"
d="M5.75809956087186 45.2540485811677 L5.75809956087185 45.2540485811677
L5.76120617275715 45.2665396079346 L5.77311335804057 45.2796969299219
L5.78789082319219 45.2875822990424 L5.81133923204524 45.2860679689598
L5.81500420827543 45.2802989493447 L5.81873991079039 45.2840159490299
L5.8272643070418 45.2960724356203 L5.83196808678188 45.2966833390924
L5.83861563389481 45.2805760148441 L5.84689900786945 45.2810953150664
L5.85265090523133 45.2852557021689 L5.85492293322526 45.2792601132337
L5.87077899171187 45.272177298594 L5.88975611449211 45.2574606027419
```



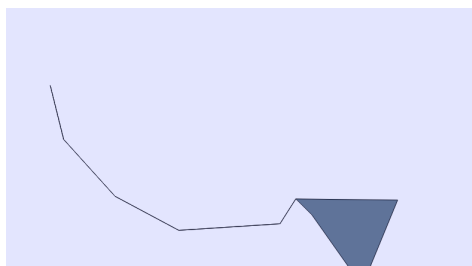
**Figure 2** – Isère. Second polygone de la première circonscription. Étrangement, il n'est pas coloré.

```
L5.88975611449209 45.2574606027419 L5.87077899171187 45.2721772985939
L5.85492293322526 45.2792601132336 L5.85265090523133 45.2852557021688
L5.84689900786945 45.2810953150663 L5.83861563389481 45.280576014844
L5.83861563389479 45.2805760148441 L5.83196808678188 45.2966833390923
L5.8272643070418 45.2960724356202 L5.81873991079043 45.2840159490299
L5.81873991079039 45.2840159490299 L5.81500420827543 45.2802989493447
L5.81500420827542 45.2802989493447 L5.81133923204524 45.2860679689597
L5.78789082319219 45.2875822990423 L5.77311335804057 45.2796969299219
L5.76120617275715 45.2665396079345 L5.75809956087186 45.2540485811677 Z "></path>
<!-- Le polygone précité a une surface de 0.00561314529280037 -->
```

Les points de départ et d'arrivée du polygone sont identiques. Nous ne comprenons donc pas pourquoi il ne se ferme pas. et le Z aurait dû faire l'affaire.

## 3.2 Réduction du polygone

On va supprimer des points du milieu. Une couleur apparaît : figure 3.

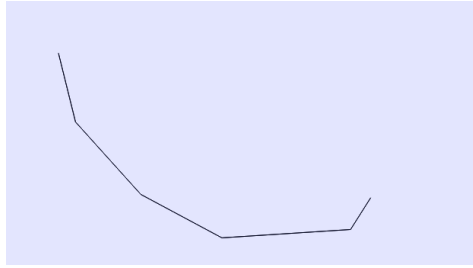


**Figure 3** – Isère. Second polygone de la première circonscription, avec quelques points du milieu supprimés.

On récidive, pour retomber sur une simple ligne : figure 4.

Voici le descriptif svg de cette figure 4.

```
<path id="3801-2" fill=... d="M5.75809956087186 45.2540485811677
L5.75809956087185 45.2540485811677 L5.76120617275715 45.2665396079346
```



**Figure 4** – Isère. Second polygone de la première circonscription, avec d'autres points du milieu supprimés.

```
L5.77311335804057 45.2796969299219 L5.78789082319219 45.2875822990424
L5.81133923204524 45.2860679689598 L5.81500420827543 45.2802989493447
L5.81500420827542 45.2802989493447 L5.81133923204524 45.2860679689597
L5.78789082319219 45.2875822990423 L5.77311335804057 45.2796969299219
L5.76120617275715 45.2665396079345 L5.75809956087186 45.2540485811677 Z "></path>
```

### 3.3 Un polygone devenu chemin aller-retour

On commence à comprendre, avec la figure 5, composée de... 5 points :

```
<path id="3801-2" d="M5.75809956087186 45.2540485811677
L5.75809956087185 45.2540485811677
L5.76120617275715 45.2665396079346
L5.77311335804057 45.2796969299219
L5.75809956087186 45.2540485811677 Z "></path>
```



**Figure 5** – Isère. Second polygone de la première circonscription, avec cinq points conservés.

Il faut plus de 20 heures de travail pour arriver un fond de carte propre. L'*open data* peut se transformer en *dirty open data*, inaccessible aux profanes.

## 4 Mode d'obtention du fichier contours

```
import geopandas as gpd
shapefile = gpd.read_file("circonscriptions_legislatives_030522.shp")
print(shapefile)
shapefile.to_csv('example.csv', index=False)
```