

GIS Analytics: quel potentiel pour les données géographiques ?

Vandy Berten Smals Research 05-05-2022

Smals Research 2022



**Innovation with
new technologies**



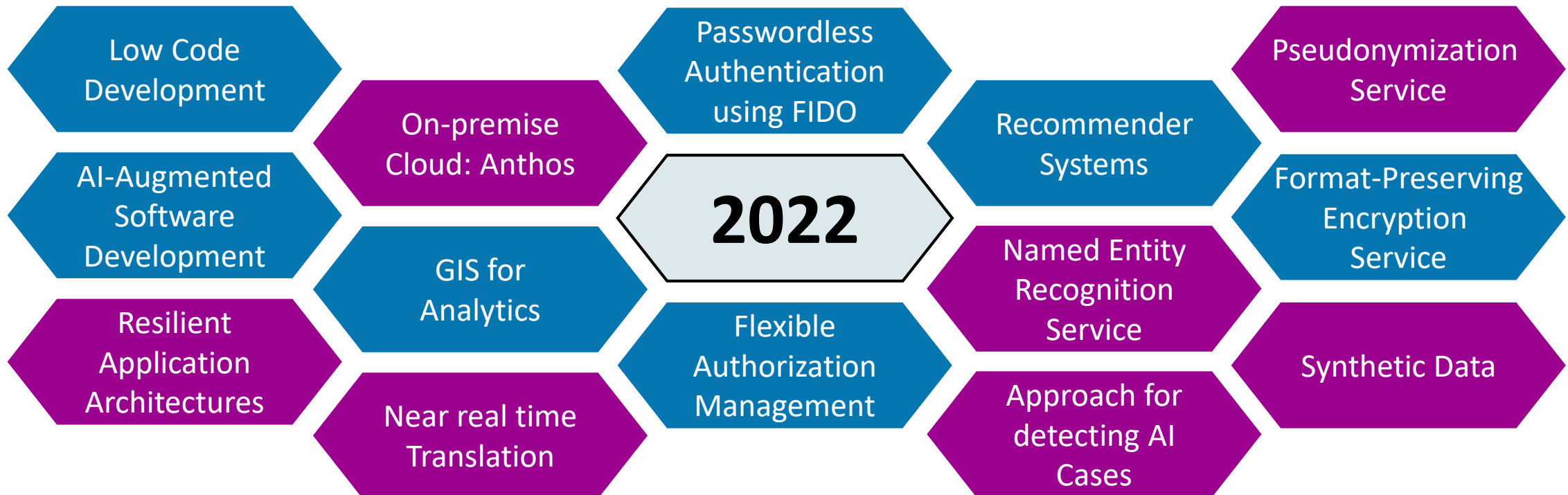
**Consultancy
& expertise**



**Internal & external
knowledge transfer**



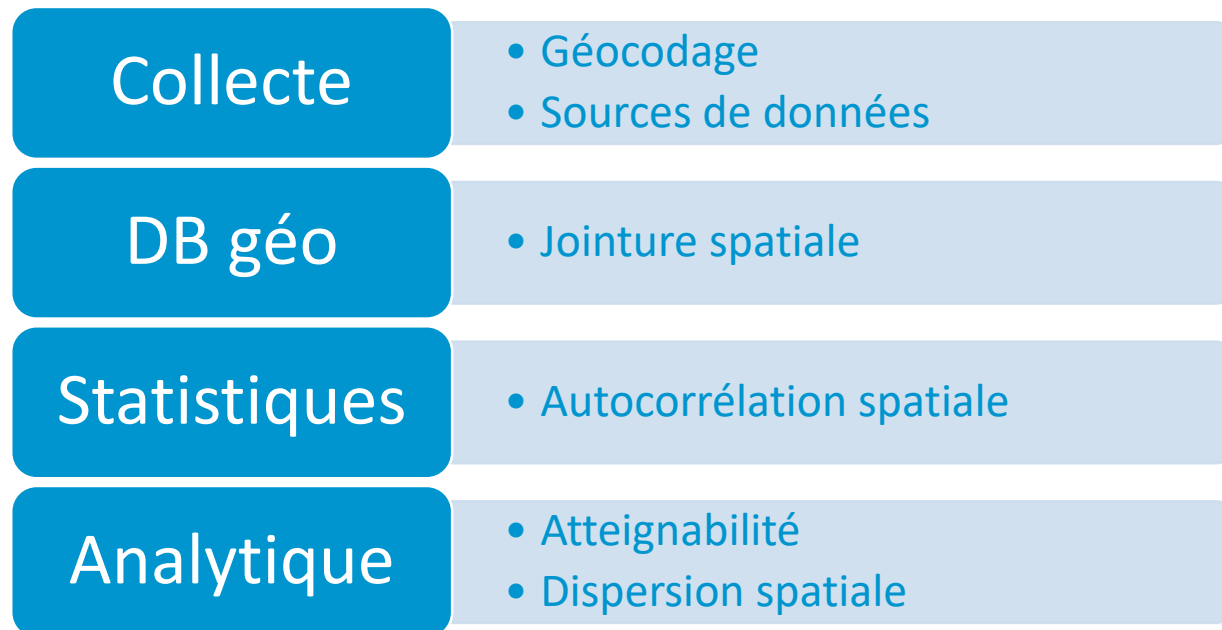
**Support for
going live**



Objectif et contenu

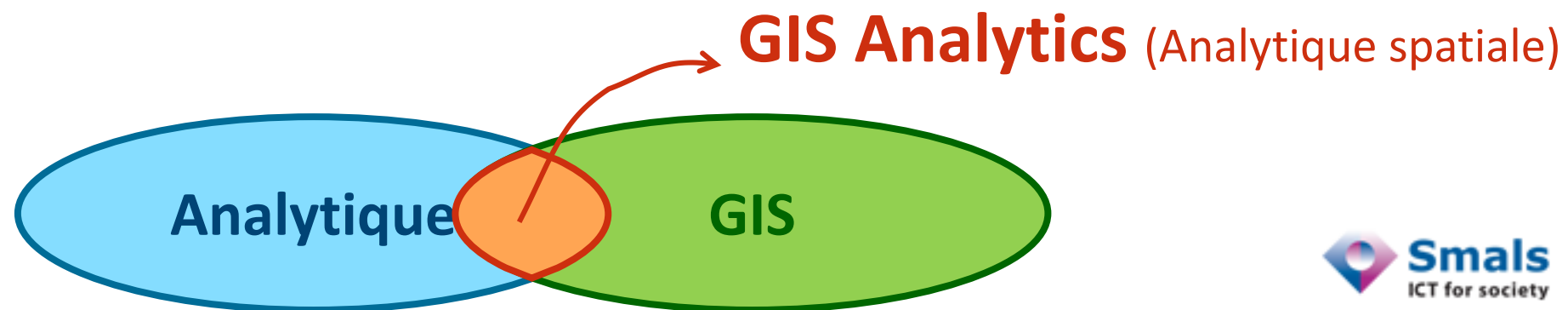
Objectif : Sensibiliser au **potentiel** qu'ont les **données géographiques**

Contenu :



Pas dans le contenu :

- Visualisation
- Cartographie (nature de sols, réseau routier, hydrologie, 3D...)
- Données géographiques complexes
- Modèles épidémiologiques/spatio-temporels
- Optimisation de tournées/itinéraires

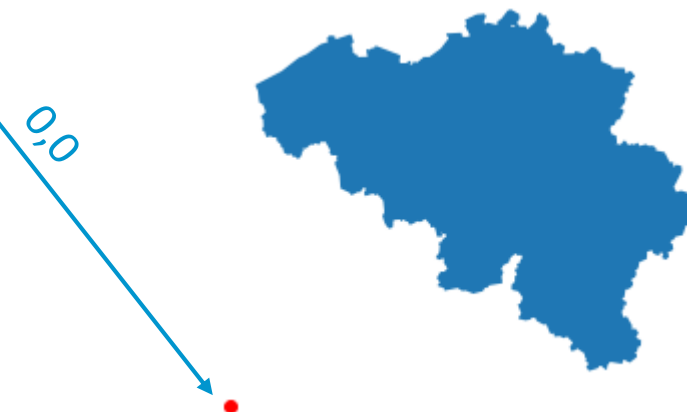


Donnée géographique

Donnée géographique : donnée que l'on peut **placer sur une carte**, directement ou indirectement

Directement :

Degrés (géographique)	50.8361263, 50°50'10.1"N,	4.3382716 4°20'17.8"E	(WGS 84 – EPSG:4326)
Mètres (cartésien)	6592292.39, 169482.70,	482944.07 147859.16	(WGS 84 – EPSG:3857) (Lambert 1972 – EPSG:31370)
	Latitude (Y)	Longitude (X)	CRS (coordinate ref. system)

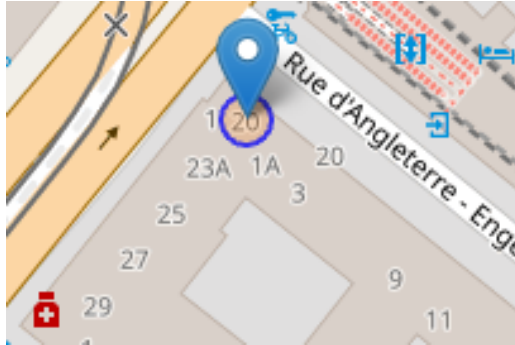


Indirectement :

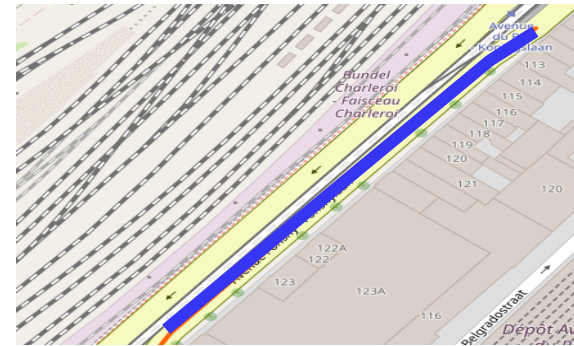
Avenue Fonsny 20, 1060 Saint-Gilles
1060
Saint-Gilles
Smals, Saint-Gilles

Objet spatial/géographique

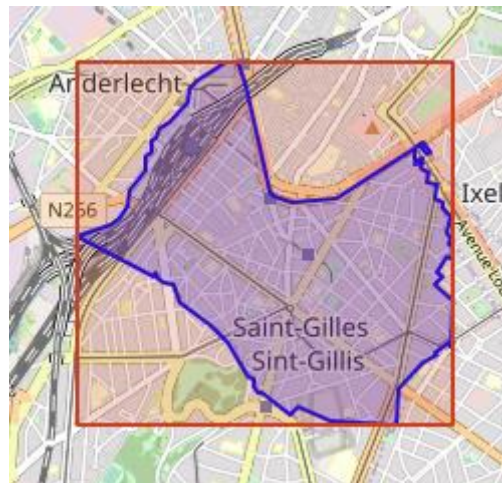
Point : localisation précise



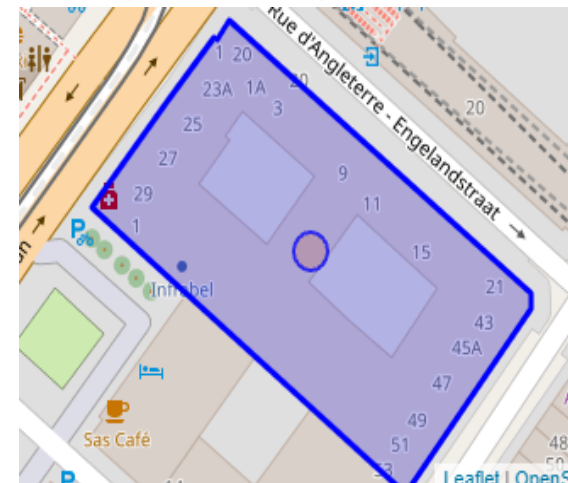
Ligne : segment de route



Rectangle : “bounding box”



Polygone : région, bâtiment





Collecte

- **Géocodage**
- Sources de données

DB géo

- Jointure spatiale

Statistiques

- Autocorrélation spatiale

Analytique

- Atteignabilité
- Dispersion spatiale

Géocodage : concept

Pour être exploitable géographiquement, un adresse doit être “géocodée”

“Av. Fonsny 20, 1060 Bruxelles”



```
{“street” : “Avenue Fonsny”,  
“housenbr” : “20”,  
“zipcode” : “1060”,  
“city” : “Saint-Gilles”,  
“country” : “Belgique”,  
“location” : [50.8358, 4.3361]}
```

Geocoding:

- Visualization
- Controls
- Spatial (GIS) analytics

...7

Address validation:

Data quality:

Standardization

Deduplication

Matching

Auto-completion
(type-ahead)

...

Or polygon (point sequence)

Géocodage : outils

Fournisseurs Cloud commerciaux :

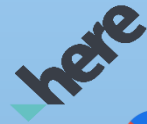
Here we go

Bing maps

Google maps

TomTom

Parfois partiellement gratuit



Google Maps

Principalement une solution open-source/open-data :

OpenStreetMap/Nominatim

Variante (avec ElasticSearch) : Photon

On-premise possible (1 Docker cmd line)
pour une région spécifique



OpenStreetMap

À venir (?) : BestAddress (WS+docker?)















@ Smals :

Trillium (batch/API to come) pour la
Belgique → sans localisation

ESRI

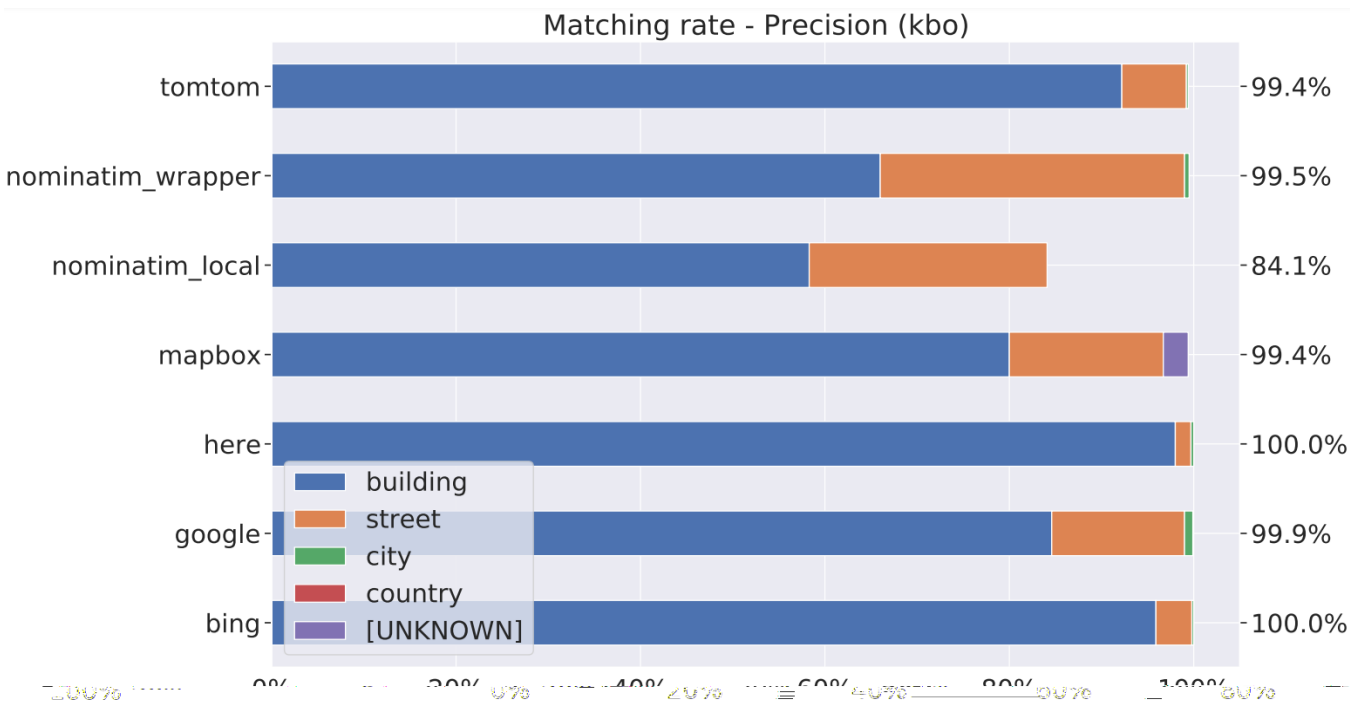
Via IGN/NGI: TomTom on premise

Géocodage : pros/cons

	Cloud solutions	On-premise solutions
Complexité à mettre en place		
Coût infra		
GDPR, Confidentialité		
Coût pour gros volumes (temps, €)		
Accès internet		
Qualité		
Couverture mondiale		

Géocodage : qualité

Dataset : 1,000 adresses (belges) de KBO OpenData



Matching rate : proportion d'adresses avec un match

Précision :

Building : localise le bâtiment

City : localise le centre de la ville

Principales faiblesses de Nominatim/OSM:

Pas très robuste

Souvent « street level »

→ Pas un problème pour la plupart de nos applications

Mais : on-premise !

NominatimWrapper¹ : wrapper « in house » autour d'un Nominatim (OpenStreetMap) on-premise



Collecte

- Géocodage
- Sources de données

DB géo

- Jointure spatiale

Statistiques

- Autocorrélation spatiale

Analytique

- Atteignabilité
- Dispersion spatiale

Sources

Enrichissement de données internes (adresses, coordonnées, noms entités/commerces...), via des **APIs** (souvent gratuitement) :

- Géocodage

- Collecte de POI

- Courbes isochrones

- Itinéraires

- ...

Utilisation de **données authentiques**/Open Sources :

- Frontières : communes, villes, quartiers, pays...

- Hiérarchies d'entités : zipcode > commune > arrondissement > ...

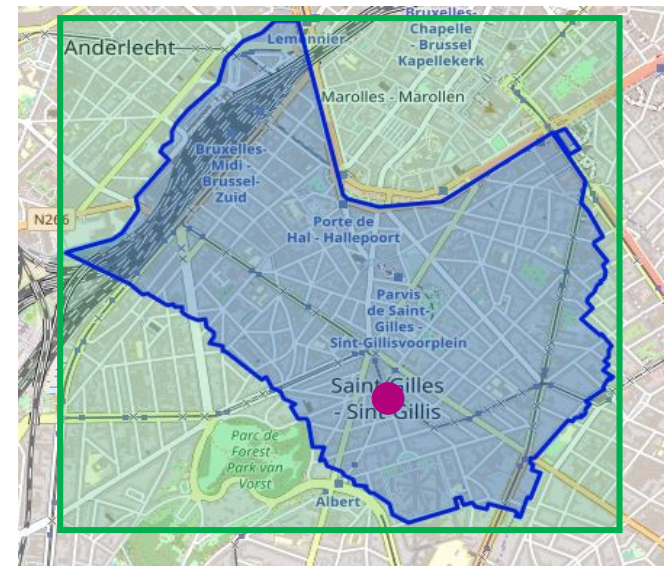
- Indicateurs socio-économiques

API Nominatim (OpenStreetMap)

https://nominatim.openstreetmap.org/search.php?q=1060+saint-gilles&format=jsonv2&addressdetails=1&polygon_geojson=1

```
[{place_id: 281729964,  
  boundingbox:  
  ["50.8200036", "50.8401825", "4.3264847", "4.3588636"],  
  lat: "50.8249958", lon: "4.3454841",  
  display_name: "Saint-Gilles, Brussels-Capital, 1060, Belgium",  
  place_rank: 14,  
  address:  
  {town: "Saint-Gilles",  
   county: "Brussels-Capital",  
   region: "Brussels-Capital",  
   postcode: "1060",  
   country: "Belgium",  
   country_code: "be"  
  }, (...)
```

```
geojson:  
{type: "MultiPolygon",  
 coordinates:  
  [ [ [ [4.326484, 50.830849], [4.330251, 50.82958], [4.33104,  
  50.829126], [4.331237, 50.829222],(...)],
```



API Nominatim (OpenStreetMap)

Base: <https://nominatim.openstreetmap.org/...>

Adresse exacte :

[\(...\)search?q=Avenue+Fonsny+20,+1060+saint-gilles&format=jsonv2&addressdetails=1](...search?q=Avenue+Fonsny+20,+1060+saint-gilles&format=jsonv2&addressdetails=1)

POI (restaurant, Saint-Gilles):

[\(...\)search?q=restaurant,+1060+saint-gilles,+Bruxelles&format=jsonv2&addressdetails=1](...search?q=restaurant,+1060+saint-gilles,+Bruxelles&format=jsonv2&addressdetails=1)

POI (restaurant, par “bounding box”):

[\(...\)search?q=restaurant&viewbox=4.35886,50.84018,4.32648,50.82&bounded=1&format=jsonv2](...search?q=restaurant&viewbox=4.35886,50.84018,4.32648,50.82&bounded=1&format=jsonv2)

Reverse:

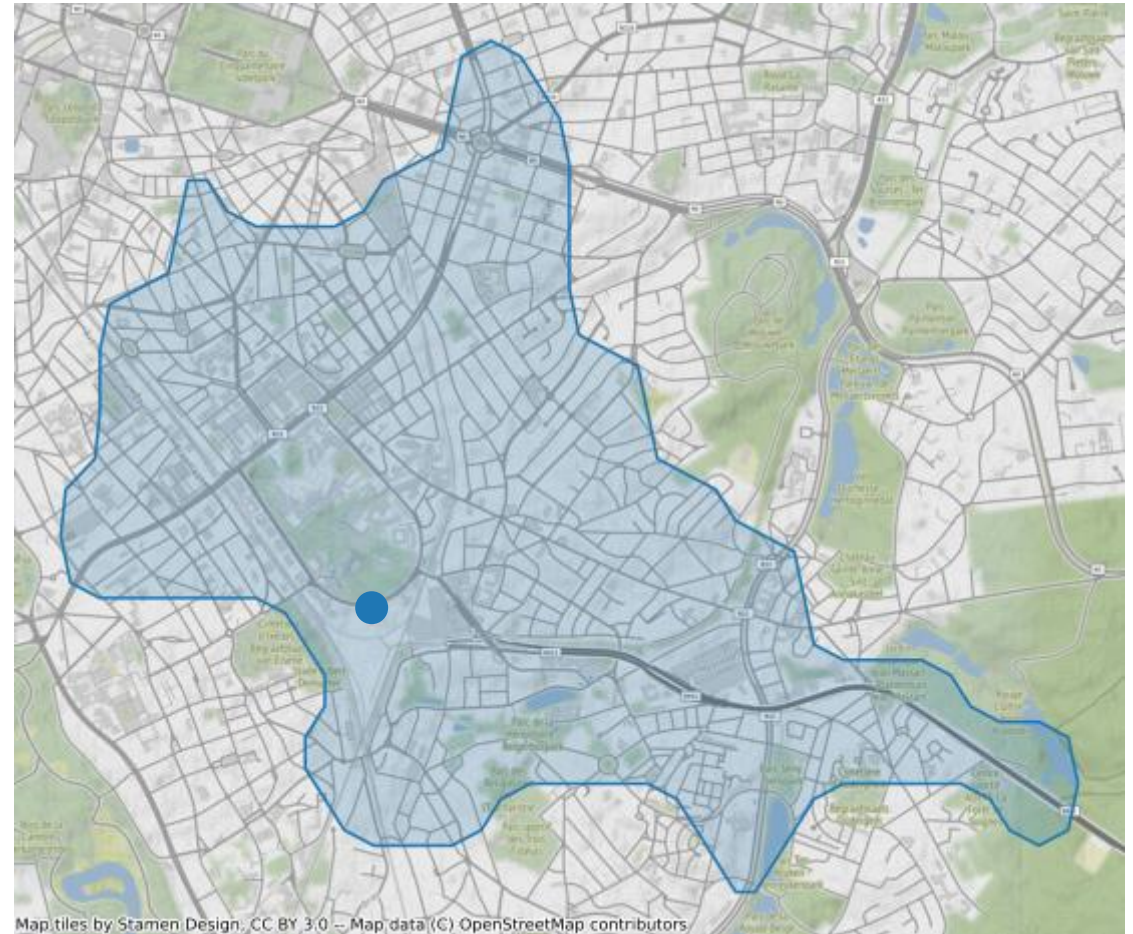
[\(...\)reverse?lat=50.83582&lon=4.33868&format=jsonv2](...reverse?lat=50.83582&lon=4.33868&format=jsonv2)

API : Here (isochrone)

Isoline : ensemble des points accessibles à partir de/vers un point, en un temps (isochrone) ou une distance définie (voiture, camion, piéton)

Résultat : polygône

Exemple : zone accessible en 5 minutes à partir de l'hôpital Delta (Auderghem, Bruxelles)



Sources officielles

Contours des codes postaux (BPost): https://bgu.bpost.be/assets/9738c7c0-5255-11ea-8895-34e12d0f0423_x-shapefile_3812.zip

Contours des secteurs statistiques (StatBel): <https://statbel.fgov.be/fr/open-data/secteurs-statistiques-2020>

Localisation de chaque adresse (BestAddress) : <https://opendata.bosa.be/index.fr.html>

Quartiers Bruxellois : <https://monitoringdesquartiers.brussels/>

Structures des codes postaux/localité/communes/province (BPost): https://www.bpost2.be/zipcodes/files/zipcodes_alpha_fr_new.xls

Indicateur socio-économique par “zone” :

<https://statbel.fgov.be/>

<https://data.belgium.be/>

<https://www.census2011.be/>

Nombreuses données à l'IGN : <https://www.geo.be/>



BestAddress : open data

BestAddress permet de télécharger toutes les adresses belges : <https://opendata.bosa.be/>

Un fichier CSV par région :

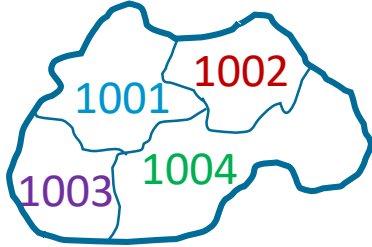
	C	D	E	G	H	J	L	O	P	Q	R	U
1	EPSG:4326	EPSG:4326	address_id	house_number	municipality	municipality	postcode	street_id	streetname	streetname_fr	streetname_nl	
2	50.88995	4.31196	212450	6	21010	Jette	1090	2559		Place Jean-Louis Thys	Jean-Louis Thysplaats	
3	50.82095	4.28556	200187	20	21001	Anderlecht	1070	3959		Square Marie Curie	Marie Curiesquare	
4	50.79599	4.41243	25899	25	21017	Watermaele	1170	4637		Avenue Delleur	Delleurlaan	
5	50.85773	4.30862	19920	31	21012	Molenbeek	1080	3811		Rue Joseph Diongre	Joseph Diongrestraat	

Difficulté : mapping “adresse en input” vs “adresse BestAddress”

À venir : REST API de géocodage (mai 2022?)

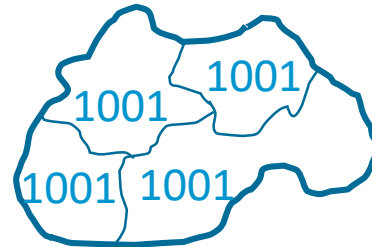
Schéma de division Belge version postale

Schéma postal (BPost) : commune principale – localité



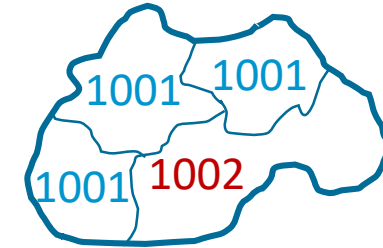
Main commune : Bruxelles(-Ville)

Zipcode	Locality
1000	Bruxelles
1020	Laeken
1120	Neder-Over-Heembeek
1130	Haren



Main commune : Jemeppe-Sur-Sambre

Zipcode	Locality
5190	Jemeppe-Sur-Sambre
5190	Spy
5190	Onoz
5190	Balâtre
5190	...



Main commune : Assesse

Zipcode	Locality
5330	Assesse
5330	Mailen
5330	Sart-Bernard
5332	Crupet
5333	...

Pas d'ID (numérique) pour une commune (seulement une dénomination) → 581 valeurs

ID de localité : code postal + nom de localité (1187 codes postaux, 2765 localités)

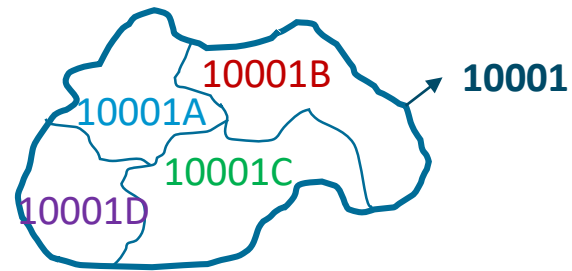
Une série de « codes spéciaux » (1110 NATO, 1044 RTBF, ...)

Schéma de division Belge version statistique

Schéma statistique/administratif (Statbel – SPF Economie) :

Commune principale (NIS5) → 581 valeurs

Sections (NIS6) → 2690 valeurs



Pas de mapping Zipcode-NIS5 (ou NIS6)

Pas de codes spéciaux

Main commune : Bruxelles(-Ville)

NIS5		NIS6
21004	BRUXELLES PENTAGONE	21004A
21004	BRUXELLES-RUE DE LA LOI	21004B
21004	BRUXELLES-LOUISE	21004C
21004	BRUXELLES-CHAUSSEE D'ANVERS	21004D
21004	BRUXELLES-LAEKEN	21004E
21004	BRUXELLES-NEDER-OVERHEEMBEEK	21004F
21004	BRUXELLES-HAREN	21004G



Collecte

- Géocodage
- Sources de données

DB géo

- **Jointure spatiale**

Statistiques

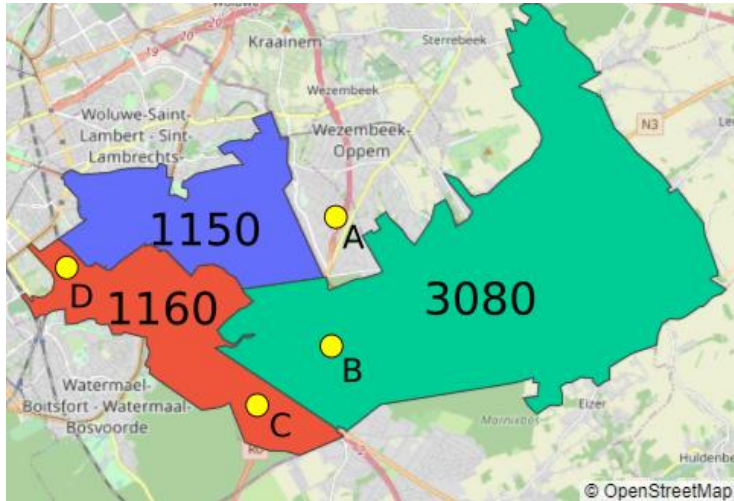
- Autocorrélation spatiale

Analytique

- Atteignabilité
- Dispersion spatiale

Jointure spatiale

Jointure “classique” :



observations

ID	ZIPCODE
A	
B	3080
C	1160
D	1160
...	...

zipcodes

ZIPCODE	PROVINCE
1150	Bruxelles-Capitale
1160	Bruxelles-Capitale
3080	Brabant Flamand
...	...

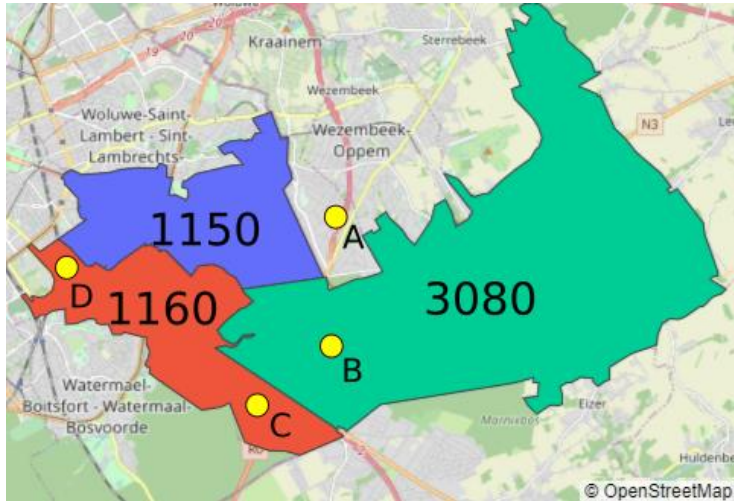
```
SELECT o.ID, o.zipcode, z.province
FROM observations o
LEFT JOIN zipcodes z
ON o.zipcode = z.zipcode
```



ID	ZIPCODE	PROVINCE
A		
B	3080	Brabant Flamand
C	1160	Bruxelles-Capitale
D	1160	Bruxelles-Capitale
...

Jointure spatiale

Jointure “spatiale” :



observations

ID	GEOM
A	point(x_a, y_a)
B	point(x_b, y_b)
C	point(x_c, y_c)
D	point(x_d, y_d)
...	...

provinces

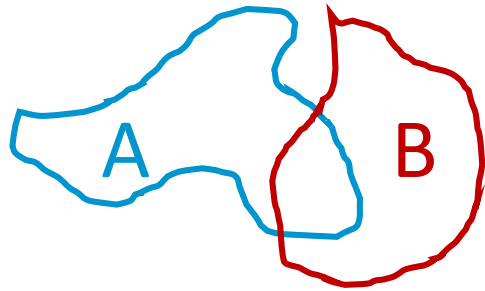
PROVINCE	GEOM
Bruxelles-Capitale	polygon(pt_1, pt_2 □
Brabant Flamand	□
Brabant Wallon	□
...	...

```
SELECT o.ID, p.province
FROM observations o
LEFT JOIN provinces p
ON ST_Contains (p.geom, o.geom)
```



ID	PROVINCE
A	Brabant Flamand
B	Brabant Flamand
C	Bruxelles-Capitale
D	Bruxelles-Capitale
...	...

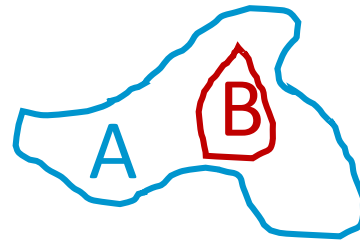
Opérations spatiales (booléennes)



Intersects(A, B)

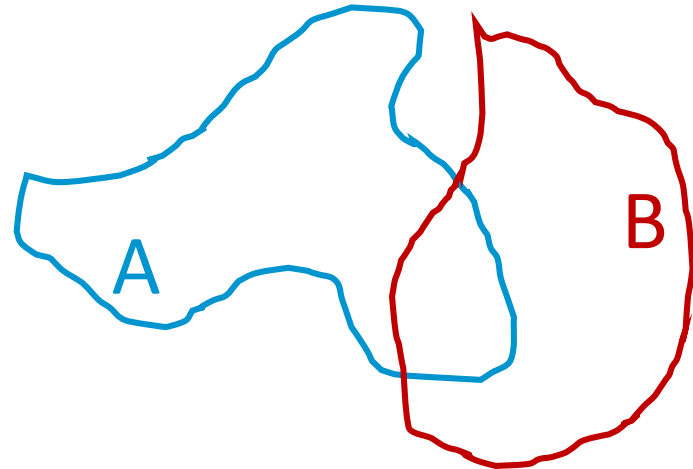


Touches(A, B)

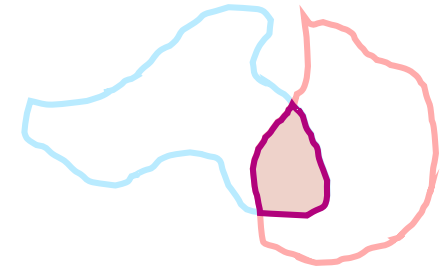


Contains(A, B)
= Within(B, A)

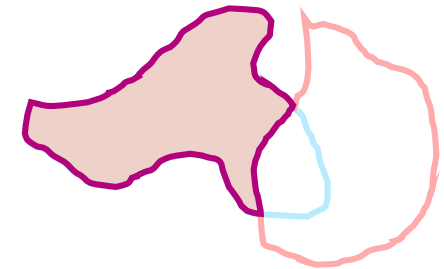
Opérations spatiales (constructives)



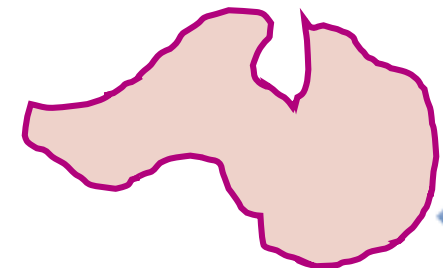
Intersection(A, B):



Difference (A, B):

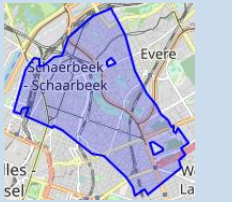


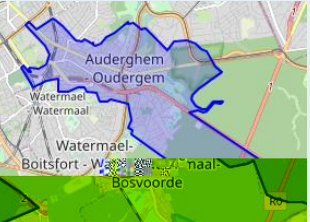


Union (A, B):




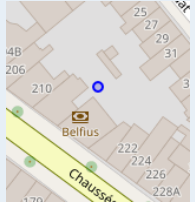
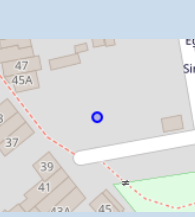

Jointure spatiale point vs polygone

Bpost – frontières postales

zipcode	geom
1030	
1060	
1130	
1160	



AdN – centres postaux

zipcode	geom
1030	
1060	
1160	
7300	






Bp.zipcode	Bp.geom	AdN.zipcode	AdN.geom
1030	...	1030	...
1060	...	1060	...
1130	...	7300	...
1160	...	1160	...

53 « mismatches » (dans AdN) :

- 1130 et 3700 Haren au même endroit
- « 8600 Driekapellen » localisé « Driekapellenstraat, 1541 St. Pt Kap. »
- « 6831 Noirfontaine » localisé « Rue de Noirfontaine, 4400 Flémalle »
- Nom de commune vs lieu dit (Laar, Schriek, Berg, Voort...)



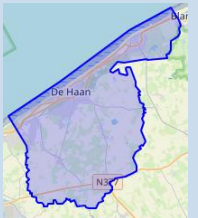
Jointure spatiale polygone vs polygone

Bpost – frontières communales

commune	geom
VILVOORDE	
HAM-SUR-HEURE	
DE HAAN	



Statbel – frontières communales

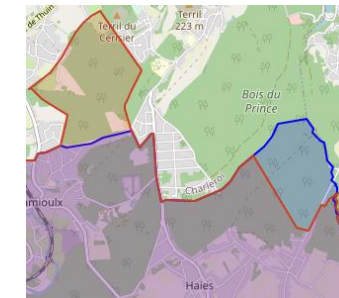
commune	geom
Vilvorde	
Ham-sur-Heure-Nalinnes	
De Haan	



AND

()

 ()



Côte : avec (Statbel) ou sans (BPost) la plage

Situations connues: 1040 = Etterbeek + Europe (Bxl) ; 1050 = Ixelles + Louise-Roosevelt (Bxl)

Incohérences ?

Détection d'anomalies:

Centre de code postal pas situé dans le bon polygone (52 cas pour l'Agence du Numérique)

Superposition de polygones :

Intrasource: territoires attribués à deux codes postaux (Bpost: ~15 CP spéciaux pas correctement "extraits")

Intersource: territoires partagés entre une commune "postale" et une commune "statistique" (~25 incohérences)

Attribuer un code postal, une commune, province, arrondissement... à chaque point d'une liste

Combinaison avec une courbe isochrone: communes accessibles en X minutes à partir d'un point (cf plus loin)



Collecte

- Géocodage
- Sources de données

DB géo

- Jointure spatiale

Statistiques

- **Autocorrélation spatiale**

Analytique

- Atteignabilité
- Dispersion spatiale

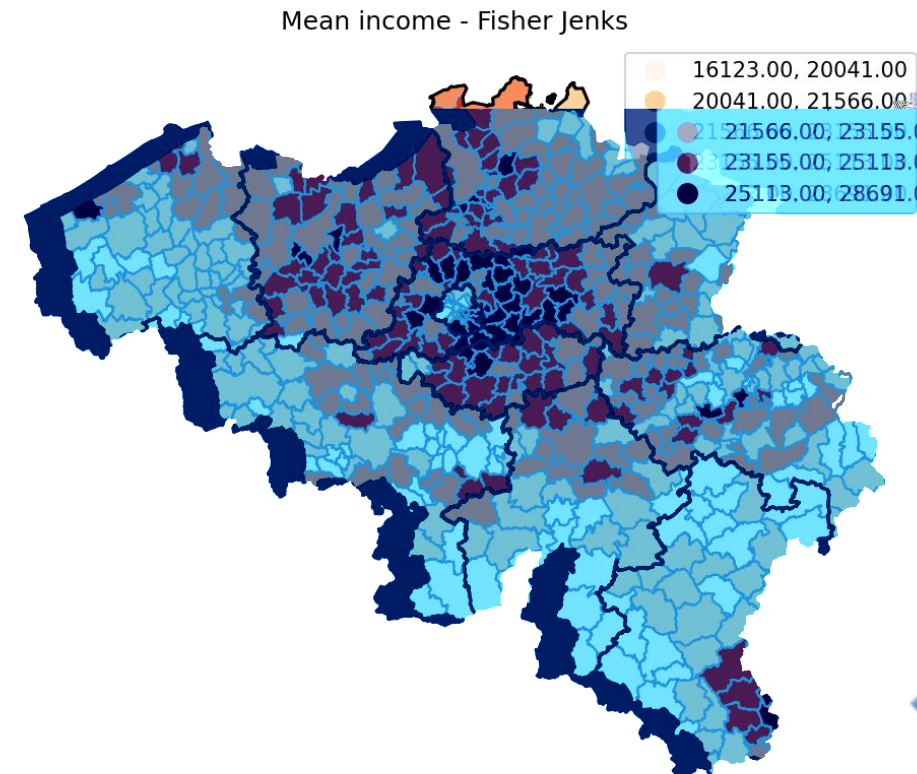
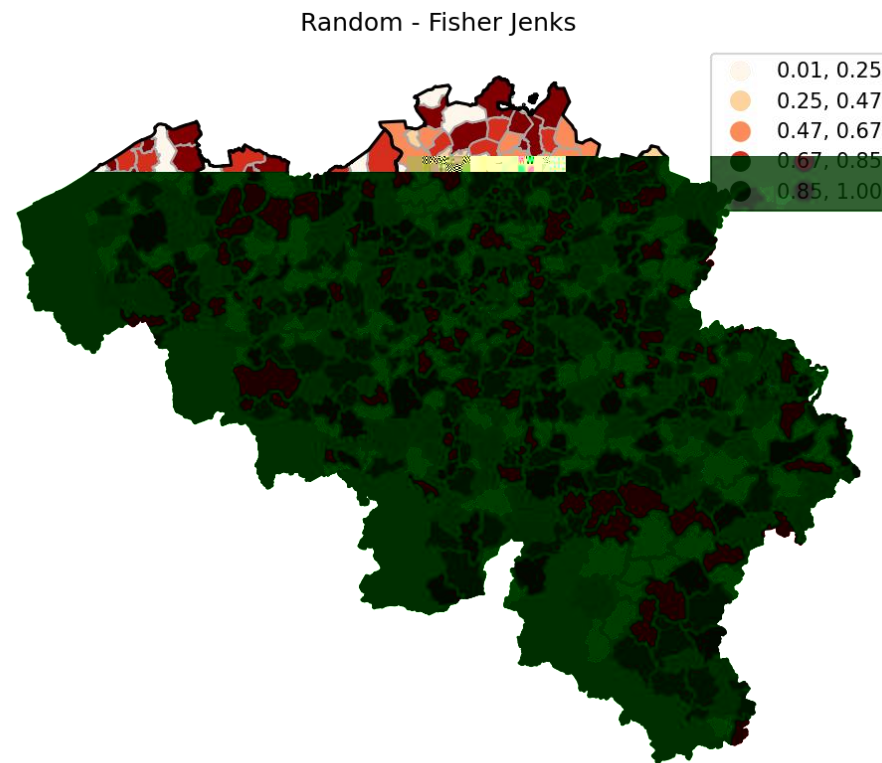
Autocorrélation spatiale

La valeur d'une observation d'une variable A est-elle liée aux observations voisines (géographiquement) ? Un indicateur a-t-il une **composante géographique** ?

A-t-on plus de chance d'observer un salaire élevé près d'une commune avec un salaire élevé ?

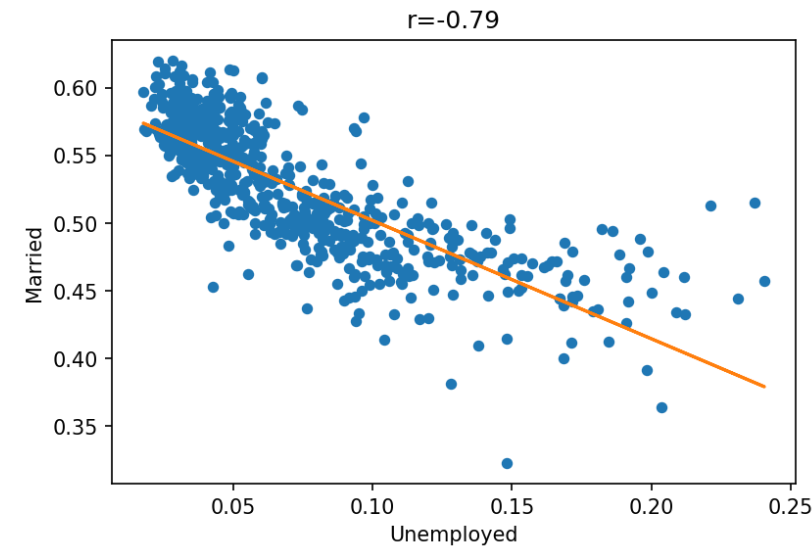
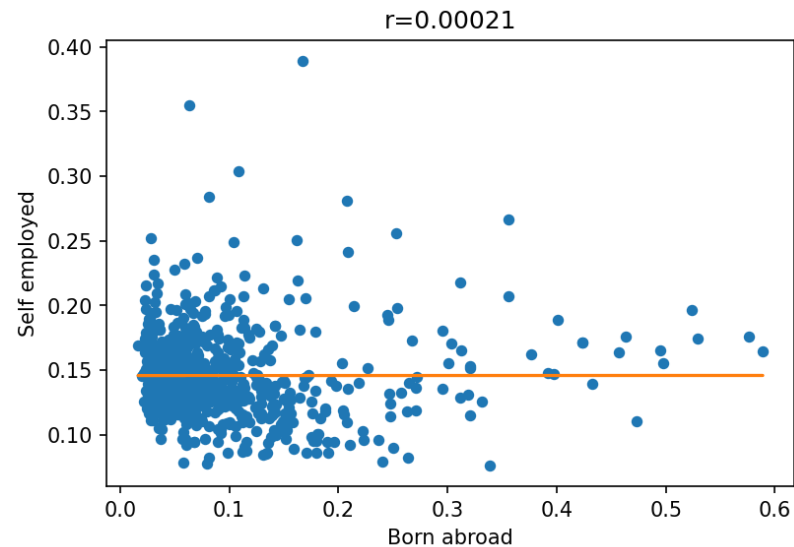
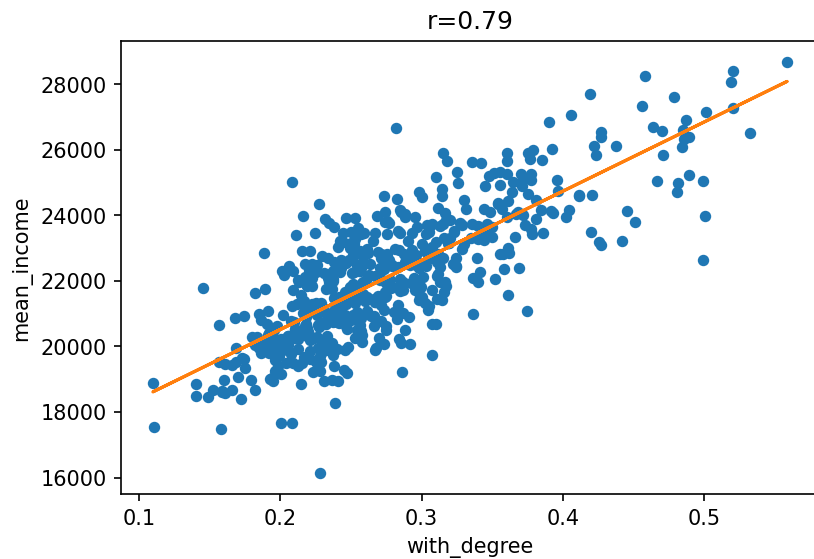
Le salaire moyen a-t-il un plus grande dépendance géographique que le niveau d'éducation ?

Visuellement assez clair, mais comment le **formaliser/quantifier** ?



Corrélation

Corrélation (classique) : caractérise le **niveau de liaison** entre deux « variables » (ou plus)



Corrélation A vs B : connaitre de A **donne une indication** de B

Causalité A vers B : modifier A aura un **impact** sur B

Coefficient de corrélation (Pearson's r) :

1= corrélation parfaite,

0= aucune corrélation,

-1= corrélation inverse

Corrélation ~~⇒~~ Causalité !!

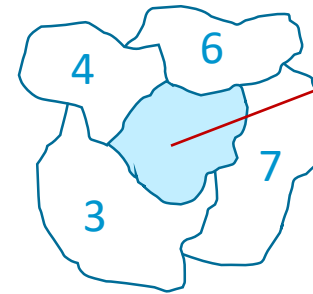
Spatial lag

Spatial lag (« décalage spatial ») : valeur **moyenne** des **zones voisines** de chaque zone

Sans tenir compte de « soi-même »

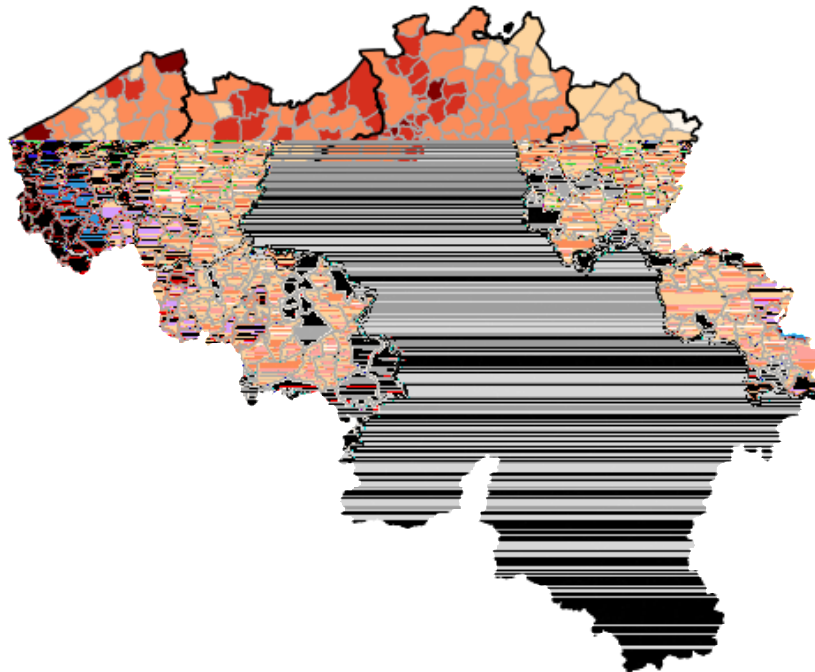
Permet de « lisser les valeurs »

Nécessite un **calcul de voisinage** : quelle sont les voisins de chaque zone ? (Queen neighborhood)

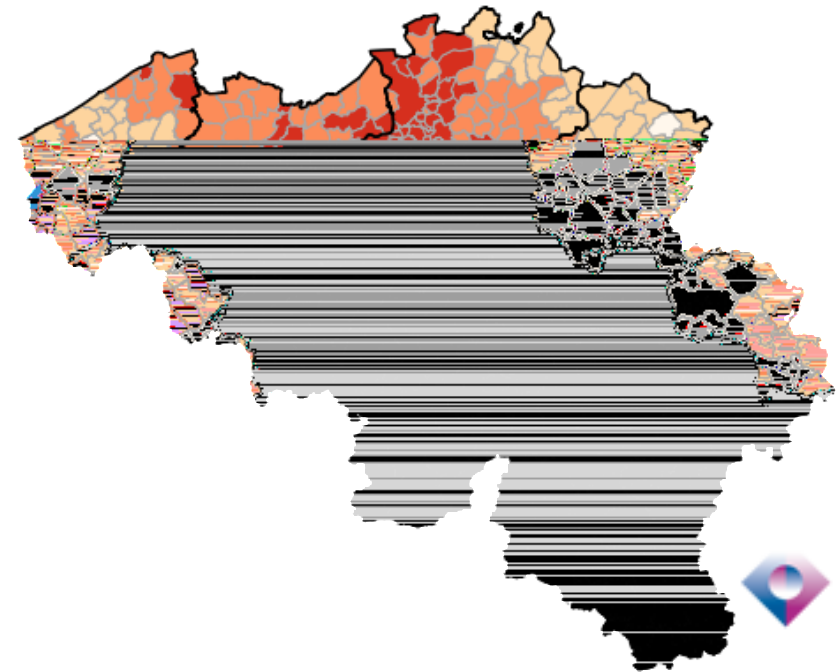


Spatial lag: $\text{avg}(3, 4, 6, 7) = 5$

Value (Fisher-Jenks)

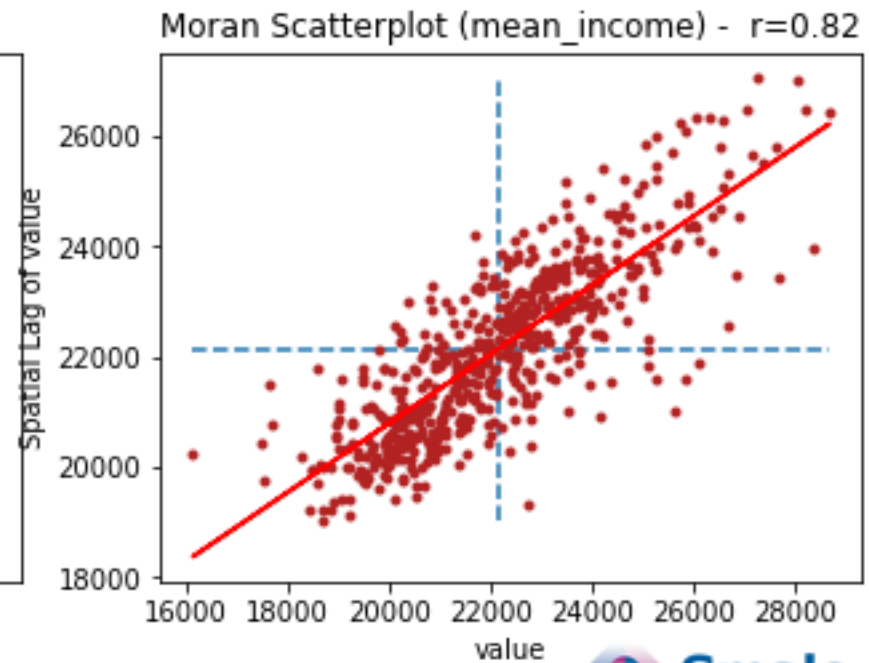
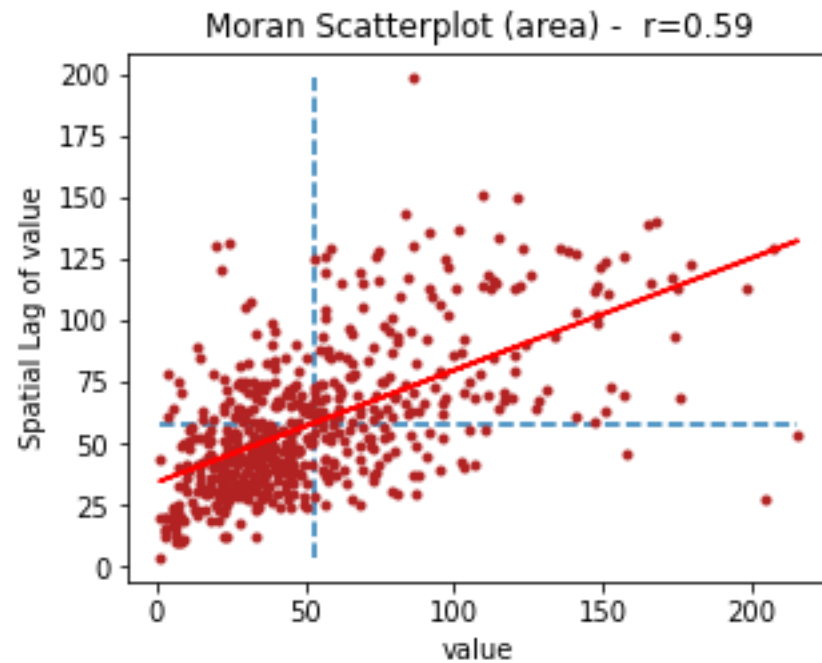
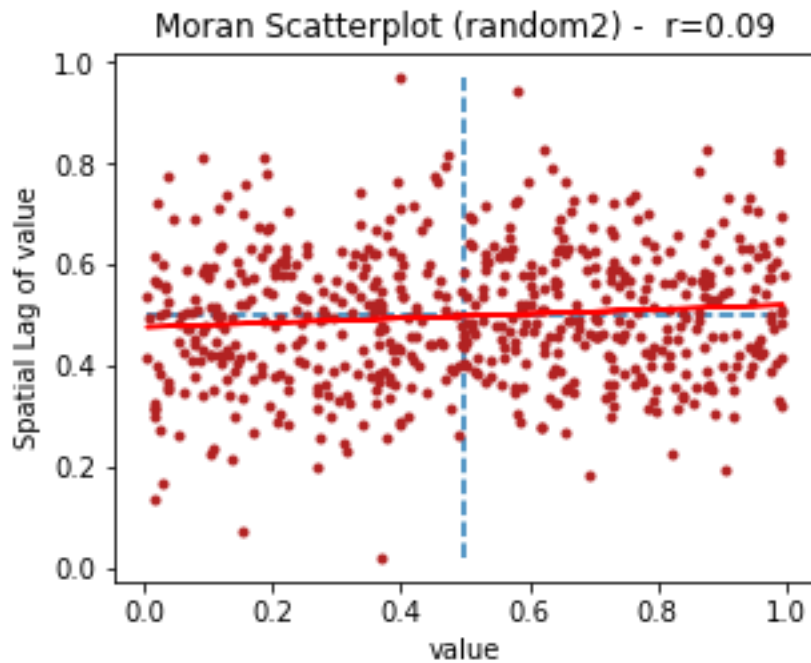


Value lag (Fisher-Jenks)



Moran scatterplot

Indicateur de dépendance spatiale : corrélation (classique) entre la valeur et son « spatial lag »
Aucune pour « random », très nette pour « mean income » → indique une forte dépendance spatiale (valeur d'une zone corrélée à la valeur de ses voisines)



Autocorrélation globale : version binaire

Autocorrélation globale : calcule d'un indicateur « global » mesurant l'intensité de la composante spatiale

On convertit chaque valeur en 0 (\leq médiane, « White ») ou 1 ($>$ médiane, « Black »)

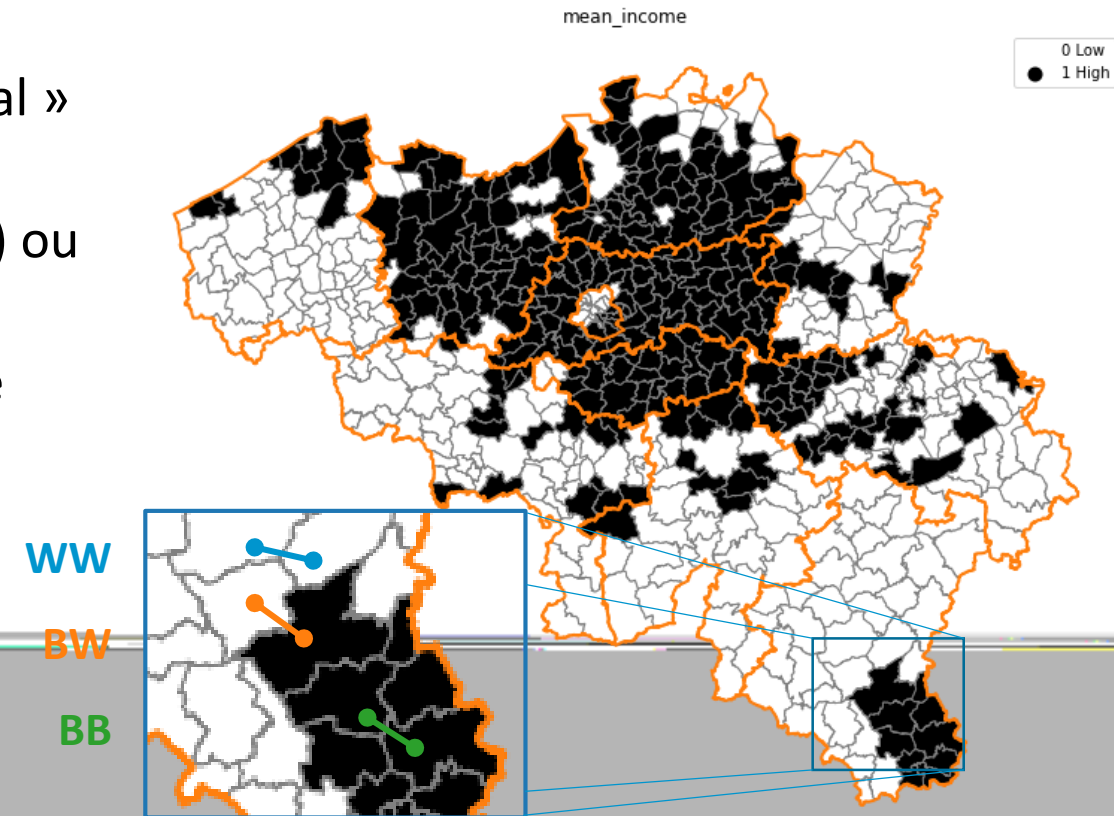
On regarde le nombre de « jonction Black-Black » (zone « Black » touchant une autre zone « Black ») :

mean_income : 646

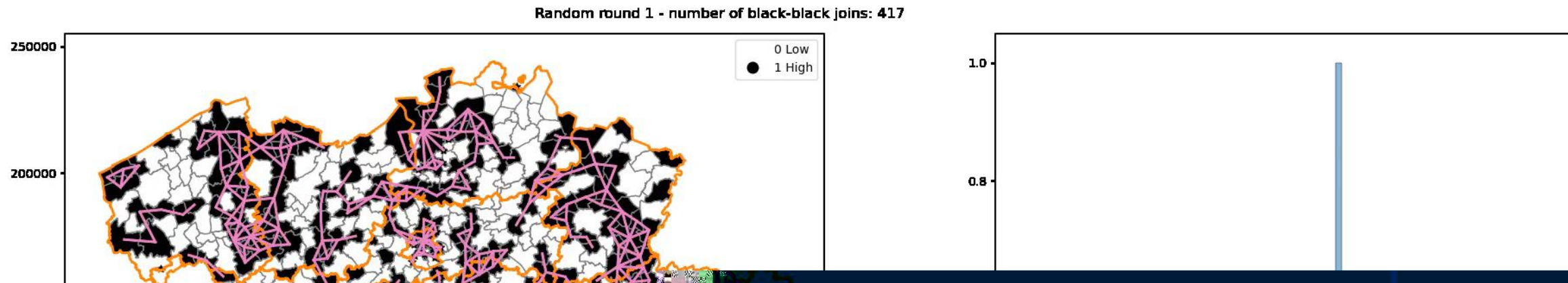
with_degree : 601

area : 580

Est-ce beaucoup ? On compare avec un indicateur « random » sans aucune composante géographique



Autocorrélation globale : version binaire



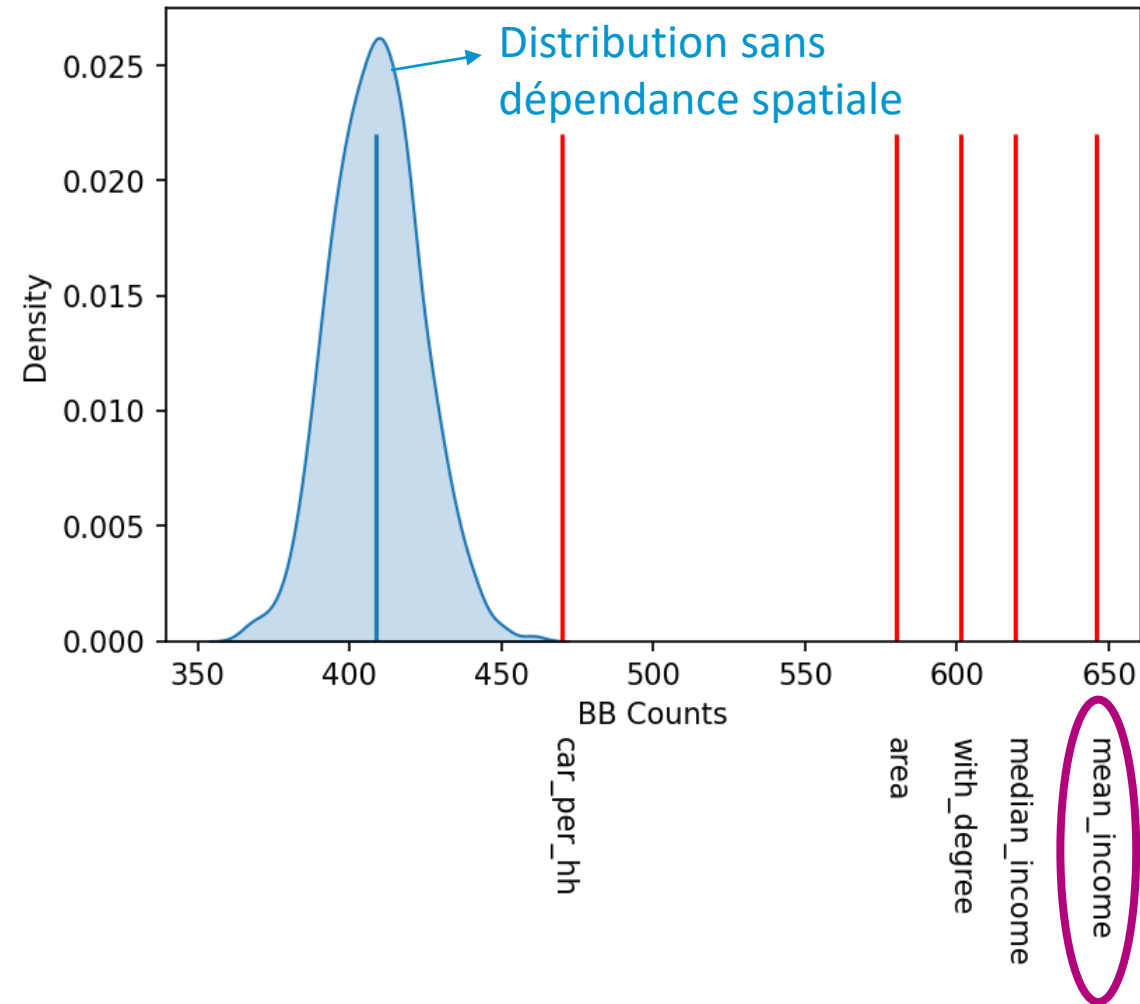
Autocorrélation globale : version binaire

On compare avec le nombre qu'on aurait eu en choisissant **la couleur au hasard** (pour la même géographie)

Cloche : distribution d'un grand nombre de **tirages aléatoires**

Plus on **s'en éloigne**, plus faible est la probabilité que la métrique soit aléatoire ... donc plus grande est qu'elle soit « **spatialement dépendante** »

Il existe une généralisation « continue » (Moran's I)



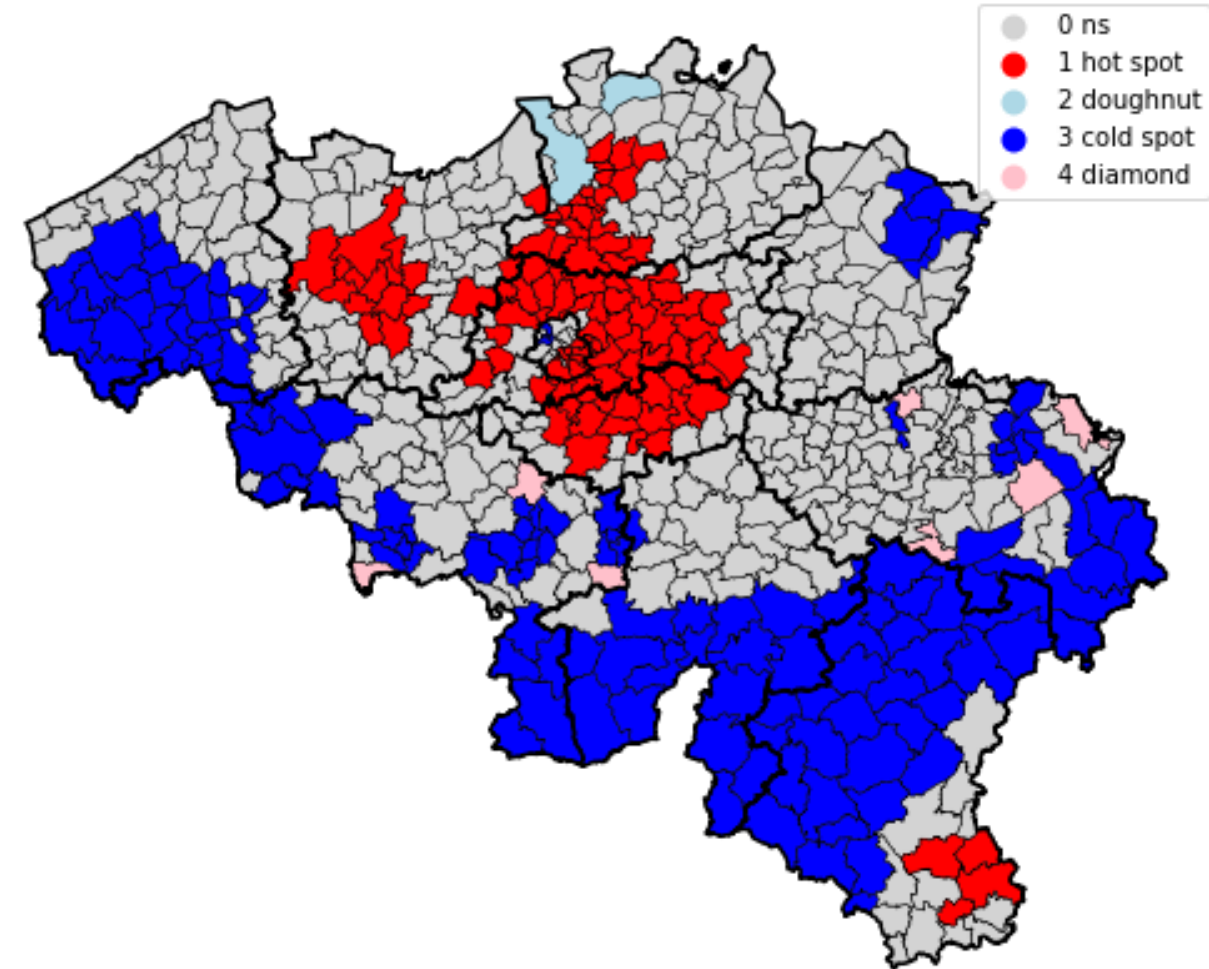
Autocorrélation locale

Autocorrélation locale : indicateur associé à chaque secteur

En simplifiant :

Si voisinage proche de la moyenne : non significatif ○

Sinon :



		REVENU MOYEN VOISINAGE			
		HAUT		BAS	
REVENU MOYEN	HAUT	Hot spot	●	Diamond	○
	BAS	Doughnut	○	Cold spot	●

Autocorrélation spatiale : conclusion

L'autocorrélation permet quantifier/objectiver la **dimension géographique d'une métrique**

3 méthodes :

Corrélation classique entre une valeur et son « spatial lag » (décalage spatial)

Autocorrélation globale : un nombre « global » pour quantifier la dépendance géographique

Autocorrélation locale : pour chaque secteur, mesure de l'importance de sa dépendance spatiale

Permet aussi d'identifier des **anomalies** (« hot spots », « donnuts »)

Ne permet pas de déduire de causalité, mais est une première étape

Métriques attribuées à des **secteurs**. Il existe aussi des méthodes pour des **points**. Exemple : corrélation entre le chiffre d'affaire/nombre de travailleurs/... d'un restaurant et celui des restaurant voisins ?



Collecte

- Géocodage
- Sources de données

DB géo

- Jointure spatiale

Statistiques

- Autocorrélation spatiale

Analytique

- **Atteignabilité**
- Dispersion spatiale

Multipoint reachability

Question centrale : **quel territoire ou quelle population est atteignable** à partir d'un ensemble de points (hôpitaux, antenne locale d'une administration, domicile des agents de terrain...)

En supposant un déplacement en voiture (possibles extensions : à pieds, en transport en commun...)

On combine des « **courbes isochrones** »

On va **comparer** plusieurs **scénarios** (ouverture/fermeture ou non d'un ou plusieurs départs)

Usecase : SMUR (MUG) à Bruxelles : 10 départs, dont 3 en alternance un semaine/3

« Disclaimer »: déplacement à allure « normale » et analyse simplifiée !

Plus de détails :

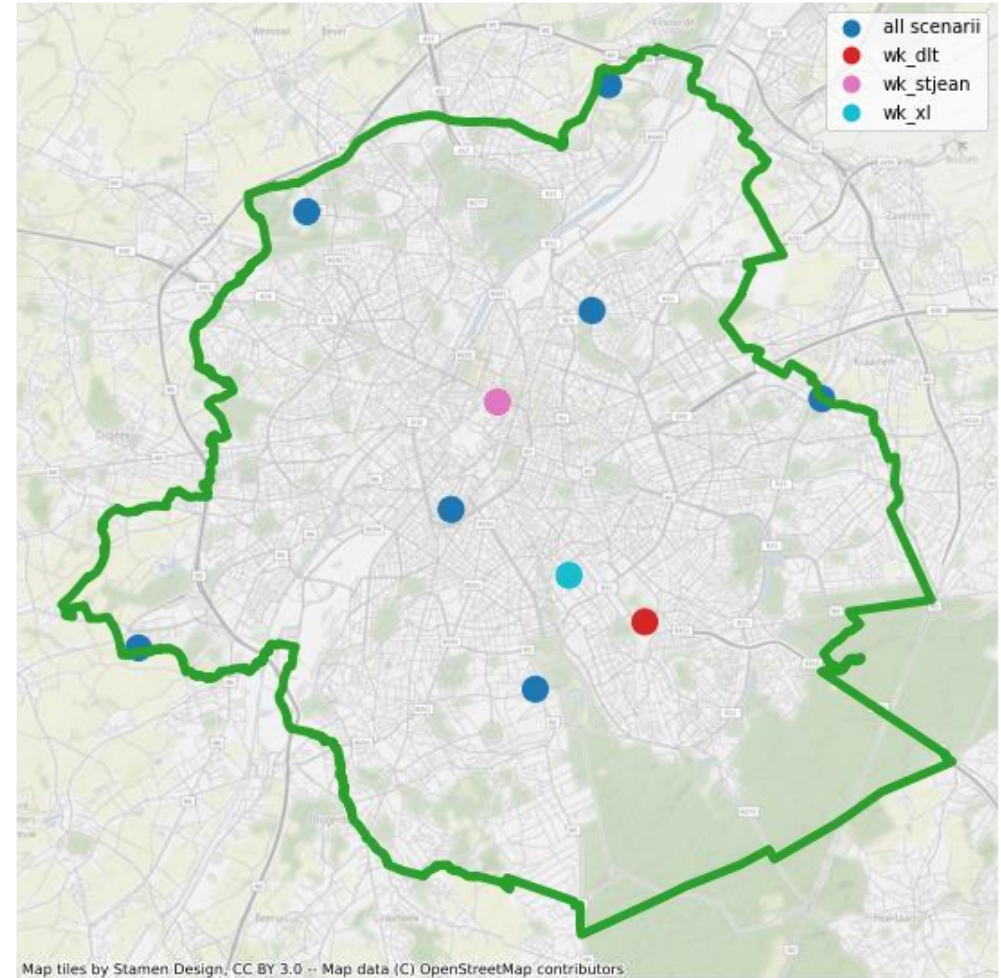
<https://www.smalsresearch.be/peut-on-toujours-atteindre-une-maternite-en-30-minutes/>

<https://www.smalsresearch.be/peut-on-toujours-atteindre-une-maternite-en-30-minutes-partie-2/>

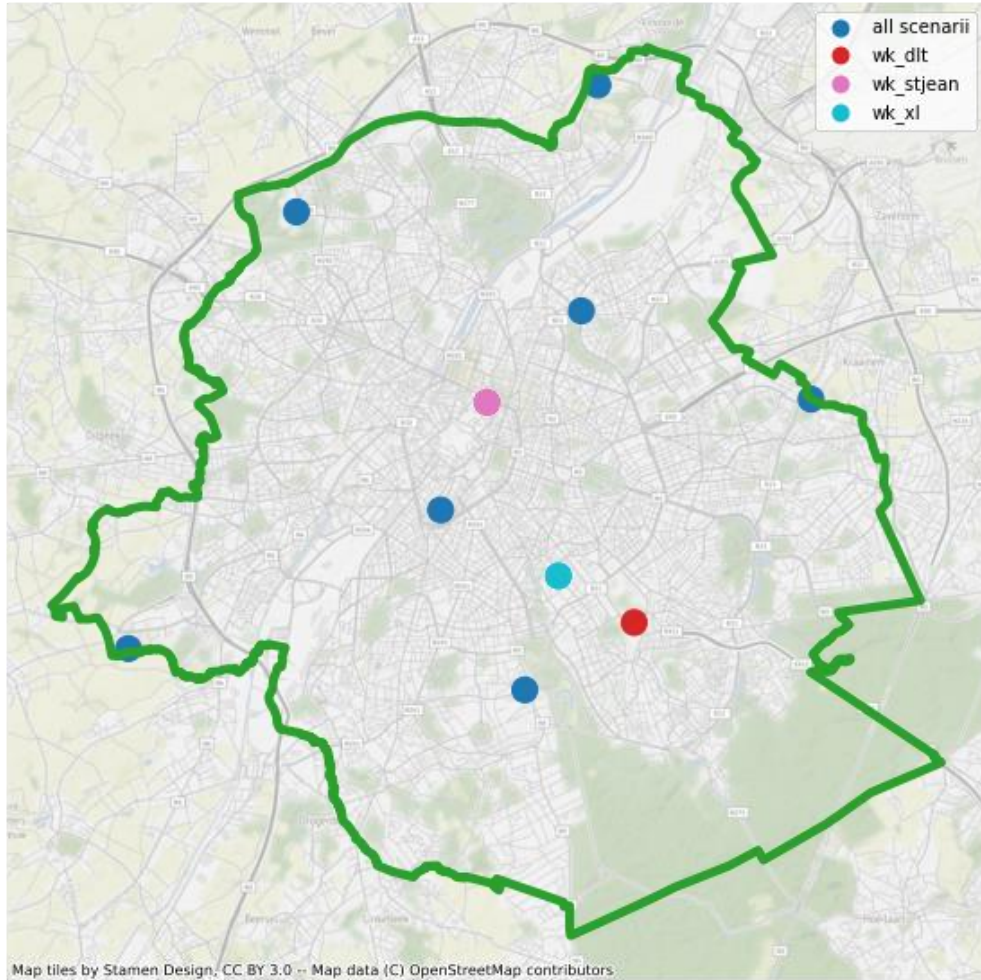
Départs

		wk_dlt	wk_stjean	wk_xl
Saint-Luc	Av. Hippocrate	x	x	x
Saint-Pierre	Rue haute	x	x	x
...	...			
Delta	Bvd du triomphe	x		
Saint-Jean	Rue du marais		x	
Ixelles	Rue Paquot			x

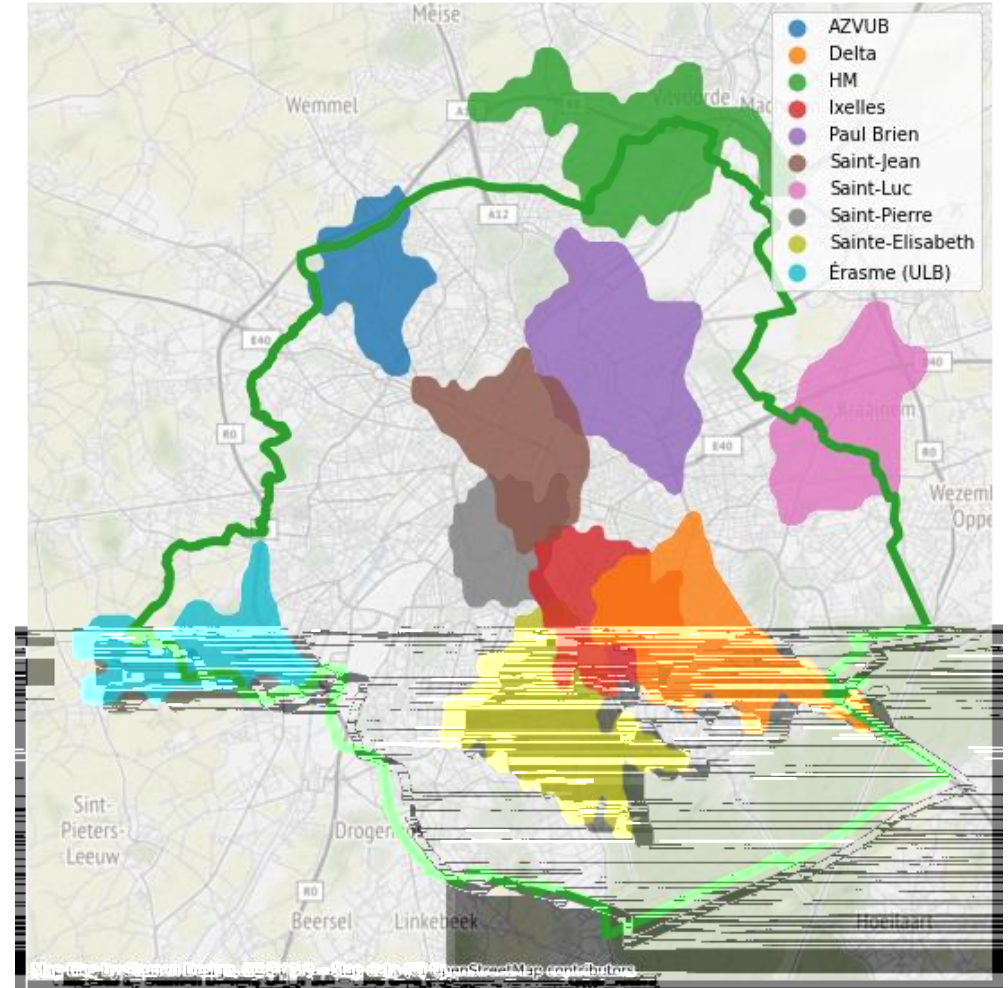
Géocodage



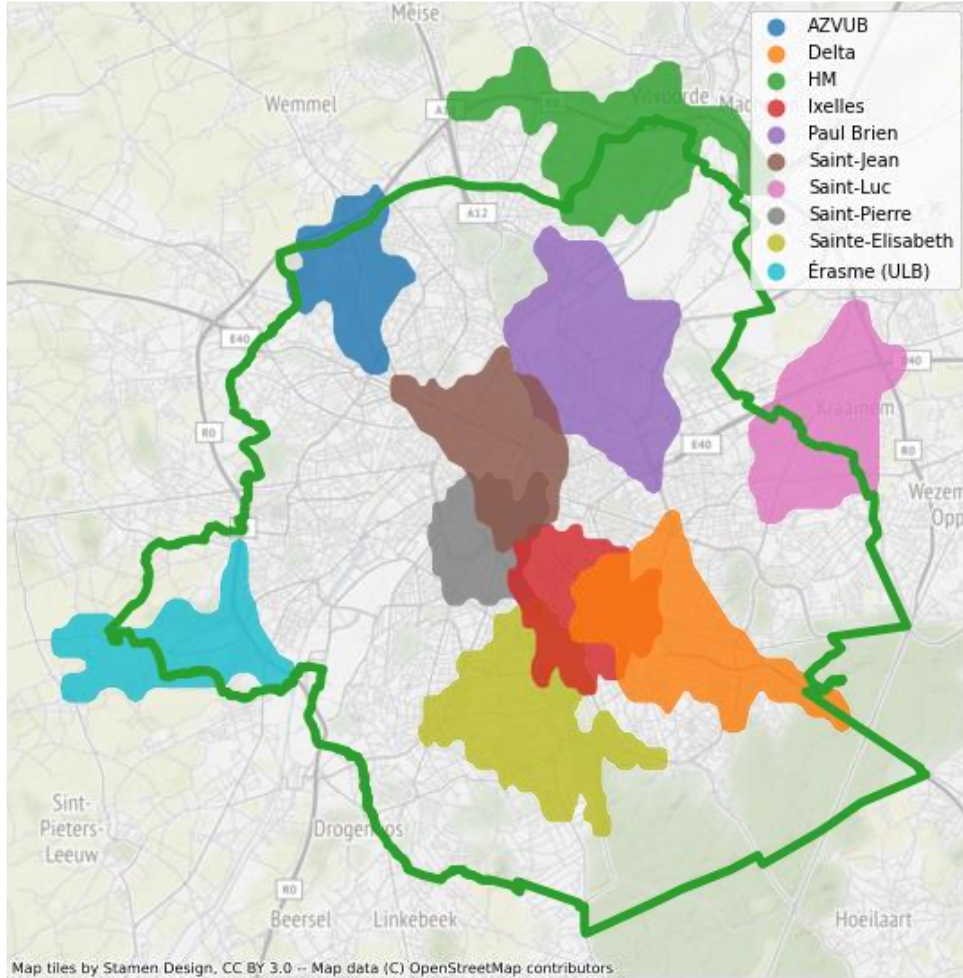
Courbes isochrones



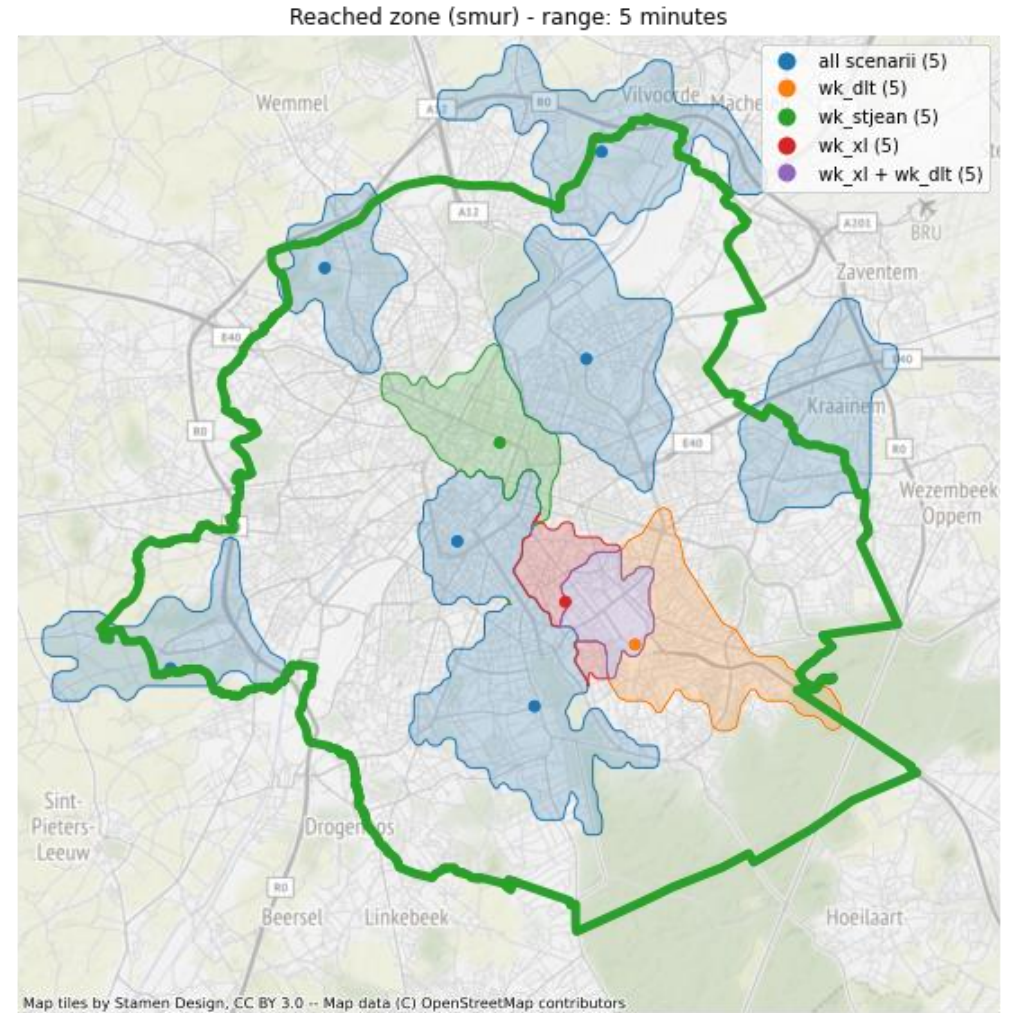
Isoline (5 min) →



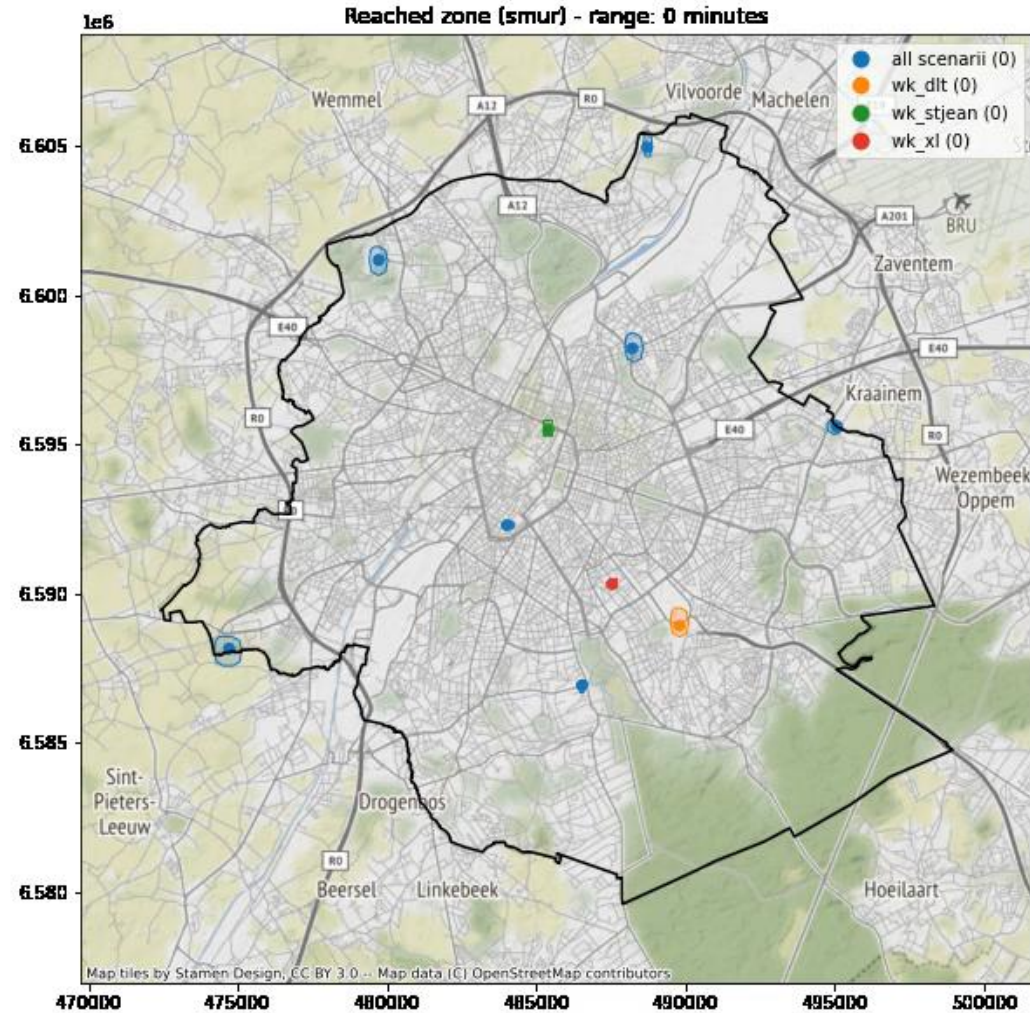
Pour un délai donné □



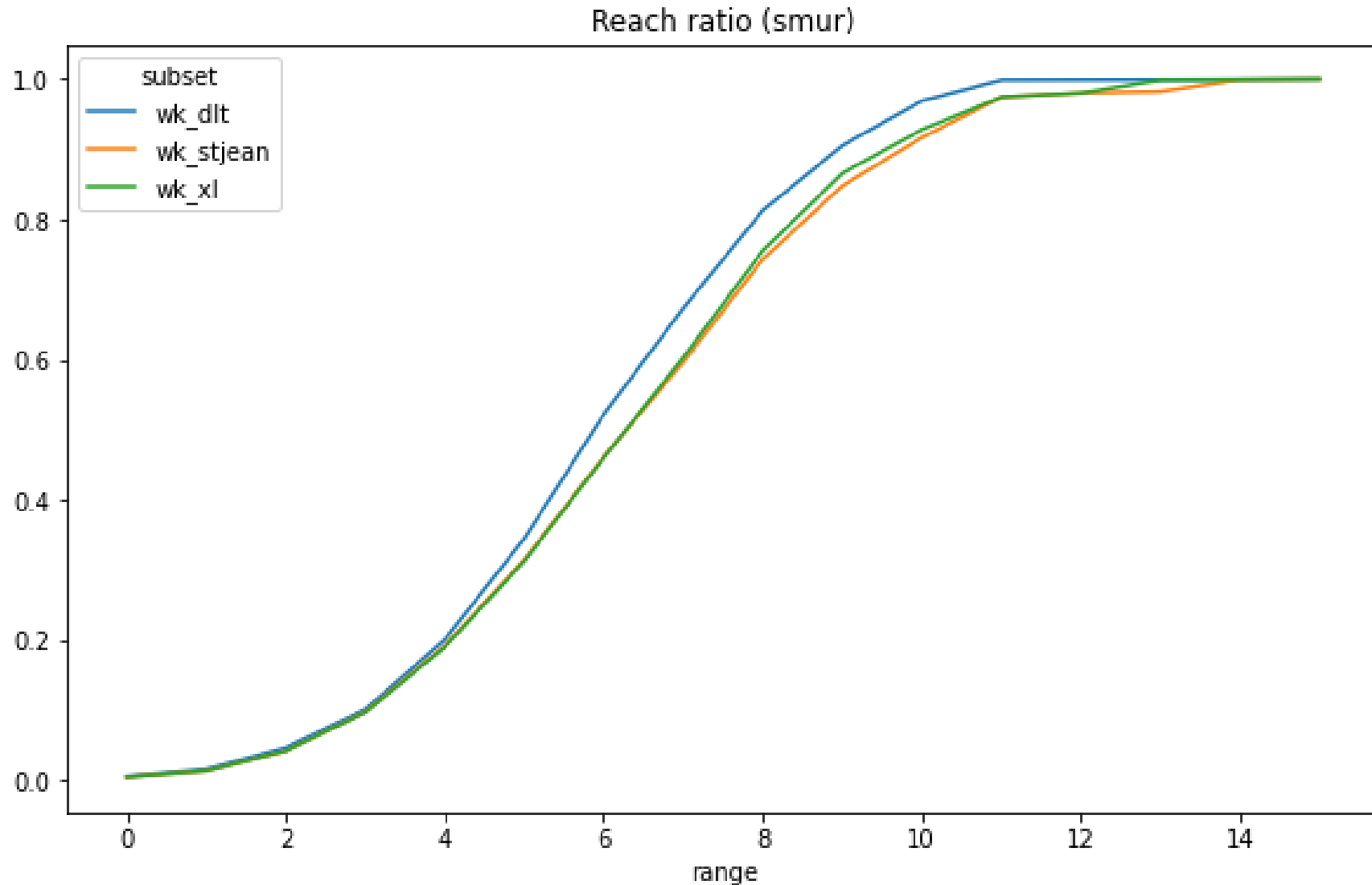
Spatial ops



Temps variable



Vue globale : reached ratio



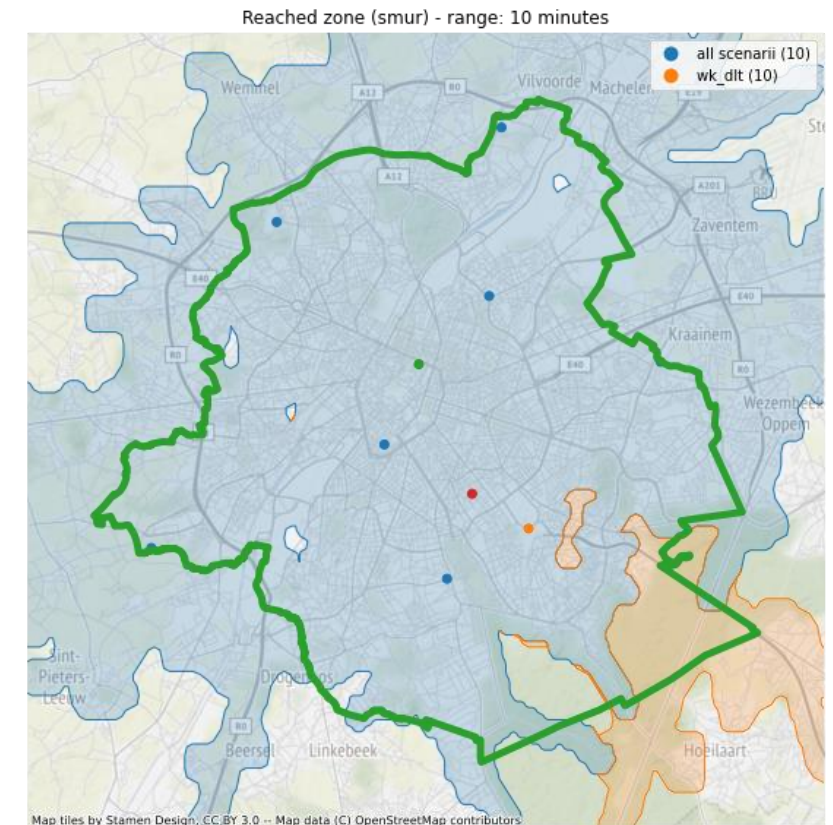
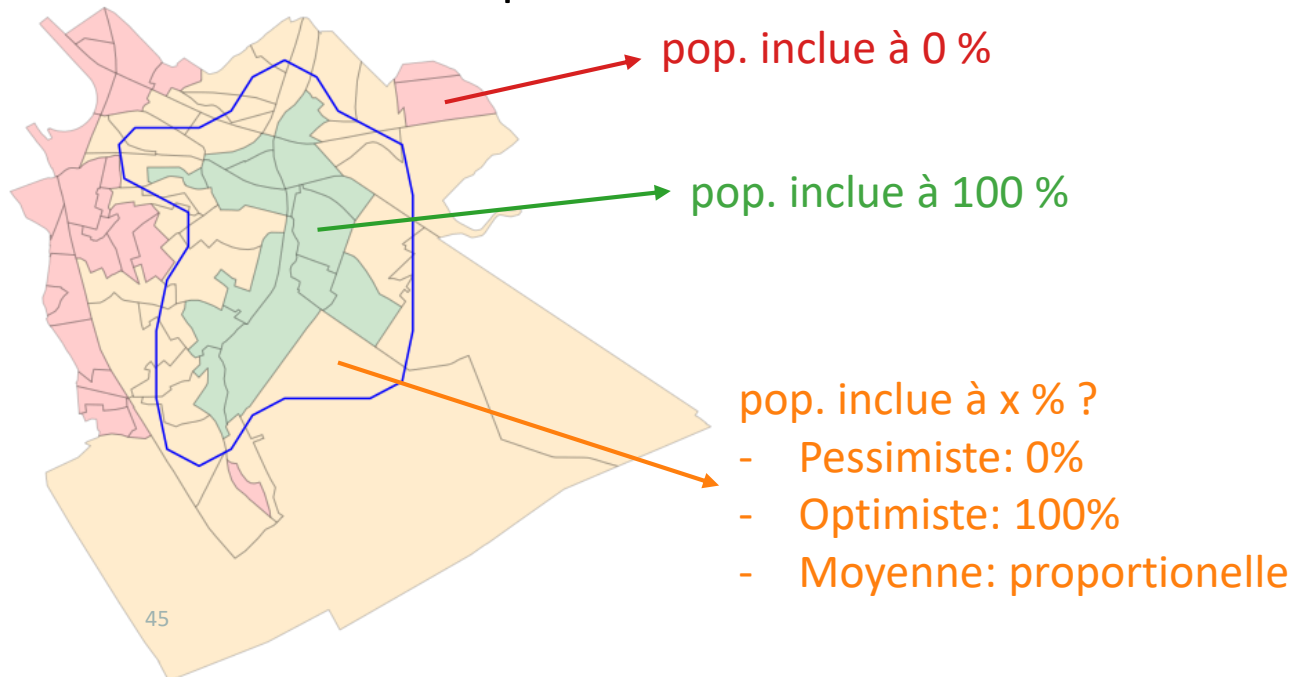
Couverture populationnelle

Couverture territorial <> couverture populationnelle !

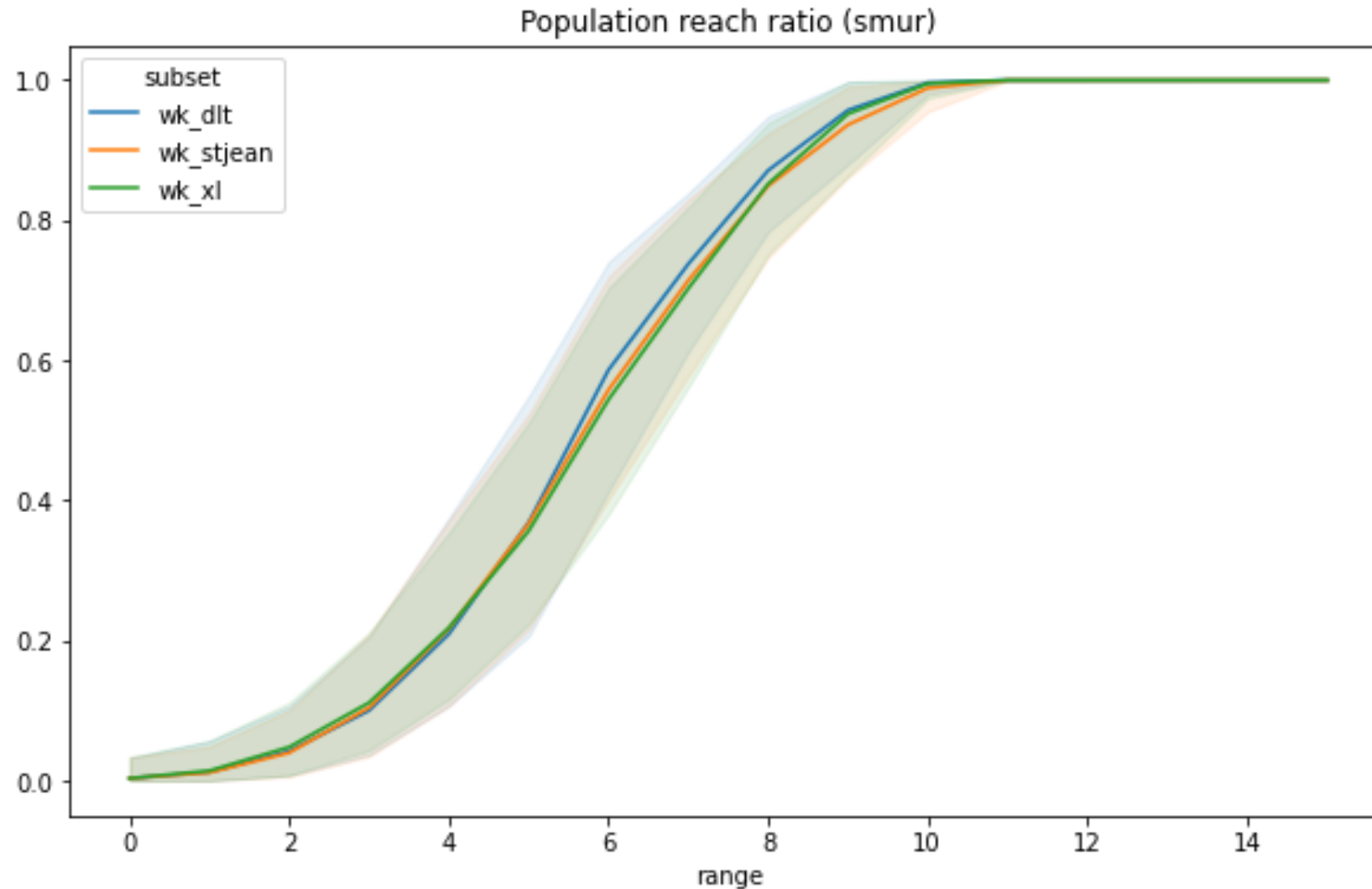
Plus-value de la semaine "Delta": couverture de ...
la forêt de Soignes !

Connaissant la population par secteur statistique (Statbel), on
peut approximer la population couverte par jointure spatiale

"Approximer" car on ne connaît pas la répartition au sein des
secteurs statistiques



Population reach ratio





Collecte

- Géocodage
- Sources de données

DB géo

- Jointure spatiale

Statistiques

- Autocorrélation spatiale

Analytique

- Atteignabilité
- **Dispersion spatiale**

Check-in **At Work** (ONSS) : déclaration de présence sur un chantier (**Place Of Work**)

Peut-être fait :

- En direct : scan d'un QR code

- À l'avance/distance

Objectif : identifier les entreprises qui systématisent le principe de “check in everybody everywhere”

Problèmes similaires :

- Identifier un médecin qui envoie des certificats à des personnes “trop distantes” (INAMI ?)

- Un indépendant qui déclare des chantiers trop distants (Fisc ?)

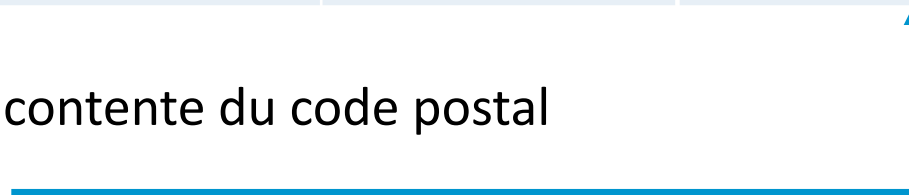
CAW : données en entrée

Données en entrée (pour une date en particulier) :

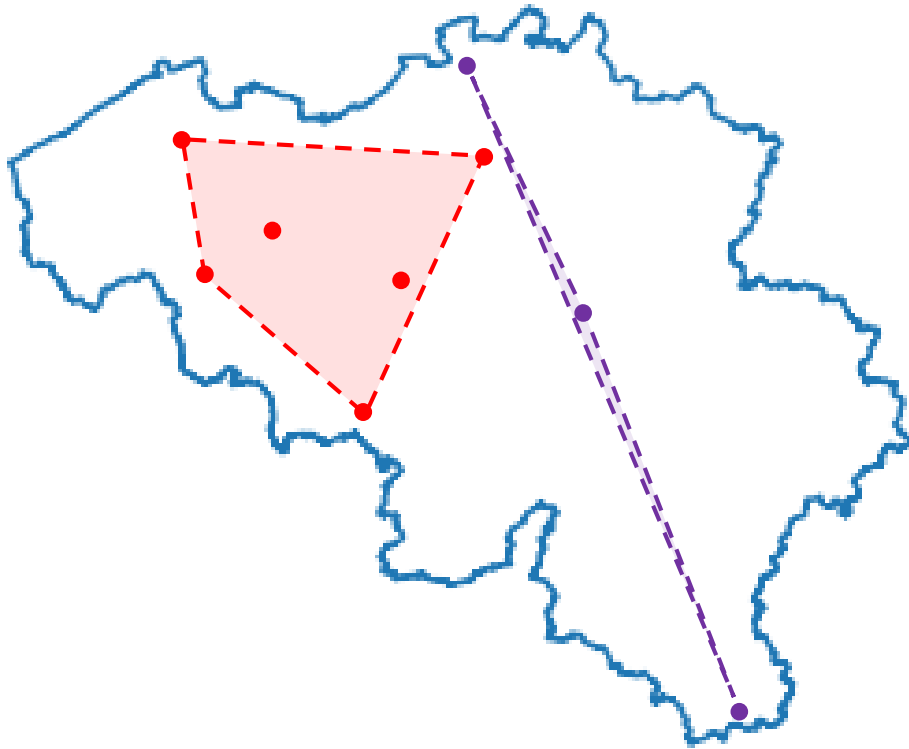
WORKER_ID	POW_ID	STATUS	ZIPCODE	GEOM
Wrk ₁	Pow ₁	OK	1060	POINT (4.34466 50.82946)
Wrk ₁	Pow ₂	OK	1160	POINT (4.43800 50.81019)
Wrk ₂	Pow ₃	NOK	5190	POINT (4.67619 50.47171)
Wrk ₃	Pow ₄	CANCELLED	6220	POINT (4.52807 50.47096)

Pour éviter du géocodage, on se contente du code postal

Jointure avec les données BPost



Enveloppe convexe



Pour chaque travailleur, on rassemble tous les points

On construit une “**enveloppe convexe**” (convex hull)

Périmètre de l’enveloppe : **borne inférieure** de la **distance à parcourir** par le travailleur (retour au départ)

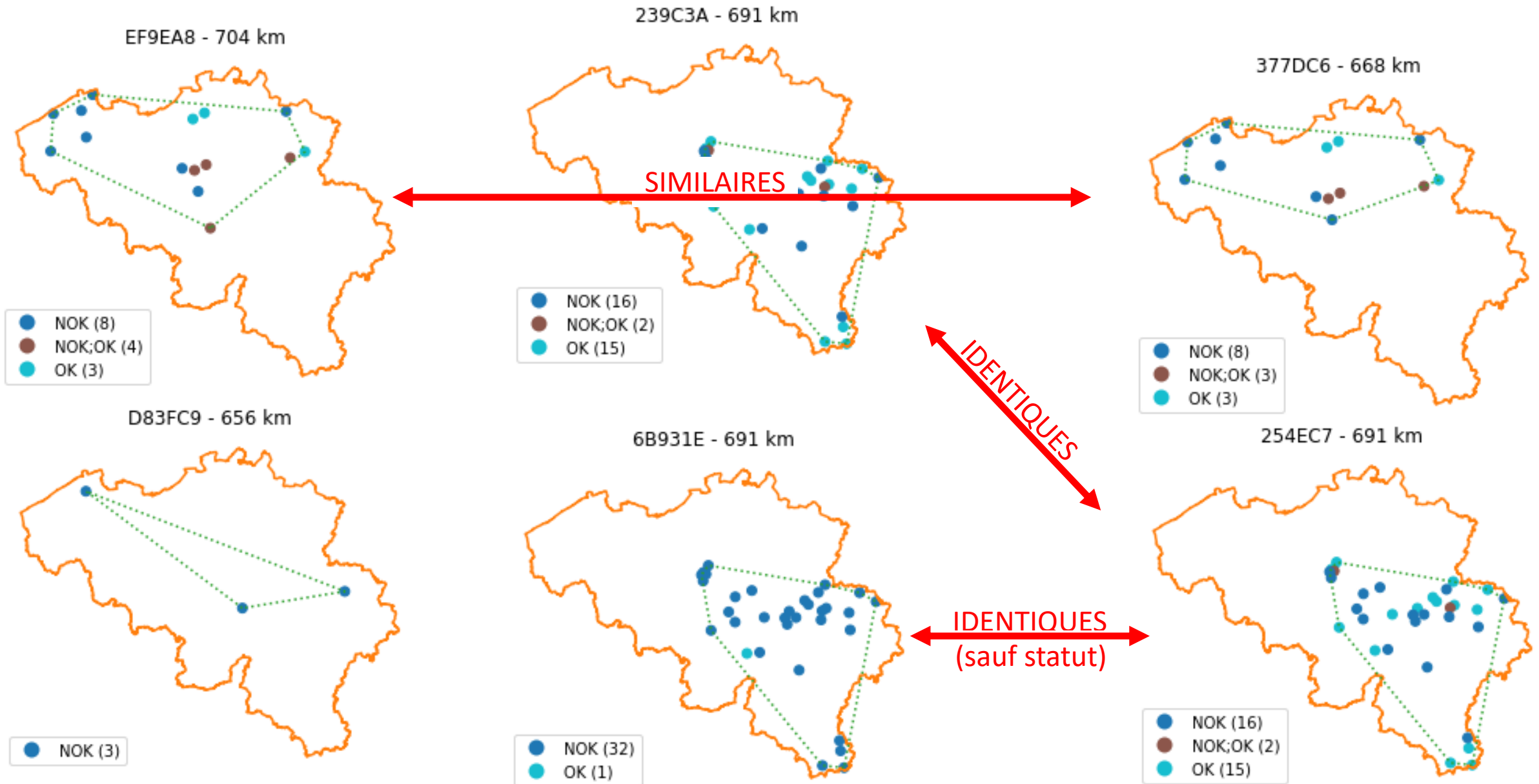
Inférieure, car :

- Ignore les points dans l’enveloppe

- À vol d’oiseau

- Supposant que le travailleur habite près d’un des points

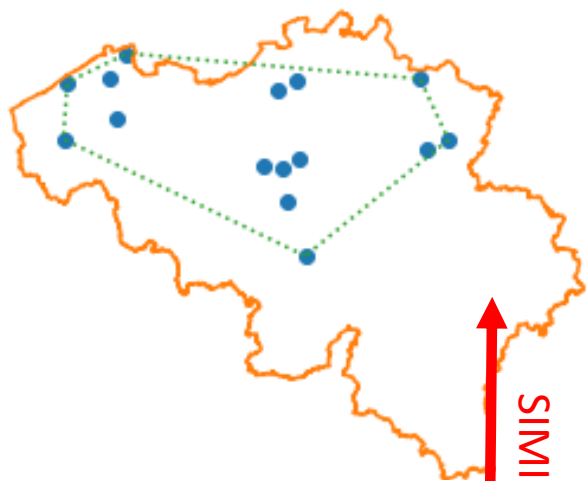
Exemples



Patterns identiques

Constat: de nombreux travailleurs ont le "même pattern" → regroupement (en ignorant le statut)

1 workers - 15 pow - 704 km



3 workers - 33 pow - 691 km



9 workers - 7 pow - 547 km



2 workers - 14 pow - 668 km



SIMILAIRES

workers - 34 pow - 691 km

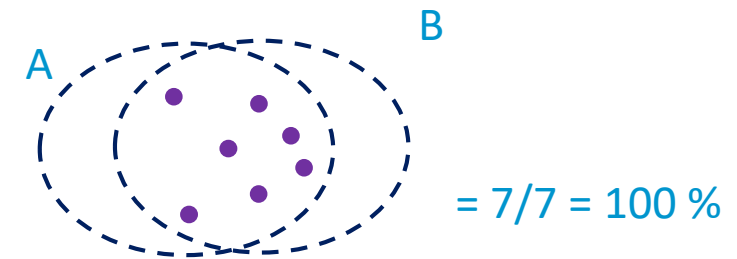
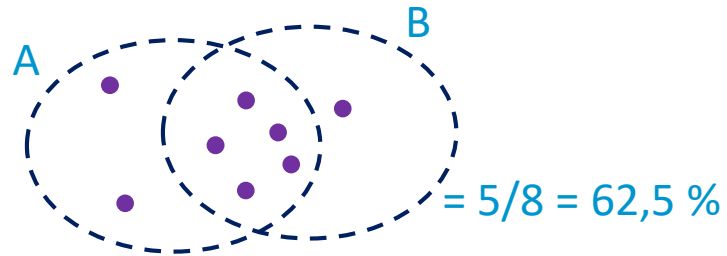
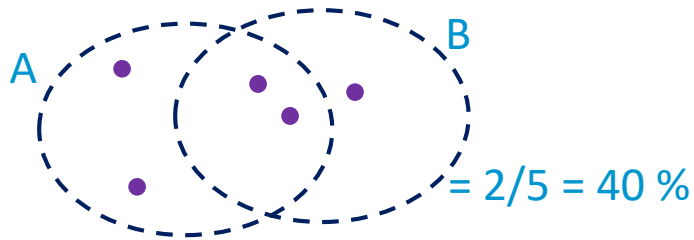
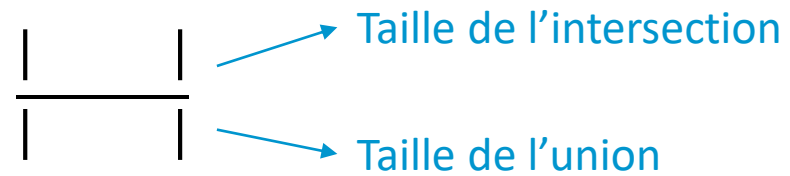


13 workers - 3 pow - 631 km



Patterns similaires

On peut évaluer la proximité entre deux ensembles A et B par la “similarité de Jaccard” :



“Point”= centre du code postal du POW

Patterns similaires

Jaccard similarity: 97.1%

3 workers - 33 pow - 691 km



1 workers - 34 pow - 691 km



● Points communs
● Différences

Jaccard similarity: 93.3%

2 workers - 14 pow - 668 km



1 workers - 15 pow - 704 km



Use case CAW : conclusions

Ceci est un simple POC “naïf”

On se contente de la localisation du code postal

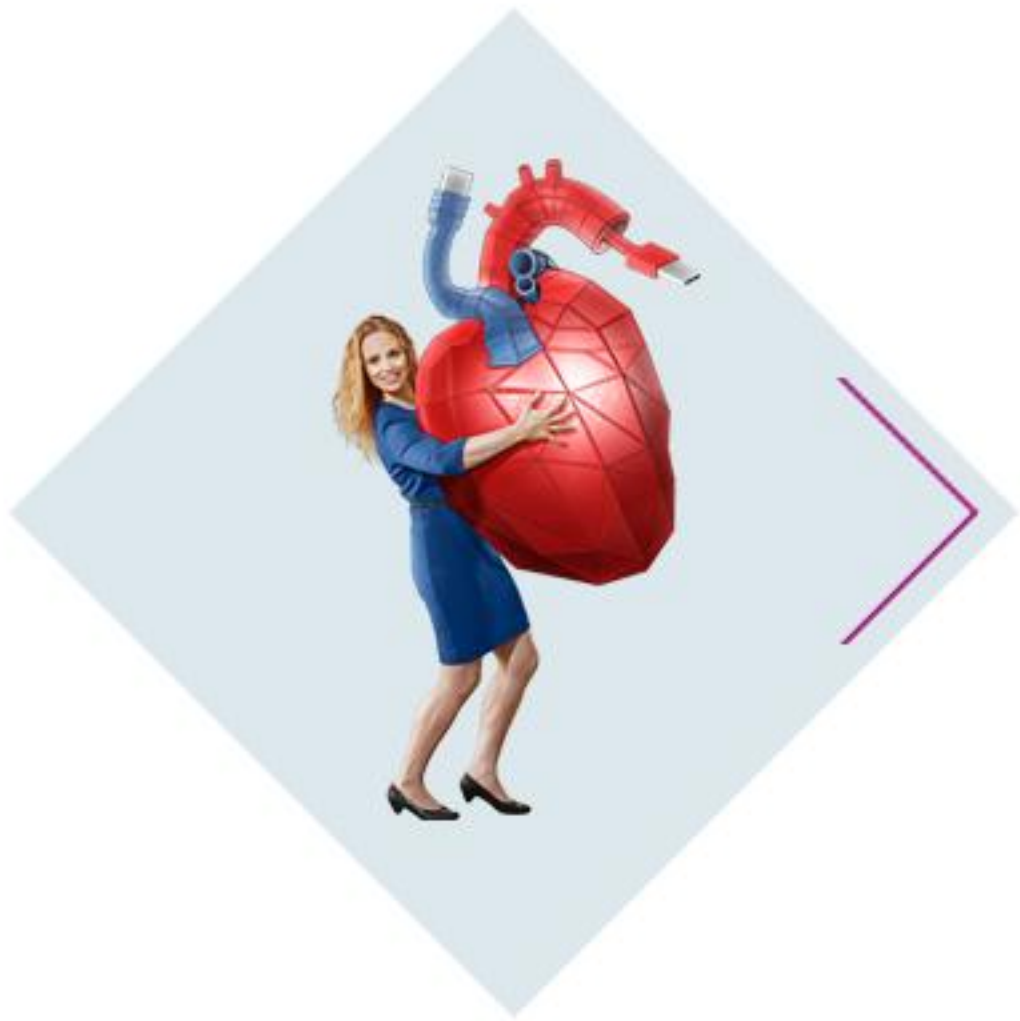
On ignore le statut

On ignore le “contenu de l’enveloppe” (Amélioration: alpha-shapes)

On ne regarde qu’un jour → l’aspect temporel est certainement crucial, mais “out of scope” de cette présentation

Sans analytique géographique, ces cas auraient été difficiles à évaluer

Extension possible : identifier les travailleurs qui ont (probablement) le même « patron »



Conclusions

Outils Open-source/ (partiellement) gratuits

Programmation :

Analytique : Python + Pandas + Geopandas/shapely

Visualisation (javascript) : leaflet.js, mapbox, google maps...

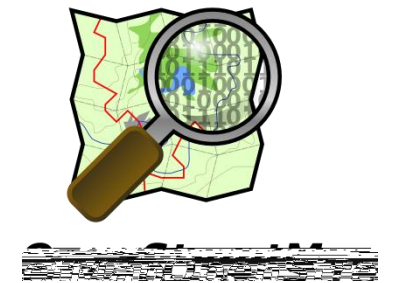
API : OSM, Here, Google maps...

Database : Postgresql + PostGIS (<https://postgis.net/>)

Desktop :

Cartographie, visualisation : QGIS (<https://www.qgis.org/>)

Analytique : GeoDa (<https://geodacenter.github.io/>)



Jointure spatiale :

Relie des polygônes/points en contact
DQ, enrichissement...

Autocorrélation spatiale :

Dimension géographique d'un indicateur : corrélation avec le voisinage
Statistiques, identification d'anomalies

Atteignabilité :

Territoire atteignable à partir d'un ensemble de point
Analyse d'impact d'une fermeture/ouverture de poste/antenne/...

Dispersion spatiale :

Identification d'étendue géographique anormale
Détection de fraude/anomalies

Conclusions

Données géographiques **très présentes** dans les données administratives

A combiner avec **des données ouvertes** (API, frontières...)

Ont des **spécificités** qui peuvent être exploitées

Ne concerne pas que les géographes/géologues/épidémiologistes/militaires !

Plus-value indéniable pour l'analyse :

- Statistiques

- Détections d'anomalies/fraudes

- Analyse d'impact (ouverture/fermeture d'antenne)

- Optimisation (tournées/affectation d'agents...)

Des outils **gratuits/Open-sources** sont souvent suffisants

Des ressources/données/idées/compétences à l'**IGN/NGI**: ne pas hésiter à les contacter!

À la portée de tout data-scientist/technicien curieux

Un peu de lecture

<https://www.smalsresearch.be/author/berten/#blogs>

Géocodage :

<https://www.smalsresearch.be/comparer-des-geocodeurs/>

<https://www.smalsresearch.be/geocodage-contourner-les-lacunes-dopenstreetmap-partie-1/>

<https://www.smalsresearch.be/geocodage-contourner-les-lacunes-dopenstreetmap-partie-2/>

Jointure spatiale :

<https://www.smalsresearch.be/la-jointure-spatiale-la-cle-de-lanalytique-geographique/>

Atteignabilité :

<https://www.smalsresearch.be/peut-on-toujours-atteindre-une-maternite-en-30-minutes/>

<https://www.smalsresearch.be/peut-on-toujours-atteindre-une-maternite-en-30-minutes-partie-2/>

Autocorrélation spatiale :

<https://www.smalsresearch.be/autocorrelation-spatiale-qui-se-rassemble-se-ressemble/>

La suite : sondage

Voulez-vous aller plus loin sur une thématique abordée durant présentation ?

Intérêt pour un second webinaire ? Sur quelle thématiques ?

Bases de données géographiques (PostGIS) : principes, possibilités, outils, exemples...

Géocodage: outils, comparaison de qualité...

Autre ?

Vandy BERTEN
vandy.berten@smals.be

Smals, ICT for society

02 787 57 11

Fonsnylaan 20 / Avenue Fonsny 20

1060 Brussel / 1060 Bruxelles