



Université de Rennes 1



RAPPORT DE GÉOMARKETING M2 MIAGE 2020 - 2021

Visualisation de données cachées réparties sur
une aire géographique

Equipe 8 :

- ❖ Mohamed EL MAHJOUB
- ❖ Mamadou Saidou Sall
- ❖ Emmanuelle Mbougou
- ❖ Morgane Bouchard
- ❖ Ali Aquil
- ❖ Issa Keita

Objectif :

En se basant sur leur nom, il faut déterminer si les différentes communes françaises sont plus influencées par la langue “oc” ou par la langue “oil”. Le résultat doit être sous forme de carte de France, avec des points de couleur pour représenter les différentes communes.

Résumé du travail effectué :

Pour commencer, nous avons discuté du sujet pour être sûr que tous les membres du groupe en aient la même compréhension. Ensuite, on a discuté des outils et langages que nous allions utiliser. Nous avons opté pour des algorithmes en python, et le logiciel Knime nous permettra de réaliser la superposition de la carte de France et du repère cartésien.

Ensuite, nous avons récupéré les données qui allaient nous être utiles. Sur ces données, nous avons appliqué un algorithme. Ce dernier enlève les données superflues, et ajoute une colonne en fonction de si la commune est plus oil ou oc. Cette dernière colonne prendra la valeur -1, 1 ou 0.

Travail réalisé:

I - Description des sources de données et traitement des données

Pour réaliser notre projet, nous avons trouvé les sources de données sur lesquelles effectuer notre étude. Nous avons identifié différentes sources de données dont :

- la liste des communes et de leurs coordonnées GPS récupérées sur la plateforme data.gouv. Ces données contiennent le code Insee de la commune, le nom de la commune, le code postal, le libellé de la commune et les données GPS de chaque commune.
- la liste des adjectifs en français, anglais, espagnol et italien récupérés sur différents sites (voir source) pour former un fichier regroupant la plupart de ces adjectifs.
- la liste des suffix oil et oc depuis site lefigaro

Pour le traitement des données, nous avons préféré nous orienter vers un programme Python, L'avantage de Python c'est la simplicité et la rapidité de développement, c'est un langage très facile à apprendre et très rapide à mettre en place. L'algorithme pour déterminer si une commune est Oil ou Oc fonctionne comme suit :

- si l'adjectif est préfixe d'une commune -1 = Oil
- si l'adjectif est suffixe d'une commune alors elle est OC avec L (score) = 1
- si le suffixe est un suffixe de commune dite oil (selon le figaro) alors la commune est oil avec L = -1
- si le suffixe est un suffixe de commune dite oc alors la commune est OC avec L = 1

Nous avons pu indexer plus de 12 000 OïL et OC au total.

```
total : 35100
oil : 6049
oc : 6167
neutre : 22884
```

Figure 1 : Résultat de L'indexation

II - Visualisation

1- Recherche du meilleur outil de cartographie

	Avantages	Inconvénients
OpenStreetMap (Extension dans Knime)	<ul style="list-style-type: none">• Les cartes sont très détaillées et à jour• Solution open source• gain de place (stockage)• Maîtrise de l'outil Knime	<ul style="list-style-type: none">• Rendu visuel pauvre• Pas de couche satellite ni de version retina clé en main
QGIS	<ul style="list-style-type: none">• Libre et Open source• Nombreuse fonctionnalités• Disponible sous Linux, Windows et Mac.	<ul style="list-style-type: none">• Formation à faire sur le logiciel• Stockage de l'espace du logiciel

Grâce à cette étude comparative, nous avons décidé d'utiliser Knime pour cartographier notre carte de France.

2- Liaison entre nos données et la Carte

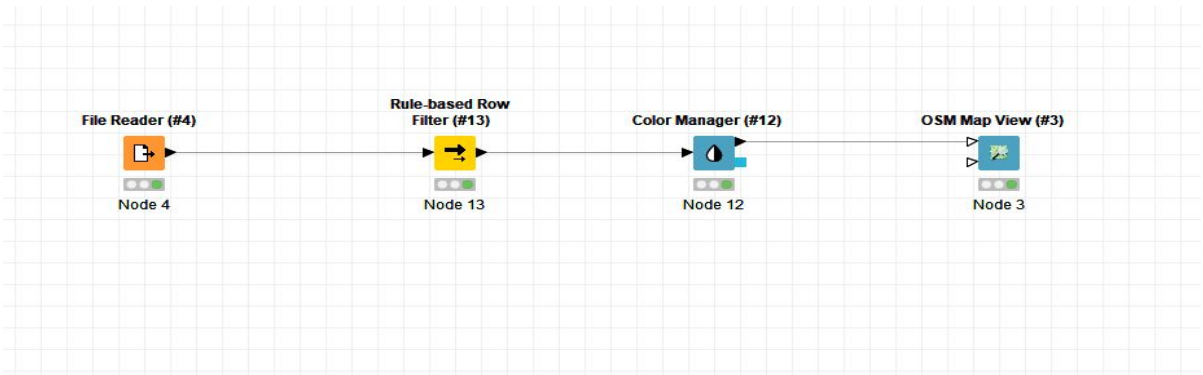


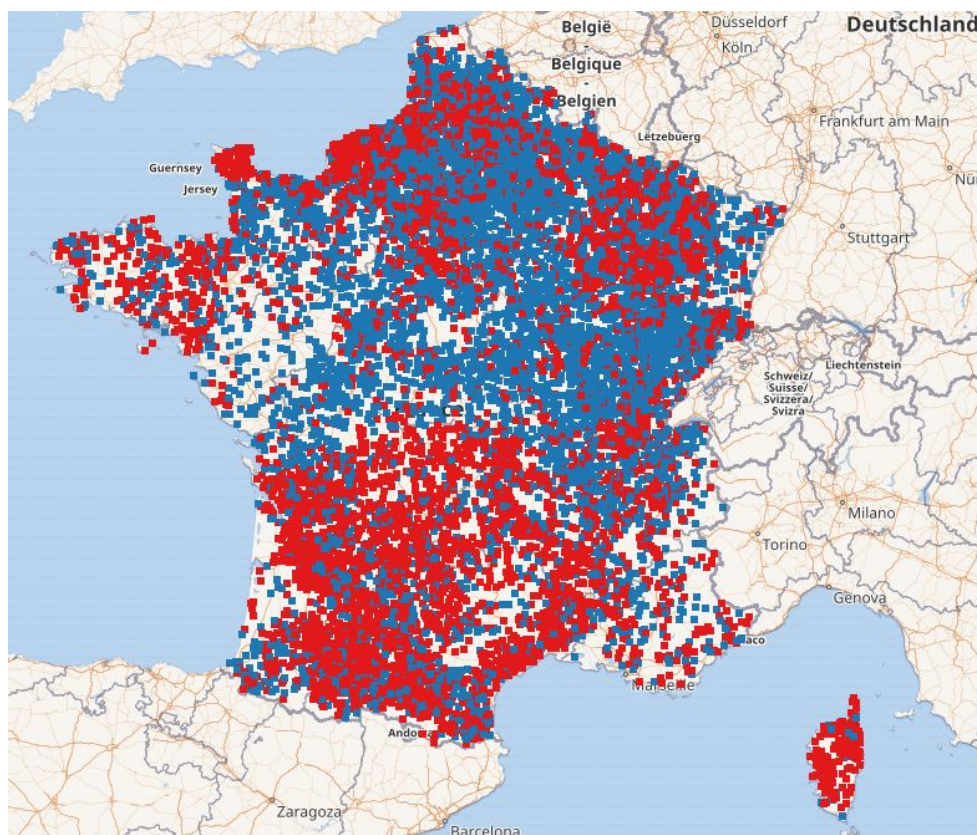
Figure 2 : Workflow Knime pour obtenir la carte finale

Grâce au logiciel Open source KNIME nous avons pu clairement visualiser nos données.

- Le noeud “File Reader” est utilisé pour définir le nouveau csv comme source de données;
- Le noeud “Rule-based Row Filter” nous permet d’effectuer un filtrage sur les communes indeterminées.
- Le noeud “Color Manager” est utile pour affecter une commune à une couleur.
- Le noeud “OSM Map View” (Extension d’OpenStreetMap) nous permet de visualiser notre carte .

III - Résultat final

Carte de la France



Bibliographie

<https://www.data.gouv.fr/fr/datasets/base-officielle-des-codes-postaux/>

<http://www.lepetitjournaldesprofs.com/charivari/wp-content/uploads/sites/6/2016/10/dubois-buysecharivari.xls>

<https://www.lefigaro.fr/fig-data/nom-communes/>

https://www.mobap.edu/wp-content/uploads/2013/01/list_of_adjectives.pdf