

```
# Chargement des packages
```

```
library(stringi)
```

```
library(dplyr)
```

```
##
```

```
## Attachement du package : 'dplyr'
```

```
## Les objets suivants sont masqués depuis 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## Les objets suivants sont masqués depuis 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(knitr)
```

```
library(kableExtra)
```

```
##
```

```
## Attachement du package : 'kableExtra'
```

```
## L'objet suivant est masqué depuis 'package:dplyr':
```

```
##
```

```
##      group_rows
```

```
library(tidyr)
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
```

```
## v forcats   1.0.0      v readr     2.1.5
```

```
## v ggplot2   3.5.1      v stringr  1.5.1
```

```
## v lubridate 1.9.4      v tibble   3.2.1
```

```
## v purrr     1.0.4
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter()      masks stats::filter()
```

```
## x kableExtra::group_rows() masks dplyr::group_rows()
```

```
## x dplyr::lag()         masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(summarytools)
```

```
##
```

```
## Attachement du package : 'summarytools'
```

```
##
```

```
## L'objet suivant est masqué depuis 'package:tibble':
```

```
##
```

```
##      view
```

```
library(gridExtra)
```

```
##  
## Attachement du package : 'gridExtra'  
##  
## L'objet suivant est masqué depuis 'package:dplyr':  
##  
##      combine
```

```
library(purrr)  
library(skimr)  
library(spdep)
```

```
## Le chargement a nécessité le package : spData  
## To access larger datasets in this package, install the spDataLarge  
## package with: 'install.packages('spDataLarge',  
## repos='https://nowosad.github.io/drat/', type='source')'  
## Le chargement a nécessité le package : sf  
## Linking to GEOS 3.10.2, GDAL 3.4.1, PROJ 8.2.1; sf_use_s2() is TRUE
```

```
library(geosphere)
```

Analyse descriptive

Description de la population

Pyramide des ages

Notre population d'étude est une population assez homogène en matière d'âge. Cependant plus on dépasse les 75 ans et moins on rencontre de personnes. D'autres part notre population est fortement masculine avec une forte proportion des hommes quelle que soit la tranche d'âge à l'exception des tranches du troisième âge.

Taux de natalité et taux de mortalité

Dans les communes étudiées, le taux de natalité et de mortalité sont un peu élevées avec la plupart des taux variant entre 5 et 15 pour 1000 en ce qui concerne la natalité et 0 et 20 pour 1000 pour la mortalité. On remarque une corrélation négative entre ces deux taux. Néanmoins cette corrélation n'a à priori aucun sens. Par ailleurs, l'observation des distributions permet de constater que la natalité est de façon générale élevée par rapport à la mortalité dans les communes étudiées.

En vue de mieux de mieux voir peut être l'effet de la mortalité sur la natalité, nous allons nous intéresser alors à une analyse de la corrélation entre les deux taux par groupe d'âge. Nous allons considérer les groupes d'âge suivants : 0-24, 25-44, 45-60, 60 et plus en fonction des variables disponibles et aussi à partir de l'information sur l'âge des femmes en âge de procréer qui est de l'ordre de 25-45 et des personnes âgées dont l'âge est de plus de 60 ans. Ne disposant pas du taux de mortalité dans chaque groupe, alors nous avons dans notre analyse opté plutôt pour le pourcentage des femmes de chaque groupe en partant du principe que la natalité est très souvent liée aux femmes et du fait que nous pouvons analyser une diminution du pourcentage comme étant dû à une mortalité. Ainsi sur la base de cette nouvelle hypothèse, voici nos nouveaux résultats.

Les résultats nous montrent un lien croissant pour les tranches d'âge 0-24 et 25-45 ans montrant ainsi que dans ces tranches d'âge si le pourcentage des femmes diminue (ce qu'on pourrait assimiler à une mort des

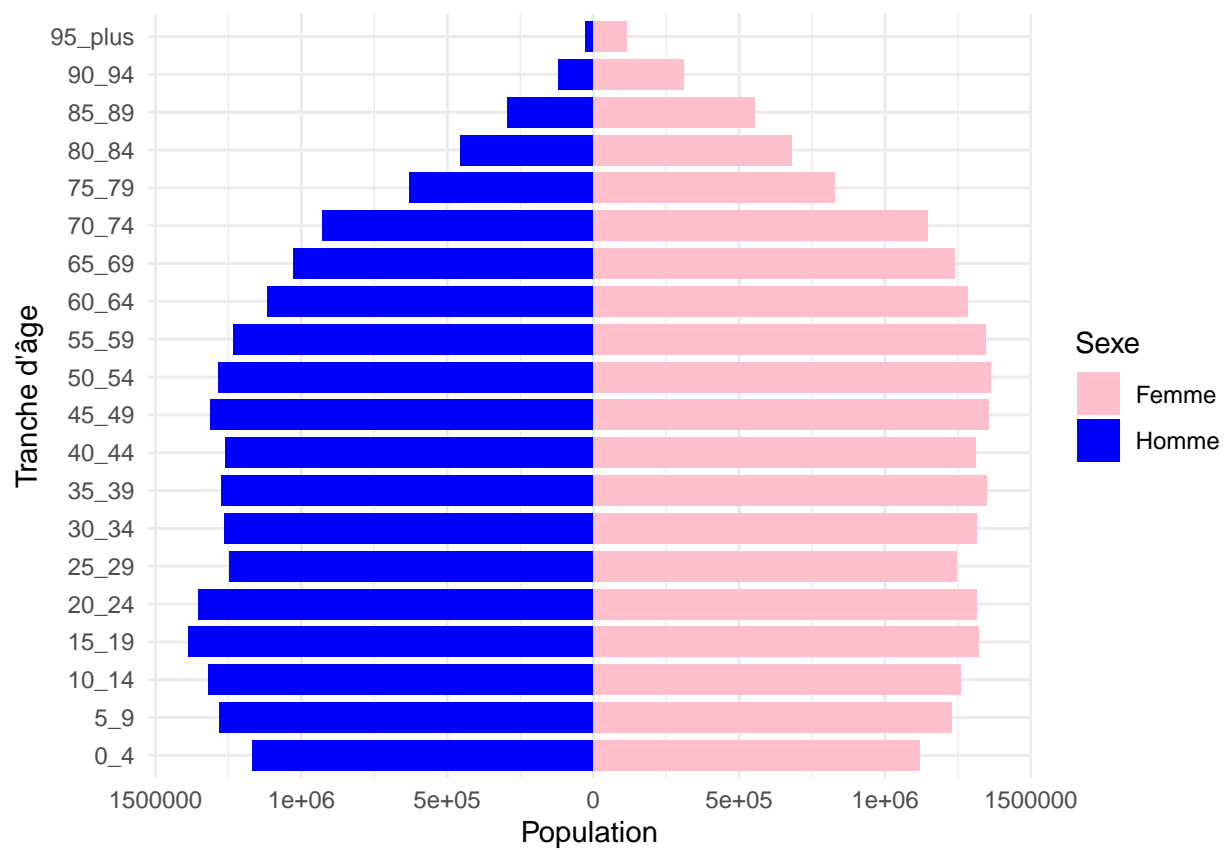


Figure 1: Pyramide des âges

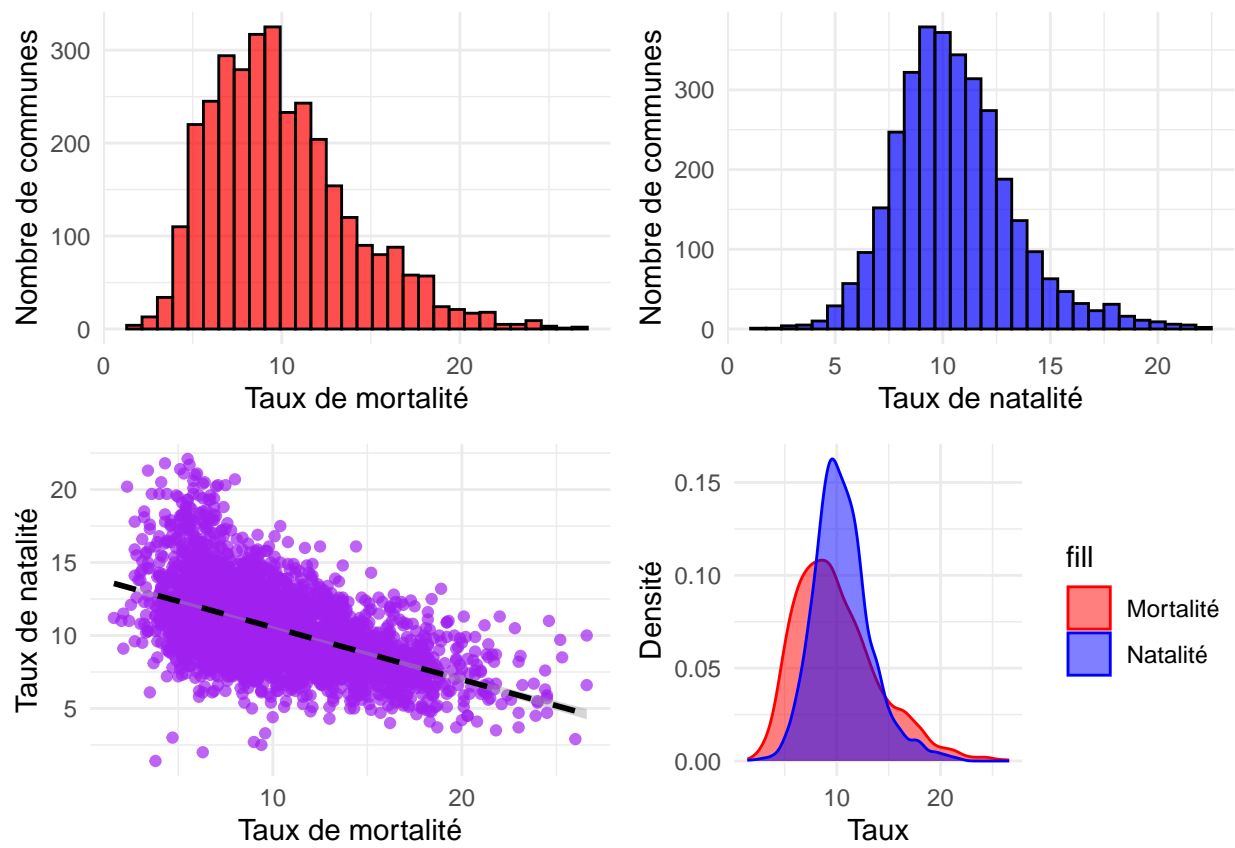


Figure 2: Taux de Natalité et Taux de Mortalité

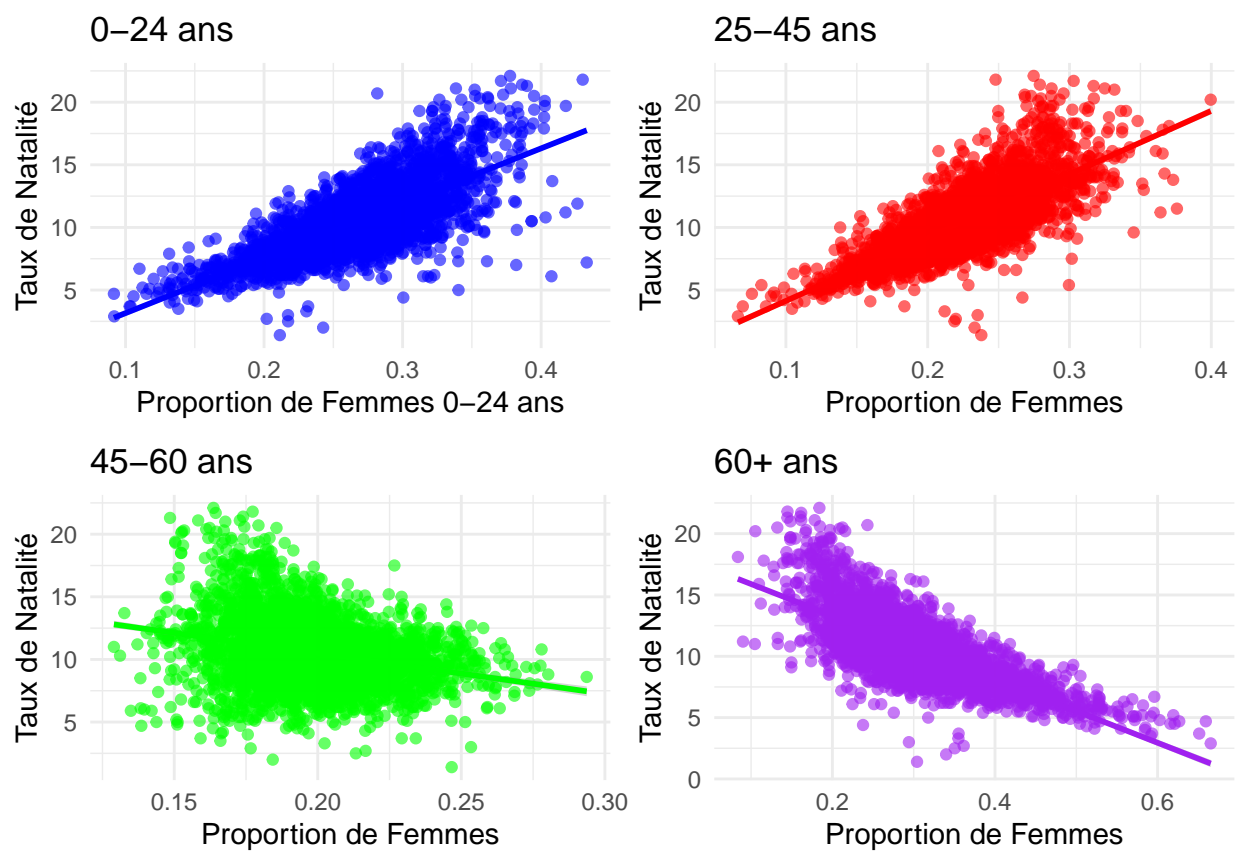


Figure 3: Taux de Natalité et Pourcentage des femmes dans chaque groupe

femmes) alors le taux de natalité diminue. Par ailleurs ceux de la tranche 45-60 semblent n'avoir aucun lien sur le taux de natalité. Enfin il a été constaté un lien négatif pour la tranche d'âge 60 ans et plus.

Taux et nombre de consultation

(Insérer les cartes à ce niveau : Richard doit refaire les cartes et les insérer)

L'analyse des statistiques descriptives sur le nombre de consultations annuelles de médecin généraliste entre 2018 et 2022 révèle une distribution fortement asymétrique à droite, avec une grande dispersion des données. La moyenne de 19130 consultations, nettement supérieure à la médiane de 9127, indique la présence de valeurs extrêmes tirant la distribution vers le haut. Cette asymétrie est confirmée par l'écart considérable entre le minimum de 1037 et le maximum de 765833 consultations par an.

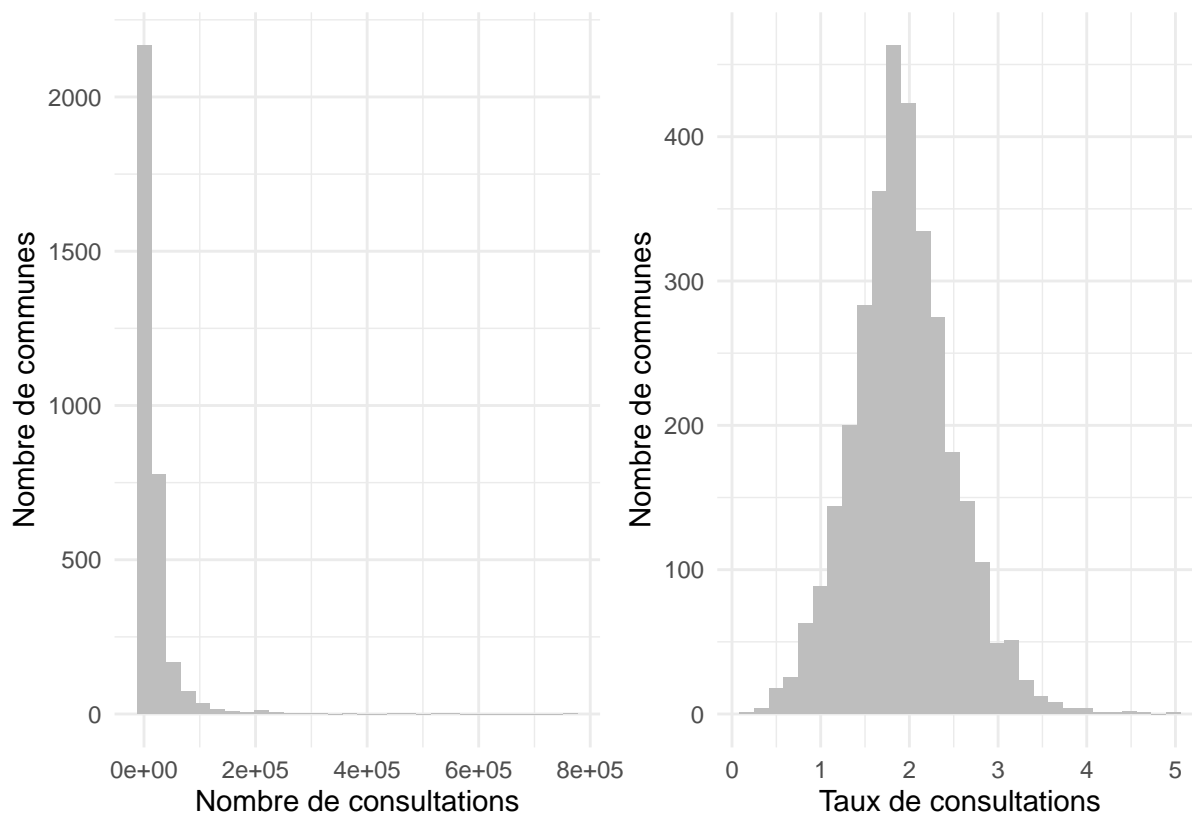


Figure 4: Distribution du nombre et du taux de consultations

La moitié des médecins généralistes effectuent entre 5993 et 17290 consultations annuellement, ce qui suggère une variabilité importante dans la charge de travail. La médiane de 9127 consultations par an, équivalant à environ 25 consultations par jour ouvrable, semble plus représentative de l'activité typique d'un médecin généraliste que la moyenne influencée par les valeurs extrêmes. Ces statistiques mettent en lumière la diversité des pratiques et des charges de travail parmi les médecins généralistes, avec potentiellement quelques cas atypiques présentant un volume de consultations exceptionnellement élevé.

Le nombre de visites pouvant potentiellement être influencé par la taille de la commune et donc par sa population, nous avons éliminé cet effet en calculant le taux de consultations qui n'est autre que le nombre de consultations moyennes par personnes.

Table 1: Résumé statistique du nombre de visites

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1037	5993	9127	19129.63	17290	765833

Taux de consultation et quelques variables démographiques et socio-économiques

Nous allons ici, voir s'il y a un lien à priori entre le taux de consultation et certaines de nos variables explicatives. Ainsi, nous avons d'abord réalisé une analyse descriptive bivariée puis nous avons calculé la corrélation de Pearson pour évaluer le lien linéaire entre le taux de consultation et des variables telles que la population totale, la part des personnes âgées (75 ans et plus), la part de quelques CSP (ouvriers et retraités).

Taux de consultation et population totale

En divisant les communes en trois groupes égaux (ou presque égaux) en fonction de la population totale, il ressort un lien clair entre la taille des communes françaises et le taux de consultations médicales, mettant en évidence une tendance où les grandes communes (> 8974 habitants) affichent un taux moyen de consultations supérieur (1,526810) par rapport aux communes moyennes (1,456356) et petites (1,383861). Cette observation suggère que l'accès facilité aux infrastructures médicales dans les zones urbaines contribue à une utilisation accrue des services de santé. En revanche, les petites communes, probablement plus isolées et moins dotées en praticiens, semblent rencontrer des barrières structurelles limitant la fréquence des consultations.

Table 2: Taux de consultations selon la taille de la commune

taille_commune	Taux de consultations
Grande (> 8974)	1.526810
Moyenne (4849 - 8974)	1.456356
Petite (≤ 4848)	1.383861

Taux de consultation et population âgée

L'analyse met en évidence que les communes françaises avec une population âgée significative (population âgée de 75 ans et plus est supérieure à la médiane soit plus de 670 habitants âgés de 75 ans et plus) présentent un taux moyen de consultations inférieur (1,410213) comparé aux communes où la population âgée est moindre (1,501111). Cette observation peut refléter des défis spécifiques aux populations plus âgées, tels que des obstacles physiques ou logistiques pour accéder aux soins médicaux, ou encore une moindre propension à consulter régulièrement en raison d'habitudes ou de conditions de santé chroniques. Ces résultats soulignent un paradoxe apparent, car les besoins en soins médicaux des personnes âgées sont en général plus importants, ce qui pourrait indiquer une inadéquation entre l'offre médicale et les besoins spécifiques de cette tranche d'âge. Cela met en lumière un enjeu crucial pour les politiques de santé visant à améliorer l'accès et l'utilisation des services médicaux pour les populations vieillissantes.

Table 3: Taux de consultations selon la population âgée

population_agee_importante	consultations_moyennes
Non (≤ 670)	1.501111
Oui (> 670)	1.410213

Taux de consultation et CSP

Aucune catégorie ne semble montrer une relation linéaire évidente avec le taux de visite. Par ailleurs, pour toutes les catégories socio-professionnelles, la majorité des communes se situent dans une plage de proportions faibles, ce qui limite la variabilité observable dans les relations. Une analyse statistique supplémentaire, comme le calcul de corrélations, serait nécessaire pour confirmer ou infirmer les relations observées visuellement.

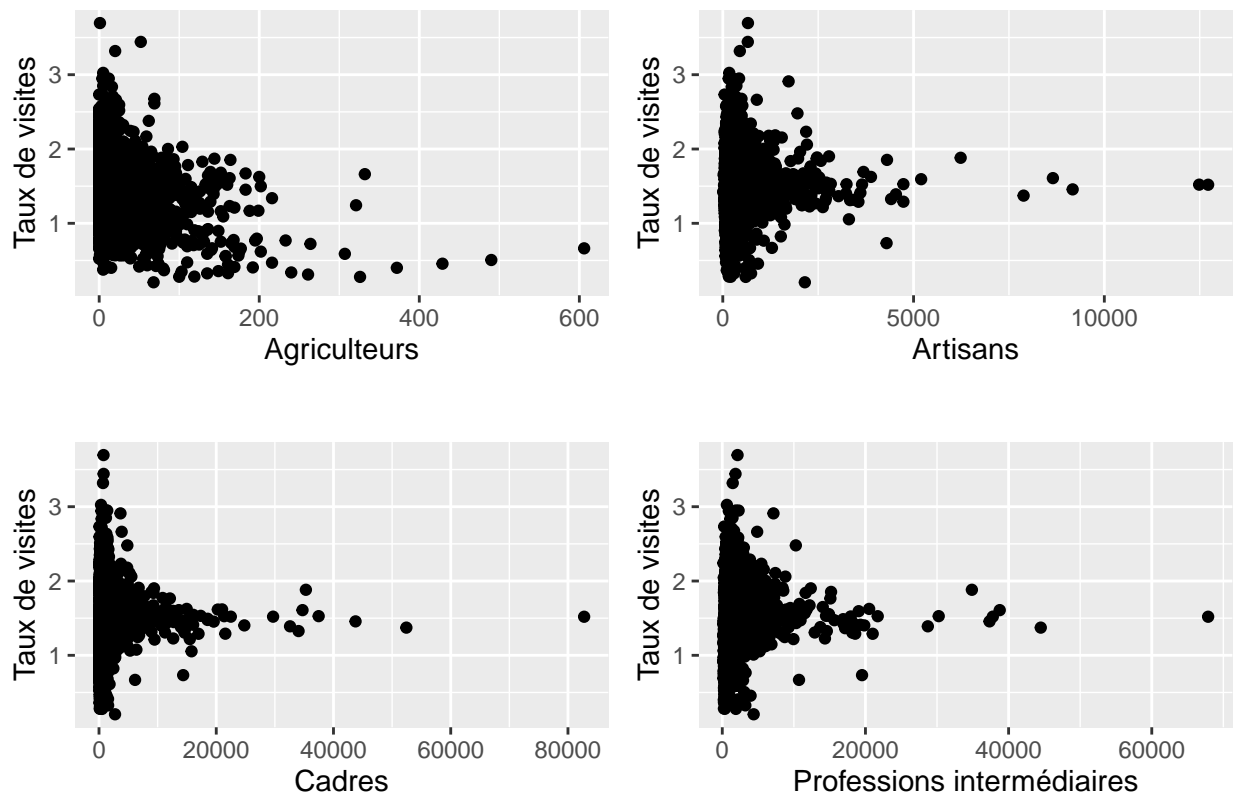


Figure 5: Relations entre le taux de consultations et certaines catégories socioprofessionnelles

(Trouver un titre pour cette section et analyser : Alex)

Analyse spatiale

Après une analyse de nos données en ne tenant pas compte de l'effet spatial, nous allons à présent poursuivre avec une analyse qui tient compte de celui-ci. Nous allons ici faire une analyse basée sur le diagramme de Moran. Le diagramme de Moran permet d'analyser la structure spatiale du taux de consultations. Il est

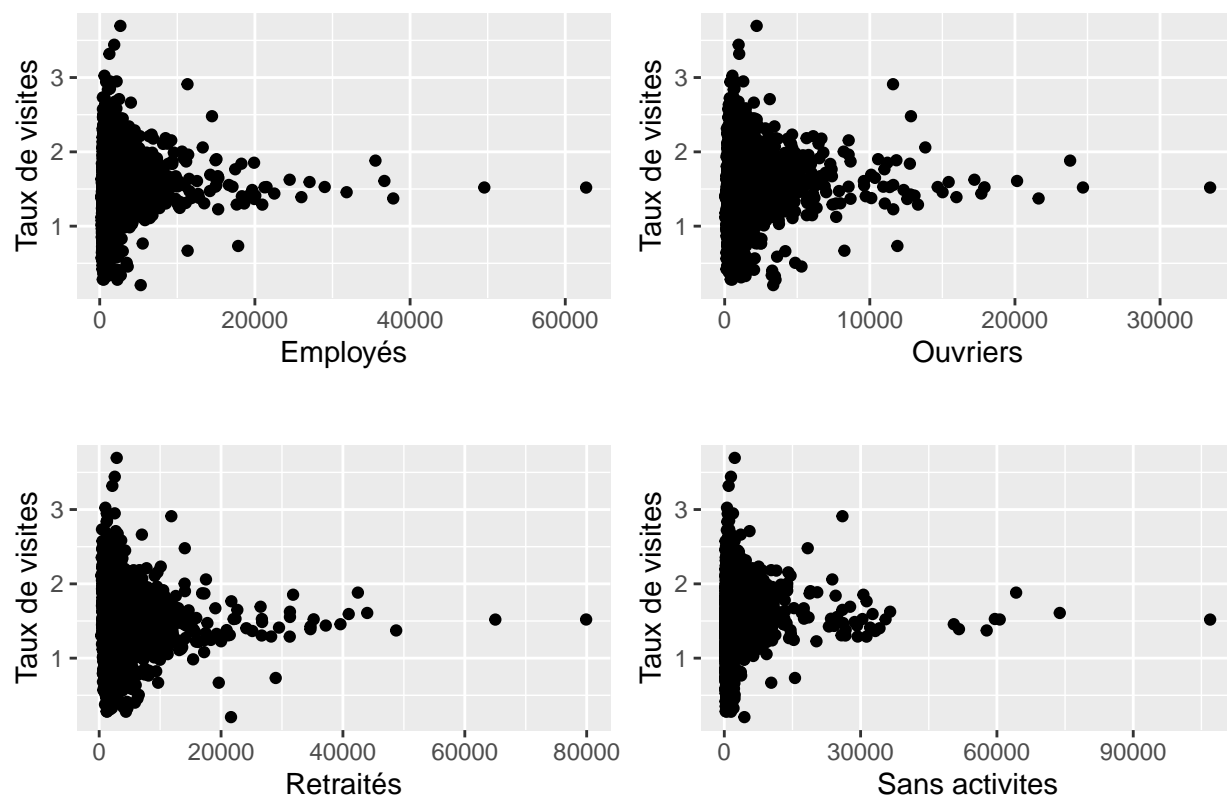


Figure 6: Relations entre le taux de consultation et certaines catégories socioprofessionnelle

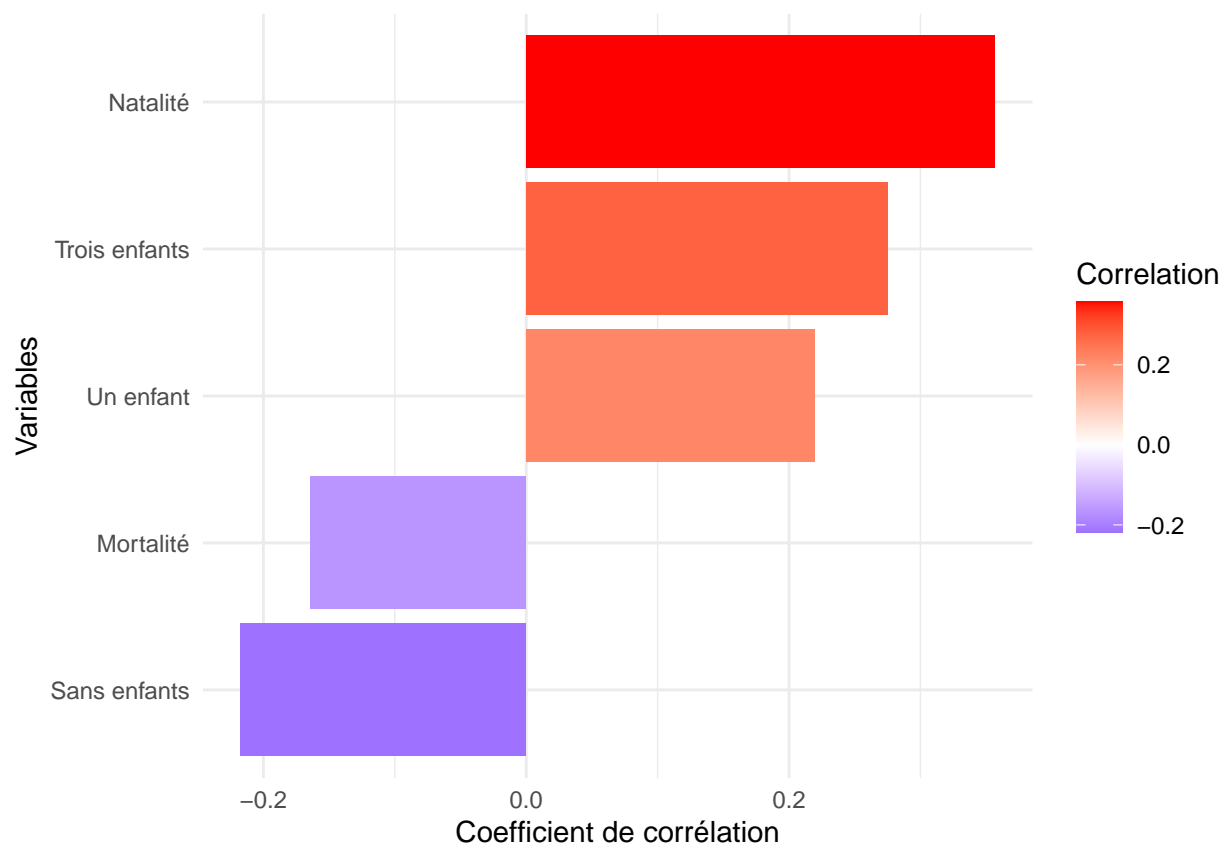


Figure 7: Corrélations entre le nombre de visite et quelques variables

constitué d'un nuage de points où l'axe des abscisses représente les valeurs centrées du taux de consultations (yy), et l'axe des ordonnées montre les valeurs moyennes du taux de consultations pour les zones voisines (Wyy), avec W étant la matrice de poids normalisée. En raison de la centration de yy et de la normalisation de W , la moyenne empirique de Wyy est égale à zéro.

Le diagramme inclut aussi une droite de régression linéaire entre Wyy et yy , ainsi que les lignes $y = 0$ et $Wyy = 0$, qui divisent l'espace en quatre quadrants.

- **Quadrant 1 (haut à droite)** : Les zones avec un taux de consultations plus élevé que la moyenne, entourées de zones présentant également un taux de consultations élevé (autocorrélation spatiale positive, structure **high-high**).
- **Quadrant 3 (bas à gauche)** : Les zones avec un taux de consultations plus faible que la moyenne, entourées de zones présentant un taux de consultations également faible (autocorrélation spatiale positive, structure **low-low**).
- **Quadrant 2 (bas à droite)** : Les zones avec un taux de consultations plus élevé que la moyenne, mais entourées de zones présentant un taux de consultations plus faible (autocorrélation spatiale négative, structure **high-low**).
- **Quadrant 4 (haut à gauche)** : Les zones avec un taux de consultations plus faible que la moyenne, mais entourées de zones avec un taux de consultations plus élevé (autocorrélation spatiale négative, structure **low-high**).

La densité des points dans chaque quadrant permet de visualiser la structure spatiale dominante du taux de consultations. Ce diagramme aide aussi à identifier les zones atypiques, qui ne suivent pas la structure spatiale générale, facilitant ainsi l'analyse des résultats que vous obtiendrez par la suite. Vu le nombre de nos communes la mise sur le graphique des noms de toutes les communes allaient être compliqué. Pour cela, nous avons choisi au hasard 4 communes par cadrants. Ainsi, Ce diagramme de Moran illustre la corrélation spatiale des taux de consultation par commune et ceux des communes voisines. La tendance générale, représentée par la droite de régression rouge, montre une relation positive entre ces taux, confirmant ainsi une autocorrélation spatiale. Autrement dit, les communes ayant un taux élevé de consultations ont tendance à être entourées par d'autres communes avec un taux similaire, et inversement.

Dans le quadrant HH (Haut-Haut), représenté en rouge, on retrouve des communes comme Saint-Saulve, Le Bouscat et Bruay-la-Buissière. Celles-ci affichent un taux de consultation élevé et sont entourées par des communes présentant également des taux élevés. Cela indique une concentration géographique des consultations médicales, qui peut s'expliquer par une offre de soins plus développée ou une demande locale particulièrement forte.

Le quadrant HL (Haut-Bas), en violet, comprend des communes comme Saint-Jacques-de-la-Lande et Le Portel. Ces communes ont un taux élevé de consultations, mais sont entourées de communes où les taux sont plus faibles. Ce contraste peut suggérer que ces villes disposent d'une offre de soins plus attractive que leurs voisines, attirant ainsi des patients des alentours.

À l'inverse, dans le quadrant LH (Bas-Haut), en vert, on trouve des communes comme Aix-en-Provence, Ribécourt-Dreslincourt et L'Isle-Adam. Ces communes affichent un faible taux de consultation, tandis que leurs voisines présentent des taux plus élevés. Ce phénomène peut s'expliquer par le fait que les habitants de ces villes se rendent dans les communes avoisinantes pour leurs consultations, soit en raison d'un manque d'infrastructures médicales locales, soit par préférence pour des services situés ailleurs.

Enfin, le quadrant LL (Bas-Bas), en orange, inclut des communes comme Saint-Jean-le-Blanc, Romorantin-Lanthenay et Le Breuil. Ces villes ont un faible taux de consultation et sont entourées de communes où les taux sont également bas. Cela peut indiquer une accessibilité réduite aux soins de santé, une moindre densité médicale, ou encore une faible demande locale pour des consultations.

En résumé, cette analyse met en évidence des disparités territoriales dans la répartition des consultations médicales. Certaines communes concentrent les services et attirent les patients des alentours, tandis que d'autres souffrent d'un accès limité aux soins, renforçant ainsi les inégalités spatiales en matière de santé.

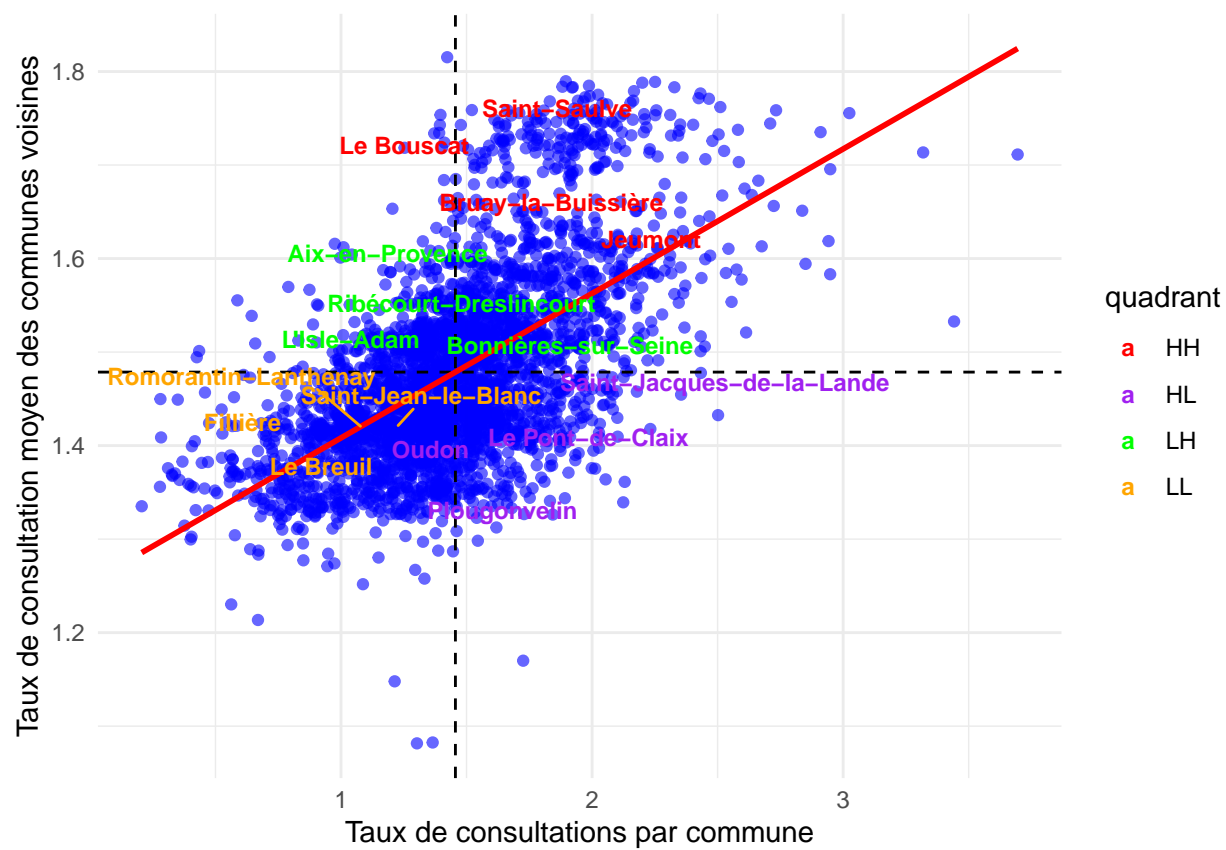


Figure 8: Moran Plot avec visibilité de quelques communes

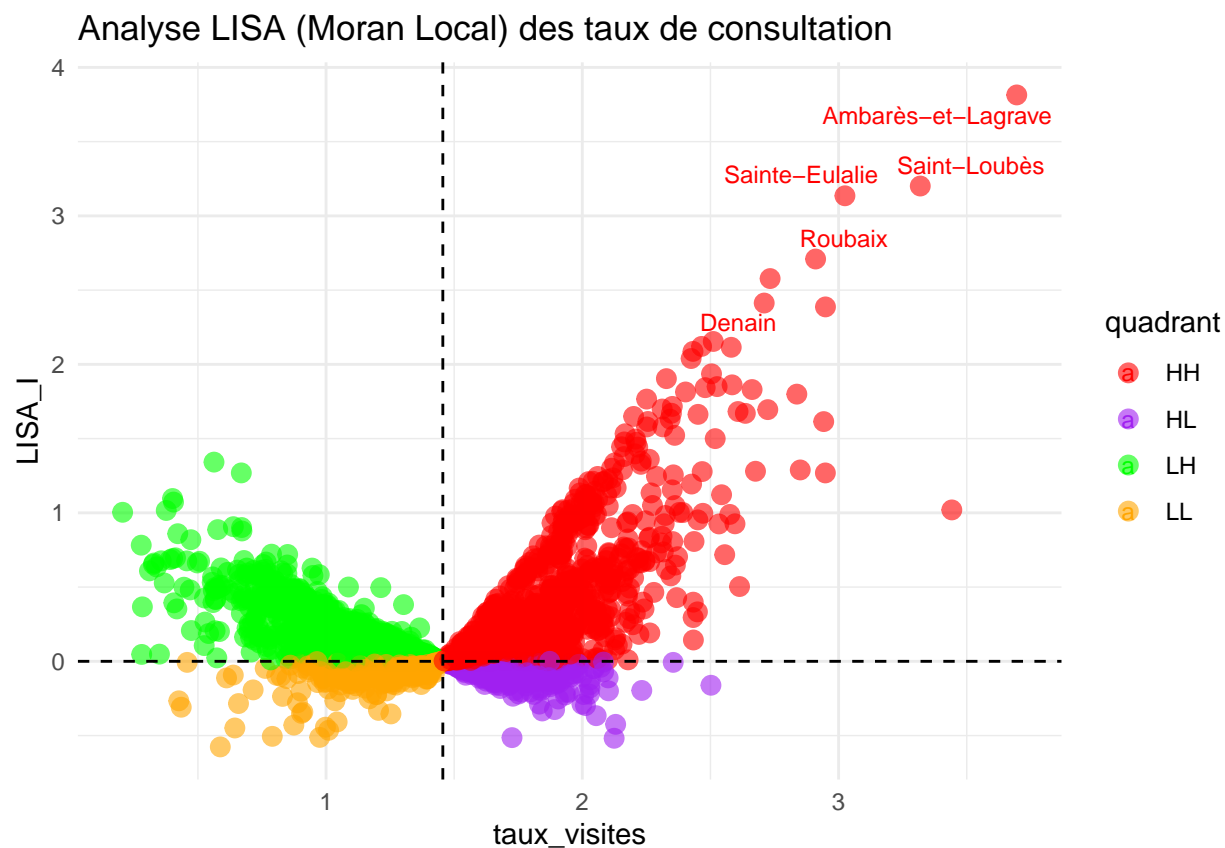


Figure 9: Clusters sur la base du LISA

En se basant sur l'analyse par cluster fourni pour le LISA (local Indicator or Spatial Association), nous pouvons remarquer que le cluster HH est celui regroupant le plus de communes suivi du cluster LH. Ceci dit une grande partie des communes ont des taux élevés et entourées par des communes de taux élevés ou encore des taux bas et entourées par des communes de taux élevés.