

Analyse quantitative des dépenses des soins de ville

Niccolo Borgioli

January 2023

1 Analyse des composantes de la hausse des dépenses de soins de ville entre 2010 et 2019.

Au niveau français, les soins de ville voient leur valeur augmenter constamment entre 2010 et 2019 passant de 40.3 à 56.6 milliards d'euros de dépenses (augmentation de +40%). Si l'évolution est plutôt constante d'année en année, aux alentours de +3%, on remarque une forte hausse entre 2018 et 2019 (+9%). Les soins de villes, dans leur entièreté entre 2010 et 2019, ont représenté une dépense considérable de 467 milliards d'euros : celle ci est majoritairement due aux Prestations en espèces et Spécialistes libéraux qui représentent à eux seuls 190 milliards d'euros, i.e. 41% de la dépense totale. Suivent ensuite les Omnipraticiens libéraux, Infirmiers libéraux et LPP, chacun contribuant pour 10% de la dépense totale. On constate enfin que tous les postes de dépense sont en hausse constante d'année en année.

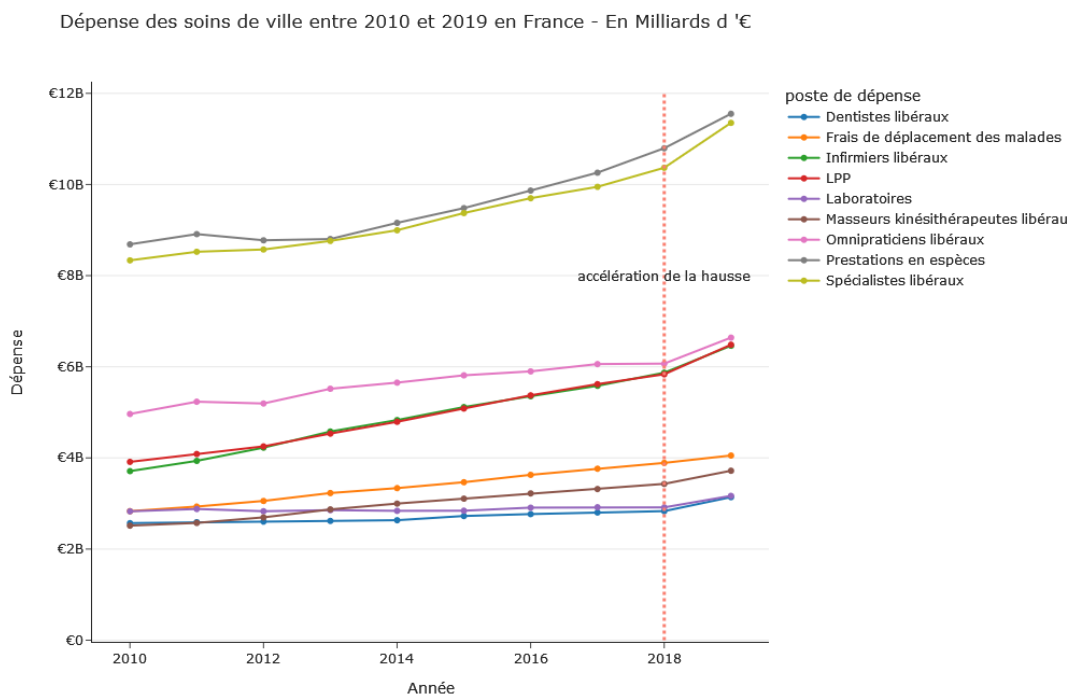


Figure 1: Evolution des dépenses en soins de ville. Les Prestations en espèces et Spécialistes libéraux représentent les dépenses les plus importante chaque année

En analysant plus en détail les dépenses de soins de ville par département (cf. Figure 2a) on remarque que deux départements se distinguent : les Bouches-du-Rhône avec des dépenses constamment en hausse et atteignant les 2.6 milliards d'euros en 2019 et le département du Nord avec une croissance similaire et 2.5 milliards d'euros en 2019. Suit ensuite Paris, chaque année, a la troisième place avec une dépense bien inférieure (60% seulement par rapport aux Bouches-du-Rhône en 2019) et le Rhône qui apparaît exceptionnellement dans le top-3 en 2018.

En s'intéressant plus spécifiquement aux dépenses de soins majoritaires par département (cf. Figure 2b) on retrouve les Prestations en espèces en tant que première dépense de soins pour environ 60 des 96 départements de France chaque année, le chiffre étant très stable entre 2010 et 2019. S'en suivent les dépenses liées aux Spécialistes libéraux comme étant la première dépense pour environ 34 départements avec une légère baisse en 2019. Moins de

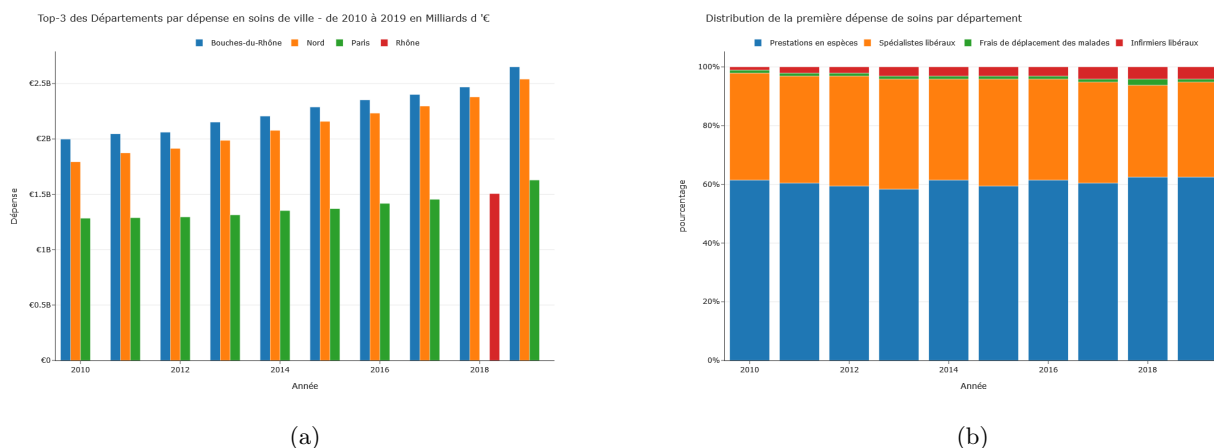


Figure 2: (a) Les Bouches-du-Rhône, le Nord et Paris sont généralement les départements avec les plus grandes dépenses (b) Les Prestations en espèces et les Spécialistes libéraux sont les premières dépenses de soins par départements chaque année

5% des départements ont une dépense primaire autre que les Prestations en espèces et les Spécialistes libéraux (les Frais de déplacement des malades et Infirmiers libéraux). De l'autre côté du spectre, on observe chaque année des départements avec des dépenses bien plus basses avec des valeurs de quelques dizaines de millions d'euros seulement. Cependant, même lorsqu'on s'intéresse aux dépenses minimales, on constate toujours une hausse des dépenses, par exemple en Lozère, avec une moyenne mobile de ses dépenses, sur les trois dernières années, ayant augmenté de 24% entre 2012 et 2019 (58 millions d'euros en 2019). En conclusion, on constate une hausse significative des soins de ville entre 2010 et 2019 avec les Prestations en espèces et Spécialistes libéraux étant les dépenses les plus importantes chaque année. Au niveau départemental, les Bouches-du-Rhône et le Nord engendrent les plus grandes dépenses et la Lozère clôt le classement avec une dépense 4x fois inférieure par rapport aux Bouches-du-Rhône. Cela démontre une hétérogénéité des dépenses en soins de ville entre les départements qui peut potentiellement s'expliquer en prenant en compte des données supplémentaires telles que les estimations de population par département, sexe et tranche d'âge ainsi que le nombre d'ALD (affections de longue durée) par département et par type d'ALD.

2 Analyse de l'hétérogénéité de la demande en soins de ville entre départements pour l'année 2019 à partir des données démographiques et des données d'affectation de longue durée pour cette même année

Après avoir uni les données liées au nombre de ALD et aux estimations de populations par département, il a été jugé judicieux d'inclure aussi le total, en 2019, des dépenses en soins de ville par département à partir du premier jeu de données. Ceci permet de définir une proxy à la demande en soins, sous l'hypothèse que les dépenses sont fortement corrélées à la demande. Dans ce but, les observables correspondantes aux DOM ont été écartées et devraient faire l'objet d'une analyse séparée en vue de leur situation géographique et démographique spécifique. Afin d'expliquer l'hétérogénéité de la demande en soins XgBoost a été choisi en tant que modèle d'apprentissage statistique, avec pour tâche de prédire la demande en soins (les dépenses totales par département en 2019, par hypothèse). Cette tâche de Régression a été transformée en une tâche de classification en créant quatre catégories de dépense : dépense totale pour un département de moins de 0.5 milliards, entre 0.5 et 1 milliards, entre 1 et 1.5 milliards et plus de 1.5 milliards. En utilisant le total des soins de ville en tant que variable dépendante et en construisant des catégories de dépenses de façon telle à avoir assez d'observables dans chacune des catégories on entraîne le modèle d'apprentissage. Même avec un training set très réduit de seulement 76 observables et un test set de 26 observables le modèle arrive à prédire correctement la catégorie de dépense 85% des fois. Plus précisément, en s'appuyant sur la matrice de confusion, on remarque que seules 3 observables donnent des prédictions fausses. De plus, la matrice de confusion nous montre aussi que des exemples pour chaque classe apparaissent dans le test set, ce qui nous indique que l'algorithme peut s'appuyer sur toutes les critères appris pour faire son choix. Trois critères sont identifiés par l'algorithme :

- Le nombre d'hommes de plus de 75 ans est inférieur à 23k
- Le nombre de femmes de 40 à 59 est moins que 143k
- L'ensemble, hommes et femmes, de 40 à 59 est inférieur à 160k

Distribution des départements par catégorie de dépense en fonction de la séniorité des individus

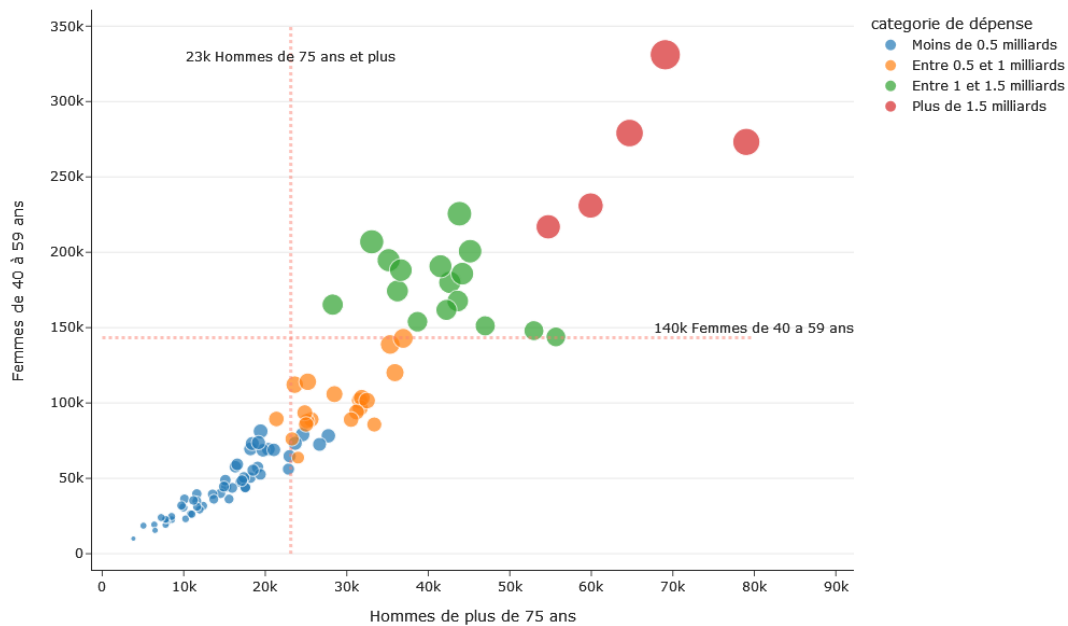


Figure 3: Les différents départements par catégorie de dépense apparaissent délimitées clairement en considérant hommes et femmes dans des tranches d'âge spécifiques

Ces critères sont vérifiables dans la Figure 3 : l'hétérogénéité de la demande s'explique lorsqu'on s'intéresse au nombre de personnes au-dessus de 75 ans, en fonction du nombre de femmes entre 40 et 59 et aussi en fonction de l'ensemble, femmes et hommes entre 40 et 59 ans (taille des cercles). On remarque clairement quatre clusters, chacun correspondant à une catégorie de dépense. En connaissant le nombre d'hommes seniors dans un département et le nombre de femmes dans une tranche d'âge spécifique, on peut estimer assez précisément la catégorie de dépense du département en question. Cela peut être particulièrement utile en support pour des prédictions de dépenses, par département, dans les années à venir en se basant sur les tendances démographiques départementales.

Afin d'améliorer le modèle plusieurs pistes pourraient être explorées :

- Changer les valeurs des catégories de dépenses en s'appuyant sur les quantiles pour plus de précision et afin d'obtenir une meilleure distribution des départements entre les différentes catégories avec un impact direct sur le modèle
- Inclure les données des ALD pour améliorer le modèle. L'algorithme n'a pas sélectionné ces données car jugées non significatives mais en changeant les hyperparamètres du modèle il est probablement possible de rendre les données ALD utiles et d'améliorer l'explicabilité du modèle
- Ajouter des nouveaux jeux de données, en s'intéressant à d'autres aspects tels que les données de soins issues des médicaments et des hôpitaux, des données sur le niveau de pauvreté, niveau d'industrialisation et de PIB pour chaque département et des données relatives aux déserts médicaux étant donné que la demande en soins est probablement directement liée à l'accès aux soins.
- Analyser les résultats de la question précédente, en se focalisant sur les Bouches-du-Rhône et le Nord mais en ce référant au données démographiques et aux tranches d'âge d'intérêt

3 Annexe

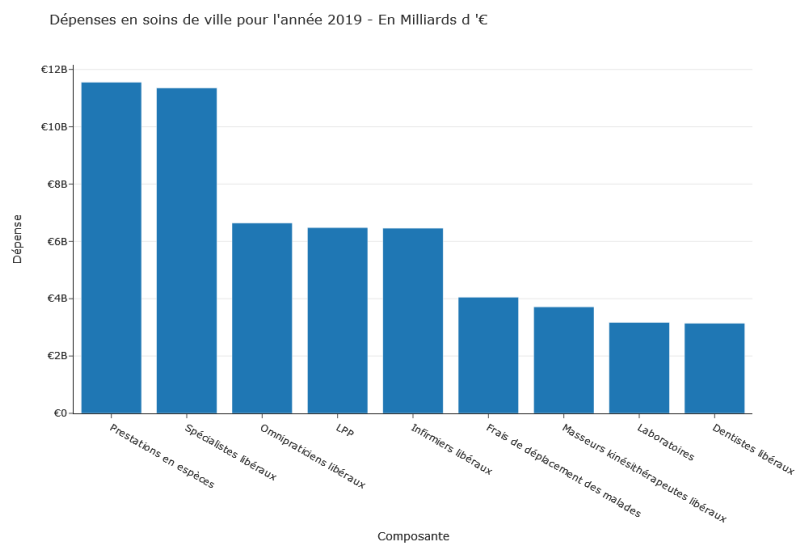


Figure 4: Total des dépenses en soins de ville en 2019 par poste

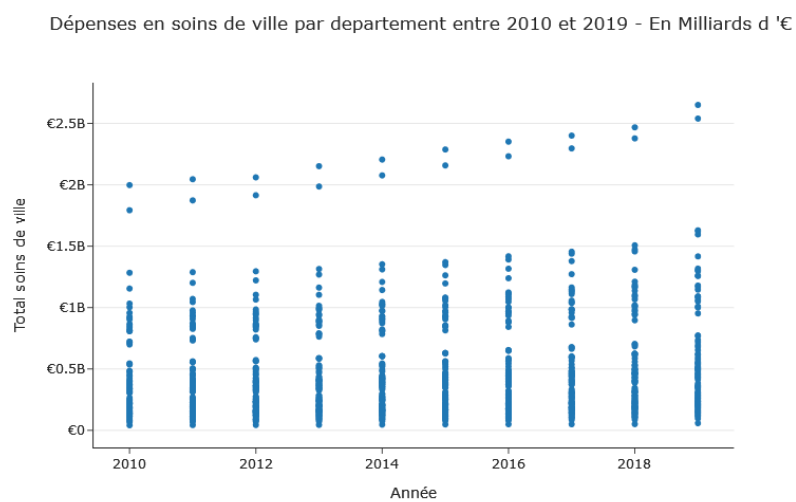


Figure 5: Hétérogénéité du total des dépenses de soins de ville

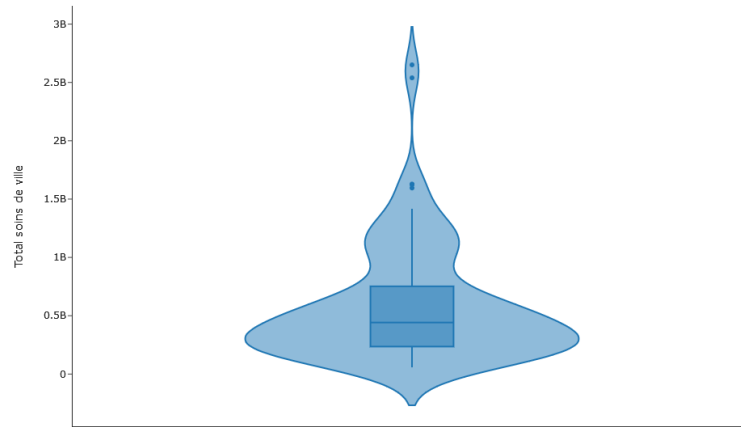


Figure 6: Distribution du total des dépenses de soins de ville

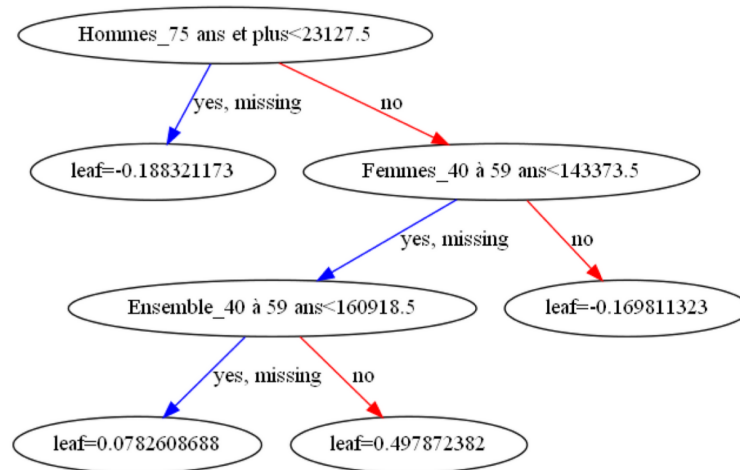


Figure 7: Arbre décisionnel et ses critères de choix