

Analyse

Moustapha Ndiaye

2025-05-31

1. Introduction Ce rapport présente une analyse descriptive et exploratoire des données de santé issues d'un échantillon d'individus. L'objectif est de mieux comprendre les déterminants des problèmes de santé, le recours aux soins, la satisfaction à l'égard des services médicaux, ainsi que les liens potentiels entre ces éléments et les caractéristiques socio-économiques des personnes enquêtées.

L'approche adoptée combine statistiques descriptives, visualisations, analyse en composantes principales (ACP) et corrélations, afin de fournir une lecture approfondie de ces interactions complexes.

2. Données et préparation Les données utilisées ont fait l'objet d'un nettoyage préalable. Les valeurs manquantes ont été imputées :

Par la médiane pour les variables numériques.

Par le mode pour les variables catégorielles.

Par ailleurs, les variables jugées essentielles à l'analyse ont été renommées pour une meilleure lisibilité, ce qui facilite leur traitement et leur interprétation ultérieure.

3. Description des problèmes de santé L'analyse montre que 23,25% des individus déclarent avoir eu un problème de santé, contre 76,75% qui n'en ont pas eu. Ce chiffre révèle une minorité significative touchée par des problèmes médicaux dans la population étudiée.

Cela permet d'identifier un groupe-cible (les malades) sur lequel les analyses suivantes seront davantage centrées, notamment en ce qui concerne les types de problèmes rencontrés, les comportements de recours aux soins et la satisfaction vis-à-vis des services.

4. Types de problèmes chez les malades Parmi les personnes ayant déclaré un problème de santé :

Le type de problème le plus fréquent est le type 1, avec 4 769 cas.

Suivent les types 14 et 11, en fréquence décroissante.

Cette répartition montre qu'une petite poignée de catégories de problèmes concentrent une large proportion des cas. Cela pourrait suggérer des pathologies ou troubles plus fréquents dans le contexte de l'étude, et potentiellement ciblables en termes de prévention ou d'amélioration des services.

5. Recours aux soins Chez les individus malades :

Environ 60% ont déclaré avoir consulté un professionnel de santé.

Tandis que 40% n'ont pas consulté malgré leur problème de santé.

Ce constat soulève une problématique importante de non-recours aux soins. Il peut être le reflet de freins financiers, culturels, géographiques ou administratifs. Ces aspects seront approfondis plus loin via l'analyse des freins et des corrélations socio-économiques.

6. Types de structures consultées Parmi les personnes ayant eu recours aux soins :

La structure la plus utilisée est de type 4, représentant 57,8% des cas.

Suivent les types 2 (12,36%) et 3 (11,08%).

Cela montre une forte préférence (ou accessibilité accrue) pour un type spécifique d'établissement, probablement des structures publiques ou communautaires. Les autres types de structures restent marginalement fréquentés, ce qui peut soulever des questions sur leur accessibilité, coût ou réputation.

7. Satisfaction vis-à-vis des soins Les résultats sont plutôt positifs :

Une majorité écrasante (81,43%) déclare un niveau moyen de satisfaction (niveau 2).

Une minorité (7,63%) est insatisfaite (niveau 1).

Même si le niveau « moyen » n'est pas optimal, cette distribution indique que les soins reçus sont globalement jugés corrects. Il reste toutefois à comprendre si ce niveau de satisfaction est lié à des attentes faibles, à des limitations dans l'offre, ou bien à de réelles limitations qualitatives.

8. Visualisations Quatre visualisations permettent d'appuyer les constats précédents :

Histogramme des problèmes de santé : confirme visuellement que la majorité de la population n'a pas rencontré de problème de santé.

Histogramme du recours aux soins : montre le taux de recours parmi les malades, confirmant le taux de 60%.

Répartition par type de structure consultée : met en évidence la prédominance de la structure de type 4.

Satisfaction des soins : illustre la prédominance de la satisfaction moyenne.

Ces graphiques permettent une lecture plus intuitive des grandes tendances observées dans les données.

9. Analyse en Composantes Principales (ACP) L'analyse en composantes principales des variables de santé révèle deux dimensions principales. La première (20,8 %) oppose un ensemble de variables liées au recours et à la satisfaction des soins (a_consulte, satisfaction_soins, type_probleme) à la variable probleme_sante, suggérant une relation inverse entre état de santé déclaré et recours aux soins. La deuxième dimension (16,1 %) distingue les variables liées à la structure et au type d'établissement de soins (type_structure, s03q08). Ces résultats indiquent que les variables de santé se structurent autour de deux axes majeurs : l'accès aux soins et le type de structure fréquentée, fournissant ainsi des pistes d'analyse sur les déterminants de la santé dans la population étudiée.

10. Corrélations entre santé et statut socio-économique

11. Les axes : Colonnes : variables socio-économiques, identifiées par des noms de type s11q01, s04q43, etc.

Lignes : variables de santé, comme a_consulte, type_probleme, satisfaction_soins, etc.

2. Les couleurs : Représentent le niveau de corrélation entre chaque paire de variables.

Interprétation indicative :

Bleu foncé : corrélation négative forte.

Bleu clair à vert : corrélation faible ou nulle.

Jaune à beige : corrélation positive modérée à forte.

3. Dendrogrammes (arbres) : En haut (colonnes) et à gauche (lignes), ils montrent le clustering hiérarchique :

Les variables regroupées ensemble sont celles qui ont des profils de corrélations similaires.

Interprétations possibles Relations santé / socio-économie : Par exemple, `probleme_sante` est fortement négativement corrélée avec `s11q32` (en bleu foncé) → Peut-être que les ménages concernés par cette modalité socioéconomique (ex. : faible niveau d'éducation ou revenu) ont plus de problèmes de santé.

À l'inverse, `satisfaction_soins` semble positivement corrélée avec certaines variables comme `s04q57` (beige/jaune) → Lien possible entre satisfaction et accès/qualité des soins lié à un meilleur statut socio-économique.

Groupes de variables : `a_consulte`, `type_probleme`, `s03q04`, etc., forment un groupe (dendrogramme serré), ce qui signifie qu'elles réagissent de manière similaire aux variations des variables socio-économiques.

Du côté socio-économique, certaines variables comme `s11q01`, `s11q04`, `s04q57` sont proches → possiblement des variables d'un même thème (ex. : revenus, éducation, logement...).

Conclusion Le graphique en heatmap met en évidence les corrélations entre les variables de santé et les variables socio-économiques. On observe que certaines variables de santé, comme `probleme_sante` ou `satisfaction_soins`, sont particulièrement influencées par des facteurs socio-économiques précis. Le regroupement hiérarchique suggère des sous-groupes cohérents, tant au niveau des indicateurs de santé qu'au niveau des caractéristiques socio-économiques, ce qui pourrait servir de base à une typologie des profils de population selon leur accès et recours aux soins.