RAPPORT:

ÉTUDE DES FACTEURS MULTIDIMENSIONNELS AFFECTANT L'ÉTAT DE SANTÉ DES INDIVIDUS



Étude réalisée par :

VOEDZO Essivi

AIT CHARKI Aymane

ZOUAGHI Ahmed



Table des matières

<u>I.</u>	<u>INTRODUCTION</u>	3
<u>II.</u>	REVUE DE LITTERATURE SCIENTIFIQUE	<u>4</u>
III.	ÉTUDE PREALABLE DES DONNEES	6
1.	CHOIX DES VARIABLES POUR NOTRE ETUDE	6
2.	VALEURS ABERRANTES ET VALEURS MANQUANTES	6
<i>3</i> .	REGROUPEMENT ET ELIMINATION DE MODALITES ATYPIQUES	6
I.	TRAITEMENT DES MODALITES ATYPIQUES CONCERNANT LA VARIABLE DEPENDANTE	6
II.	TRAITEMENT DES MODALITES ATYPIQUES CONCERNANT LA VARIABLE EXPLICATIVE	6
<u>IV.</u>	STATISTIQUES DESCRIPTIVES	<u>7</u>
ı.	ANALYSE DES TENDANCES ET DISTRIBUTIONS DES DONNEES	7
II.	ANALYSE DESCRIPTIVE ET EXPLORATION DES RELATIONS ENTRE LES VARIABLES EXPLICATIVES	8
<u>V.</u>	MODELISATION	10
1.	CHOIX DES MODELES	10
2.	INTERPRETATION	11
<i>3</i> .	COMPARAISON DES MODELES	11
<u>VI.</u>	CONCLUSION	<u> 13</u>
VII	!. ANNEXES	14



I. Introduction

L'état de santé des individus constitue un indicateur fondamental du bien-être global d'une société. Il reflète non seulement les conditions de vie, mais aussi l'efficacité des politiques publiques et la capacité des populations à accéder aux soins de santé. Les déterminants de l'état de santé sont complexes et multidimensionnels, englobant des facteurs socio-économiques, environnementaux, comportementaux et démographiques.

Comprendre ces déterminants est crucial pour identifier les inégalités en matière de santé et orienter les interventions visant à améliorer les conditions de vie des populations.

Dans ce contexte, l'analyse économétrique des variables qualitatives offre un cadre méthodologique robuste pour explorer et interpréter les relations entre un état de santé déclaré et ses facteurs explicatifs. En particulier, l'utilisation de modèles probabilistes tels que les modèles logit et probit permet d'évaluer les effets des différents déterminants et de quantifier leur influence sur l'état de santé des individus.

L'état de santé est influencé par une multitude de facteurs interdépendants. Cependant, l'identification précise de ces déterminants reste une tâche complexe, compte tenu des disparités socio-économiques, des différences comportementales, des caractéristiques environnementales et des variations démographiques. Cette étude cherche donc à répondre à la question suivante :

Quels sont les déterminants socio-économiques, comportementaux, environnementaux et démographiques qui influencent l'état de santé des individus ?

Pour répondre à cette problématique, ce projet repose sur une analyse économétrique des données. Les objectifs principaux sont :

- Identifier les facteurs ayant une influence significative sur l'état de santé des individus.
- Explorer les relations entre une variable dépendante catégorielle (état de santé) et des variables explicatives regroupées en quatre dimensions : démographique, socio-économique, comportementale et environnementale.
- Tester les hypothèses prédéfinies afin de confirmer ou réfuter les liens supposés entre les différentes variables.



II. Revue de littérature scientifique

L'état de santé des individus est influencé par un ensemble complexe de facteurs qui vont au-delà des soins médicaux et des traitements spécifiques. Ces facteurs, appelés déterminants de la santé, englobent des dimensions socio-économiques, comportementales, environnementales et démographiques. Les politiques publiques, les choix individuels et les conditions socio-économiques jouent un rôle majeur dans l'évolution de la santé des populations. L'objectif de cette revue est d'examiner les principaux déterminants qui influencent la santé des individus et d'identifier les mécanismes par lesquels ces facteurs interagissent.

Les facteurs socio-économiques représentent l'un des déterminants les plus influents de la santé avec 45% en France et 19,1%. Ils incluent des variables telles que le revenu, l'éducation, l'emploi et les conditions de logement. Il est bien établi que les individus ayant un revenu élevé, un niveau d'éducation supérieur, un emploi stable et un logement adéquat bénéficient généralement d'une meilleure santé. En revanche, les populations défavorisées sur le plan socio-économique sont souvent exposées à des risques accrus de maladies, à une espérance de vie réduite et à une qualité de vie moins favorable.

Les comportements individuels jouent un rôle crucial dans l'état de santé des individus. Ces comportements incluent l'alimentation, l'activité physique, la consommation de substances (comme le tabac et l'alcool), ainsi que d'autres choix de mode de vie. En particulier, une alimentation déséquilibrée, un manque d'exercice physique, le tabagisme et l'abus d'alcool sont associés à un risque accru de maladies cardiovasculaires, de diabète, de cancers et de maladies respiratoires.

Les comportements de santé sont influencés non seulement par des choix personnels, mais aussi par des facteurs externes tels que les conditions socio-économiques et l'environnement. Par exemple, des individus issus de milieux moins favorisés peuvent avoir moins de moyens pour adopter des comportements de santé positifs, comme l'accès à des aliments sains ou à des infrastructures permettant une activité physique régulière. Des études montrent également que des habitudes de vie comme l'exercice physique régulier et une alimentation saine peuvent réduire significativement le risque de mortalité prématurée et améliorer la qualité de vie.

Les facteurs environnementaux comprennent la qualité de l'air, du logement, l'exposition aux polluants (comme les particules fines et le bruit), ainsi que l'accès à des espaces verts et à des services publics. Ces éléments influencent directement la santé en agissant sur des aspects comme la respiration, les maladies cardiaques, les troubles mentaux et les maladies chroniques. L'exposition à la pollution de l'air, par exemple, a été fortement associée à des risques accrus de maladies respiratoires et cardiovasculaires, tout comme l'exposition au bruit, qui peut perturber le sommeil et affecter la santé mentale.

En outre, les conditions d'habitat et les infrastructures jouent également un rôle essentiel. Des logements mal entretenus ou insalubres, une mauvaise gestion des déchets peuvent nuire à la santé.



De même, l'accès limité à des espaces publics de qualité ou à des moyens de transport appropriés peut également limiter les options de style de vie sain pour certaines populations.

Les facteurs démographiques, tels que l'âge, le sexe et la génétique, contribuent également de manière significative à l'état de santé des individus. Par exemple, les personnes âgées sont généralement plus vulnérables aux maladies chroniques et aux troubles cognitifs, tandis que les femmes et les hommes peuvent présenter des prédispositions différentes en fonction des maladies en raison de différences biologiques.

Les prédispositions génétiques jouent un rôle important dans la susceptibilité à diverses maladies, telles que les cancers héréditaires et les troubles cardiovasculaires. Bien que ces facteurs soient en grande partie hors du contrôle des individus, la recherche médicale permet désormais d'identifier des stratégies de prévention adaptées en fonction des antécédents familiaux et des risques génétiques.

En conclusion, l'état de santé des individus est le résultat d'une interaction complexe entre facteurs socio-économiques, comportementaux, environnementaux et démographiques. Les politiques publiques jouent un rôle essentiel en modifiant les conditions sociales et environnementales afin de promouvoir un environnement favorable à la santé. Des interventions ciblées dans des domaines comme l'éducation, l'accès à des services de santé de qualité, l'aménagement du territoire et la lutte contre la pollution peuvent significativement améliorer la santé des populations. Par ailleurs, les comportements individuels restent un facteur clé, et des actions de prévention et de promotion de la santé sont nécessaires pour encourager des habitudes de vie saines, en particulier dans les groupes les plus vulnérables.

<u>Les hypothèses :</u>

Hypothèse 1 : Les individus avec des revenus élevés, un emploi stable et un haut niveau d'éducation ont un meilleur état de santé grâce à un accès facilité aux soins et des comportements plus sains.

Hypothèse 2 : La pratique régulière d'une activité physique réduit les risques de maladies chroniques et améliore la santé générale.

Hypothèse 3 : Des consultations médicales fréquentes permettent de prévenir et traiter efficacement les problèmes de santé, réduisant ainsi les complications.

Hypothèse 4 : Des logements inadéquats (mal chauffés, surpeuplés, ou pollués) augmentent les risques de maladies et de stress.

Hypothèse 5 : L'âge avancé est lié à une dégradation de l'état de santé en raison de l'augmentation des maladies chroniques et des limitations physiques.



III. Étude préalable des données

La base de données utilisée provient du Panel Suisse des ménages de 2014, portant sur les adultes de 18 à 60 ans. Ce panel annuel vise à observer les changements sociaux et les dynamiques des conditions de vie en Suisse. La base étudiée comprend 10 129 observations et 117 variables couvrant diverses thématiques, telles que la démographie, l'économie, l'éducation, le logement et la santé.

1. Choix des variables pour notre étude

Variable dépendante : état de santé (P14C01)

Notre étude s'appuie sur des variables regroupées en quatre grandes catégories :

- Démographiques (AGE14, SEX14, REGION14).
- Environnementales (H14H14, H14H29, H14H27).
- Comportementales (P14A01, P14A04, P14C12, P14C15).
- Socio-économiques (I14PTOTN, EDYEAR14, WSTAT14, CSPMAJ14).

2. Valeurs aberrantes et valeurs manquantes

Nos données ne contiennent pas de valeurs aberrantes ni de valeurs manquantes.

3. Regroupement et élimination de modalités atypiques

Lors de l'analyse, nous avons identifié des modalités atypiques susceptibles de biaiser nos résultats, telles que "Ne sait pas", "Pas de réponse" ou encore "Inapplicable", qui reviennent fréquemment et pourraient affecter la précision de nos modèles.

i. Traitement des modalités atypiques concernant la variable dépendante

Nous avons observé que la modalité « Inapplicable », lorsqu'elle apparaît pour notre variable dépendante, est systématiquement présente dans toutes les autres variables. Par conséquent, nous avons décidé de supprimer toutes les lignes où elle figure (voir annexe 1).

De plus, nous avons remarqué que notre variable dépendante comporte trois observations atypiques : deux avec la modalité "Ne sait pas" et une avec la modalité "Pas de réponse". Étant donné que leur impact est négligeable, nous avons également décidé de les supprimer.

ii. Traitement des modalités atypiques concernant la variable explicative

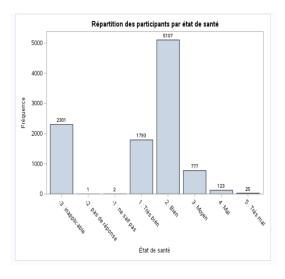
Pour nos variables explicatives, nous avons décidé de regrouper toutes les modalités atypiques en une seule modalité, nommée "Autres" et codée par la valeur (0). Cette approche a été adoptée afin de maximiser la fiabilité de nos modèles lors de la phase de modélisation, en réduisant l'impact des valeurs peu représentatives.



IV. Statistiques descriptives

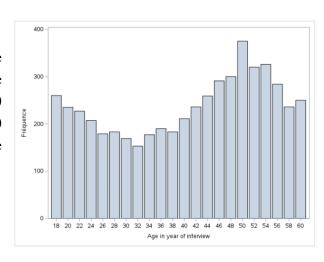
Analyse des tendances et distributions des données

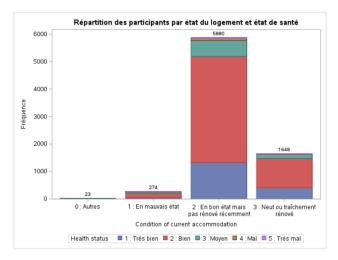
Dans un premier temps, nous réaliserons des histogrammes pour mettre en évidence les principales tendances des données. Ces visualisations nous aideront à déterminer quelle modalité est la plus fréquente et à identifier le nombre de modalités pour chaque variable.



La variable dépendante de notre analyse reflète l'état de santé des participants, réparti en cinq catégories : « Très bien » (1), « Bien » (2), « Moyen » (3), « Mal » (4) et « Très mal » (5). La majorité des participants se déclarent en « Très bien » (1 793) ou « Bien » (5 107), représentant une proportion dominante. À l'inverse, les catégories « Moyen » (777), « Mal » (123) et « Très mal » (25) sont marginales. Pour garantir la fiabilité des modélisations, les modalités « Inapplicable » (-3), « Pas de réponse » (-2) et « Ne sait pas » (-1) ont été exclues (voir section Étude Préalable). Cette répartition confirme que la population étudiée est majoritairement en bonne santé.

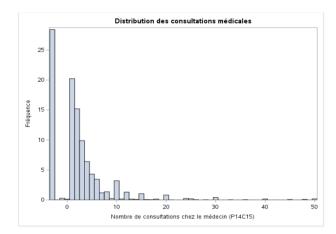
La répartition des participants (18-60 ans) montre une baisse des fréquences entre 30 et 36 ans, suivie d'une augmentation jusqu'à un pic à 50 ans (~380 participants). Une légère baisse se stabilise jusqu'à 60 ans, indiquant une forte présence des individus d'âge moyen, en particulier autour de 50 ans.





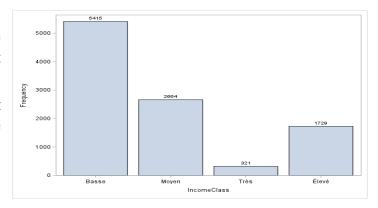
État des logements: La majorité (7800 individus) habite des logements "en bon état mais récemment rénovés". Les logements "neufs" comptent environ 2100 personnes, tandis que les catégories "mauvais état" ou "inapplicable" sont marginales.





Fréquences des consultations médicales : La plupart des individus ont eu peu de consultations médicales au cours des 12 derniers mois. Un petit groupe, avec un nombre élevé de consultations, reflète des besoins spécifiques liés à la santé ou aux conditions socio-économiques.

Revenu : Plus de la moitié (53.46 %) appartient à la classe de revenu 'Basse' (< 40,000 CHF). En revanche, seulement 3.17 % se trouvent dans la classe 'Très élevée' (≥ 150,000 CHF), confirmant une forte asymétrie dans la répartition des revenus.

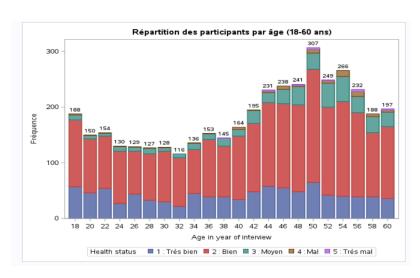


ii. Analyse descriptive et exploration des relations entre les variables explicatives

Nous allons maintenant examiner dans cette section les statistiques descriptives des variables explicatives, tout en visualisant les relations qui les lient.

1. Corrélation entre Âge et Santé : Une Approche Statistique

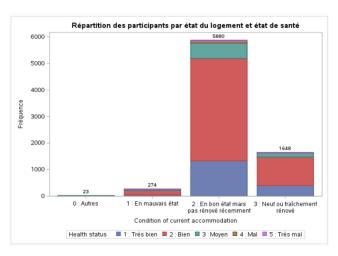
Le graphique met en évidence une relation entre l'âge et la santé : bien que les jeunes (18-30 ans) et les aînés (50-60 ans) présentent une proportion similaire en très bonne santé (« Très bien »), la santé se dégrade avec l'âge, les catégories « Moyen », « Mal » et mal Très **>>** devenant fréquentes. La majorité participants ayant plus de 40 ans, cette tendance est renforcée. Les tests statistiques (khi-deux = 218,8670, p =



0,0001 ; rapport de vraisemblance = 225,0637, p = 0,0001) confirment un lien significatif, positionnant l'âge comme un facteur explicatif clé.

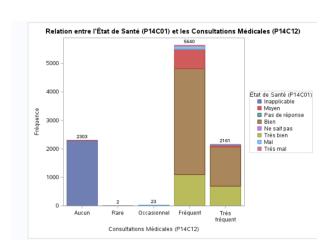


2. Relation entre l'état de logement et la santé



Le graphique montre une relation significative entre l'état du logement et la santé. La majorité des participants vivent dans des logements en bon état mais pas récemment rénovés, avec une bonne santé prédominante. Les logements neufs ou rénovés sont associés à une meilleure santé globale, tandis que les logements en mauvais état (catégorie « 1 ») montrent une proportion accrue de santé moyenne ou mauvaise. Les tests statistiques confirment cette relation : khi-deux (62,6846, p-value = 0,0001) et test de vraisemblance (50,7186, p-value = 0,0001).

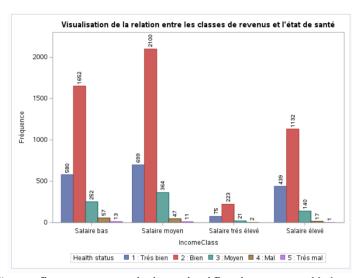
3. Relation entre la fréquence de consultation médicales et l'état de santé



Le graphique révèle une relation entre l'état de santé (P14C01) et les consultations médicales (P14C12). Les individus sans consultation déclarent majoritairement une très bonne santé (« Très bien »), tandis que ceux ayant des consultations fréquentes présentent des états de santé plus variés, incluant des perceptions « Moyen » ou « Mal ». Ces résultats mettent en évidence le lien entre un besoin accru de soins médicaux et une perception plus dégradée de la santé.

4. Relation entre le revenu et l'état de santé

Le graphique illustre une relation entre le revenu et l'état de santé : les classes "Élevée" et "Très élevée" affichent des perceptions majoritairement positives de la santé (« Bonne » ou « Très bonne »), tandis que la classe "Basse" se distingue par une proportion élevée de réponses "Inapplicable" (-3, 42.49 %). Cette tendance reflète une amélioration des perceptions positives avec l'augmentation du revenu, mettant en lumière l'influence des barrières socio-économiques dans la classe "Basse". Les tests statistiques (Chi-Square =



2643.97, p < 0.0001; Cramer's V = 0.295) confirment une relation significative et modérée, positionnant le revenu comme un facteur explicatif clé de l'état de santé.



V. Modélisation

1. Choix des modèles

Dans cette analyse, nous avons choisi d'utiliser les modèles MPL ainsi que les modèles Logit et Probit multinomiaux ordonnés pour tester nos hypothèses. La variable dépendante P14C01, mesurée sur une échelle ordonnée de 1 à 5, est particulièrement adaptée à ces modèles. Le modèle MPL permet d'analyser simultanément plusieurs variables dépendantes corrélées, tandis que les modèles logit et probit multinomiaux ordonnés estiment les probabilités cumulées des différentes catégories de la variable dépendante, tout en respectant l'ordre des catégories. Ces approches sont idéales pour gérer les variables ordinales et prendre en compte les relations entre les catégories de la variable à expliquer.

i. Modèle MPL

D'après <u>l'annexe 2</u>, on constate que l'écart entre les valeurs observées et prédites est relativement faible, avec une valeur de 2736.6308, ce qui indique que le modèle ajuste bien les données. De plus, la déviance normalisée est de 7825.0000, et son ratio proche de 1 suggère également un bon ajustement. En ce qui concerne les tests de Khi-2 de Pearson et Pearson normalisé, leurs valeurs respectives de 2736.6308 et 7825.0000, avec des ratios proches de 0.35, montrent que le modèle ne présente pas de grandes erreurs dans l'ajustement des données. Par conséquent, ces résultats indiquent que le modèle semble être bon, avec un ajustement satisfaisant aux données observées.

ii. Modèle Logit multinomial ordonné

D'après <u>l'annexe 3</u>, le modèle logit multinomial est significatif (p < 0.0001) et présente un ajustement global satisfaisant. L'indicateur R² max-rescaled de 15.80 % montre que le modèle explique une proportion modérée de la variabilité des données. Les critères d'ajustement, comme l'AIC (13,491.921) et le -2 Log L (13,419.921), indiquent une nette amélioration par rapport au modèle avec uniquement l'intercept (AIC = 14,548.266). Le test de Khi-2 global (Likelihood Ratio Chi-Square = 1,120.3447, p < 0.0001) confirme la contribution significative des variables explicatives.

La concordance globale de 68.2 % et un c-statistic de 0.688 témoignent d'une capacité prédictive modérée mais satisfaisante. Ces résultats placent le modèle logit comme un choix robuste pour modéliser les déterminants de l'état de santé, tout en offrant une interprétation claire des effets des covariables.

iii. Modèle probit multinomial ordonné

D'après <u>l'annexe 4</u>, le modèle probit cumulatif est significatif (p < 0.0001) et présente un ajustement global satisfaisant. L'indicateur R² max-rescaled de 15.41 % montre que le modèle explique une proportion modérée de la variabilité des données. Les statistiques d'ajustement, comme l'AIC (13527.928) et le -2 Log L (13449.928), indiquent une amélioration notable par rapport au modèle avec uniquement l'intercept. Le test de Khi-2 global (Likelihood Ratio Chi-Square = 1090.3374, p < 0.0001) confirme que les variables explicatives ont un effet significatif. La concordance de 67.8 % et un c-statistic de 0.683 indiquent une performance modérée pour la prédiction des probabilités. Ces résultats suggèrent que le modèle probit offre un bon ajustement global, bien qu'il reste limité dans sa capacité prédictive.



2. Interprétation

i. Modèle MPL

D'après <u>l'annexe 5</u>, le modèle présente plusieurs variables significatives, telles que AGE14, qui a un effet positif significatif sur la variable dépendante, augmentant celle-ci de 0.0088 par unité d'augmentation de l'âge, et certaines catégories de REGION14, comme la région 1 et la région 7, qui influencent respectivement la variable dépendante de -0.0732 et 0.0592. De même, des variables comme H14H14, P14A04, P14C12, et P14C15 montrent des effets significatifs avec des p-values très faibles. En revanche, des variables comme SEX14, certains niveaux de REGION14, ainsi que plusieurs niveaux de H14H14, P14A01, WSTAT14, et CSPMAJ14 n'ont pas d'effet statistiquement significatif, indiquant qu'elles n'apportent pas de contribution importante au modèle.

ii. Modèle Logit multinomial ordonné

D'après <u>l'annexe 6</u>, le modèle logit multinomial met en évidence plusieurs variables significatives influençant l'état de santé. AGE14 a un effet positif significatif, augmentant la variable dépendante de 0.0295 par année (p < 0.0001). Certaines régions, comme la région 1 (-0.2889, p < 0.0001) et la région 7 (0.2728, p = 0.0384), montrent des impacts opposés. L'activité physique hebdomadaire (P14A04) a un effet négatif (-0.0651, p < 0.0001), tandis que le nombre de consultations médicales (P14C15) a un effet positif (0.0647, p < 0.0001). En revanche, des variables comme SEX14 et certains niveaux de CSPMAJ14 ne sont pas significatives. Ce modèle permet ainsi d'identifier les principaux déterminants de la santé, tout en écartant ceux ayant un faible impact.

iii. Modèle probit multinomial ordonné

D'après <u>l'annexe 7</u>, le modèle probit cumulatif identifie l'âge (coef = 0.0169) et l'état du logement (coef = 0.2693) comme des facteurs clés avec des effets significatifs (p < 0.0001). Certaines régions, comme la région 7 (coef = 0.1450, p = 0.0454), montrent un effet positif, tandis que d'autres, comme la région 6, ont un effet négatif (coef = -0.1051, p = 0.0305). Les variables comportementales et professionnelles ont également un impact important, mais le sexe et les années d'éducation n'ont pas d'effet significatif.

3. Comparaison des modèles

Pour comparer les modèles probit et logit, nous nous appuierons sur le pourcentage de concordance ainsi que sur les graphiques des effets estimés générés avec la commande "effectplot fit" sous SAS, permettant d'évaluer visuellement les différences dans les prédictions.



i. Graphiques des effets estimés

Les courbes des modèles Probit et Logit (voir <u>annexe 8</u>) sont très similaires car les deux utilisent des fonctions sigmoïdes pour modéliser des réponses discrètes. Bien que les modèles reposent sur des distributions d'erreurs différentes (normale pour le probit, logistique pour le logit), cette différence est peu marquée ici, ce qui rend les courbes visuellement proches. Les pourcentages de concordance similaires (67.8% pour le probit et 68.2% pour le logit) montrent que les deux modèles fournissent des prédictions presque équivalentes. Ce phénomène est également observé pour toutes les variables analysées.

Le graphique du modèle MPL (voir <u>annexe 9</u>) présente des courbes linéaires uniformes, indiquant que l'âge (AGE14) influence directement l'état de santé (P14C01) de manière constante pour chaque catégorie de santé. Cette relation linéaire suggère que l'effet de l'âge est uniforme à travers les différentes catégories de l'état de santé, ce qui est typique pour des modèles simples.

Avec notre variable dépendante discrète, comme l'état de santé, les modèles Probit et Logit multinomial ordonné sont plus adaptés que le modèle MPL. En effet, les modèles Probit et Logit permettent de capturer des relations non linéaires entre les variables explicatives et l'état de santé, et sont mieux adaptés pour gérer des données discrètes, contrairement au modèle MPL qui suppose une relation linéaire simple. Les courbes des modèles Probit et Logit reflètent cette capacité à modéliser des variations plus complexes et non linéaires, offrant ainsi une meilleure flexibilité pour des variables discrètes.

ii. Autres indicateurs de performance des modèles

Ce tableau comparatif (voir <u>annexe 10</u>) met en évidence que le modèle Logit offre les meilleures performances globales parmi les trois modèles testés. Avec un AIC de 13,497.568, le Logit présente l'ajustement le plus précis, légèrement devant le modèle Probit (AIC = 13,527.928) et bien devant le MPL (AIC = 14,059.4726). Le Logit se distingue également par un R² de 0.1334 et une concordance de 68.2 %, indiquant une meilleure capacité à expliquer et prédire la variable dépendante.

Le Probit, bien que proche, a une concordance légèrement inférieure (67.8 %) et un AUC de 0.683 contre 0.688 pour le Logit. En revanche, le modèle MPL, malgré un ratio de déviance proche de 1, affiche des critères d'ajustement moins compétitifs. En conclusion, le modèle Logit est le plus adapté pour cette analyse, suivi du Probit, tandis que le MPL reste moins performant dans ce contexte.



VI. Conclusion

Cette étude a confirmé plusieurs hypothèses initiales sur les déterminants de l'état de santé des individus à travers l'analyse économétrique réalisée avec les modèles MPL, Logit multinomial ordonné et Probit multinomial ordonné. Les résultats ont validé les hypothèses selon lesquelles l'âge (H5), l'état du logement (H4), l'activité physique (H2), et les consultations médicales fréquentes (H3) jouent un rôle significatif dans la détermination de l'état de santé. En revanche, certaines hypothèses comme l'impact du sexe (H6) ou des catégories socio-professionnelles (H1) ont été partiellement ou faiblement confirmées.

Les modèles Logit et Probit ont démontré des performances similaires et supérieures pour modéliser les relations entre variables ordinales, avec des scores de concordance de 68.2 % et 67.8 % respectivement. Ces résultats indiquent que ces modèles sont particulièrement adaptés pour des données discrètes. Le modèle MPL, bien qu'il ait offert une perspective linéaire, a montré un ajustement global moins compétitif, confirmant sa pertinence moindre dans ce contexte.

Les résultats obtenus soulignent l'importance des déterminants socio-économiques, comportementaux et environnementaux sur la santé, et valident en grande partie les hypothèses initiales. Ils appellent à des politiques ciblées visant à améliorer les conditions de logement, à promouvoir un mode de vie actif, et à faciliter un accès équitable aux soins médicaux. Cette étude démontre également la pertinence d'une approche économétrique multi-modèles pour une compréhension approfondie des facteurs de santé, offrant ainsi une base solide pour des interventions en santé publique.

Cependant, cette analyse présente certaines limites qui méritent d'être soulignées. La population étudiée est majoritairement composée d'individus en bonne santé, avec un âge moyen de 44 ans, ce qui pourrait biaiser les conclusions en sous-estimant l'impact de certains facteurs sur la santé. Pour renforcer la validité et la généralisation des résultats, il serait essentiel d'élargir la base de données en incluant davantage de personnes en mauvaise santé.

Cela permettrait d'identifier plus précisément les facteurs influençant l'état de santé, notamment chez les populations plus vulnérables. Les recherches futures pourraient également explorer des sous-groupes spécifiques en fonction de la gravité des maladies, afin d'enrichir les recommandations en matière de politiques de santé publique.



VII. Annexes

Bibliographie:

 $\frac{https://www.auvergne-rhone-alpes.developpement-durable.gouv.fr/IMG/pdf/02-epures_fiche_determinants_janv23-2023.pdf$

 $\underline{https://www.economiesuisse.ch/fr/articles/sante-quatre-facteurs-qui-pesent-plus-dans-la-balance-que-les-soins-et-les-traitements}$

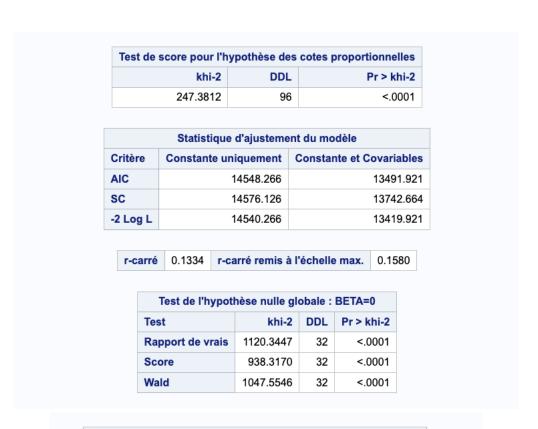
https://www.bfs.admin.ch/bfs/fr/home/statistiques/sante/determinants.html

Annexe 1 : Modalité 'Innaplicable' (-3)

										Rela	ation entre	SATIS_VIE	et les autr	es variat	oles											
Identification number of person		Sex	Age in year of interview	in year of	Number of children	Classification of Education	Years of Education based on ISCED Classification	Actual occupation, from grid	Foreigners residence permit	Since	Birth in Switzerland	Satisfaction with life in general		Illness, accident: reason		Illness, accident (closely related person): yes, no	(of a closely related	of closely related person:	death of closely	close relationship:	Suffering from termination of close relationship		or among	Problems with own children: yes, no	problems with	event T
13101	131	1	42	2	-3	32	12	1	-3	-5	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
74103	741	2	28	1	-3	32	12	4	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
118102	1181	2	40	2	-3	32	12	2	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
118101	1181	1	49	2	-3	52	16	1	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
183102	1831	2	42	2	-3	51	18	7	7	2012	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
251103	2511	- 1	25	1	-3	20	9	4	-3	1989	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
251104	2511	- 1	26	1	-3	10	8	1	-3	-5	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
258102	2581	- 1	53	2	-3	20	9	2	-3	1985	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
337104	3371	2	24	1	-3	20	9	2	- 1	2007	2	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
337102	3371	2	52	2	-3	20	9	7	-3	1984	2	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
343104	3431	2	19	1	-3	10	8	4	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
355102	3552	1	25	1	-3	32	12	1	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3		-3	-3	-3	-3
390102	3901	2	57	2	-3	33	10	7	-3	1957	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
396102	3961	2	36	2	-3	31	13	2	-3	1978	2	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
397102	3971	2	59	2	-3	32	12	7	-3	-5	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
397103	3971	1	28	1	-3	20	9	6	-3	-6	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3			-3	-3	-3	-3
397202	3972	1	41	2	-3	51	18	1	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
436102	4361	2	60	2	-3	32	12	2	-3	-5	1	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3
438102	4381	1	55	2	-3	32	12	1				-3	-3	-3					-3	-3	-3	-3	-3	-3	-3	-3
453102		- 1		2				1						-3		-3			-3	-3	-3		-3	-3	-3	
475102		2	46	2				2			-3			-3		-3			-3	-3			-3	-3	-3	
530102				2				1			-3			-3						-3			-3	-3	-3	
545102		1	43	2				1			1			-3		-3			-3	-3	-3		-3	-3	-3	
549104				2				1						-3		-3			-3	-3	-3		-3	-3	-3	
549302				-2										-3		-3			- 3	-3			-3	-3	-3	
577102			59	5				7			2			-3						-3			-3	-3	-3	
587102				1				4			1			-3		-3			-3	-3			-3	-3	-3	
606101		1		2				1											-3	-3			-3		-3	
		2			-3		9	4	-					-3		-3			-3	-3	-3		-3	-3 -3		
618104		_		2				2						-3		-3				- 3					-3	
618102							21							-3									-3	-3	-3	
645103				1				4						-3		-3			-3	-3	-3		-3	-3	-3	
645105				1				4						-3		-3			-3	-3			-3	-3	-3	
645101				2				7						-3		-3			-3	-3			-3	-3	-3	
650104		1	2.0	1				4			1			-3		-3			-3	-3	-3		-3	-3	-3	
650101		1		2				1						-3		-3			-3	-3	-3		-3	-3	-3	
650103		- 1	26	1				4			1			-3		-3			-3	-3			-3	-3	-3	
689103	6891	- 1	40	2	-3	32	12	1	1	-5	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3	-3



Critères d'évaluation de l'adéquation									
Critère	DDL	Valeur	Valeur/DDL						
Ecart	7792	2736.9080	0.3512						
Déviance normalisée	7792	7825.0000	1.0042						
Khi2 de Pearson	7792	2736.9080	0.3512						
Pearson normalisé X2	7792	7825.0000	1.0042						
Log-vraisemblance		-6993.1326							
Log-vraisemblance complète		-6993.1326							
AIC (préférer les petites valeurs)		14054.2652							
AICC (préférer les petites valeurs)		14054.5707							
BIC (préférer les petites valeurs)		14291.0779							







Association des probabilités prédites et des réponses observées									
Pourcentage concordant	67.8	D de Somers	0.366						
Pourcentage discordant	31.1	Gamma	0.370						
Pourcentage lié	1.1	Tau-a	0.187						
Paires	15657422	С	0.683						

Test de score pou	r l'hypothès	e des pentes égales
khi-2	DDL	Pr > khi-2
269.1500	96	<.0001

	Statistique d'ajusteme	ent du modèle
Critère	Constante uniquement	Constante et Covariables
AIC	14548.266	13522.255
sc	14576.126	13772.998
-2 Log L	14540.266	13450.255

	r-carré	0.1300	r-carré remis à l'échelle max.	0.1541
--	---------	--------	--------------------------------	--------

Test de l'hypot	hèse nulle gl	obale :	BETA=0
Test	khi-2	DDL	Pr > khi-2
Rapport de vrais	1090.0107	32	<.0001
Score	1070.3879	32	<.0001
Wald	926.4802	32	<.0001



Paramètre		DDL	Estimation	Erreur type	Intervalle de confia	ance de Wald à95%	Khi-2 de Wald	Pr > khi-
Intercept		1	1.6130	0.0579	1.4994	1.7265	775.38	<.000
AGE14		1	0.0088	0.0006	0.0076	0.0099	226.50	<.000
SEX14	2	1	-0.0171	0.0149	-0.0463	0.0122	1.31	0.252
REGION14	1	1	-0.0732	0.0211	-0.1145	-0.0319	12.07	0.000
REGION14	3	1	0.0179	0.0223	-0.0259	0.0616	0.64	0.423
REGION14	4	1	0.0359	0.0217	-0.0066	0.0784	2.74	0.097
REGION14	5	1	-0.0001	0.0234	-0.0460	0.0458	0.00	0.996
REGION14	6	1	-0.0516	0.0245	-0.0996	-0.0036	4.44	0.035
REGION14	7	1	0.0592	0.0370	-0.0133	0.1317	2.56	0.109
H14H14	0	1	-0.2477	0.2416	-0.7213	0.2259	1.05	0.305
H14H14	1	1	0.1510	0.0370	0.0785	0.2234	16.69	<.000
H14H14	3	1	-0.0156	0.0166	-0.0482	0.0169	0.89	0.346
H14H27	0	1	-0.0278	0.2276	-0.4738	0.4182	0.01	0.902
H14H27	1	1	0.0661	0.0250	0.0172	0.1151	7.01	0.008
P14A01	-3	1	0.7884	0.5924	-0.3726	1.9494	1.77	0.183
P14A01	-2	1	0.8104	0.5954	-0.3566	1.9774	1.85	0.173
P14A01	-1	1	-0.0716	0.4196	-0.8940	0.7509	0.03	0.864
P14A01	2	1	0.0890	0.0305	0.0292	0.1487	8.52	0.003
P14A04		1	-0.0152	0.0041	-0.0232	-0.0072	13.81	0.000
P14C12	-3	0	0.0000	0.0000	0.0000	0.0000		
P14C12	-2	1	0.1517	0.5927	-1.0100	1.3134	0.07	0.798
P14C12	-1	1	0.3858	0.1268	0.1373	0.6343	9.26	0.002
P14C12	2	1	-0.0922	0.0164	-0.1243	-0.0600	31.61	<.000
P14C15		1	0.0166	0.0008	0.0151	0.0181	475.64	<.000
I14PTOTN		1	-0.0000	0.0000	-0.0000	-0.0000	11.50	0.000
EDYEAR14		1	-0.0045	0.0033	-0.0109	0.0018	1.94	0.163
WSTAT14	-3	1	0.2094	0.4197	-0.6132	1.0319	0.25	0.617
WSTAT14	2	1	0.1957	0.0550	0.0880	0.3035	12.67	0.000
WSTAT14	3	1	0.1330	0.0376	0.0593	0.2066	12.52	0.000
CSPMAJ14	-3	1	-0.0137	0.0332	-0.0788	0.0513	0.17	0.679
CSPMAJ14	1	1	0.0629	0.0730	-0.0802	0.2060	0.74	0.389
CSPMAJ14	2	1	0.0269	0.0526	-0.0762	0.1300	0.26	0.608
CSPMAJ14	3	1	-0.0270	0.0275	-0.0810	0.0269	0.96	0.326
CSPMAJ14	4	1	-0.0022	0.0240	-0.0492	0.0448	0.01	0.927
CSPMAJ14	6	1	0.0274	0.0229	-0.0174	0.0723	1.44	0.230
CSPMAJ14	7	1	-0.0179	0.0330	-0.0827	0.0468	0.29	0.587
CSPMAJ14	8	1	0.0983	0.0336	0.0325	0.1641	8.57	0.003
Echelle	-	1	0.5914	0.0047	0.5822	0.6007	5.51	2,000



Analyse des valeurs estimées du maximum de vraisemblance											
Paramètre		DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2					
Intercept	5	1	-7.3916	0.3081	575.4667	<.0001					
Intercept	4	1	-5.3471	0.2273	553.3513	<.0001					
Intercept	3	1	-3.1249	0.2090	223.5432	<.0001					
Intercept	2	1	0.4704	0.2048	5.2741	0.0216					
AGE14		1	0.0295	0.00210	196.8868	<.0001					
SEX14	2	1	-0.0448	0.0528	0.7192	0.3964					
REGION14	1	1	-0.2889	0.0745	15.0522	0.0001					
REGION14	3	1	0.0941	0.0794	1.4047	0.2359					
REGION14	4	1	0.0969	0.0771	1.5812	0.2086					
REGION14	5	1	-0.0311	0.0828	0.1412	0.7071					
REGION14	6	1	-0.1794	0.0864	4.3077	0.0379					
REGION14	7	1	0.2728	0.1318	4.2868	0.0384					
H14H14	0	1	-0.9902	0.8301	1.4229	0.2329					
H14H14	1	1	0.4461	0.1308	11.6374	0.0006					
H14H14	3	1	-0.0364	0.0588	0.3833	0.5359					
H14H27	0	1	-0.0728	0.7842	0.0086	0.9261					
H14H27	1	1	0.2256	0.0889	6.4418	0.0111					
P14A01	-3	1	2.0984	1.9881	1.1141	0.2912					
P14A01	-2	1	2.4121	1.8962	1.6181	0.2034					
P14A01	-1	1	-0.2079	1.5482	0.0180	0.8932					
P14A01	2	1	0.2113	0.1084	3.7980	0.0513					
P14A04		1	-0.0651	0.0144	20.3825	<.0001					
P14C12	-3	0	0								
P14C12	-2	1	0.6958	2.1893	0.1010	0.7506					
P14C12	-1	1	1.4381	0.4419	10.5901	0.0011					
P14C12	2	1	-0.2718	0.0600	20.5260	<.0001					
P14C15		1	0.0647	0.00351	338.7223	<.0001					
I14PTOTN		1	-1.62E-6	5.208E-7	9.7067	0.0018					
EDYEAR14		1	-0.0211	0.0115	3.3272	0.0681					
WSTAT14	-3	1	0.8475	1.4809	0.3275	0.5671					
WSTAT14	2	1	0.6318	0.1955	10.4424	0.0012					
WSTAT14	3	1	0.3681	0.1333	7.6303	0.0057					
CSPMAJ14	-3	1	-0.0503	0.1176	0.1830	0.6688					
CSPMAJ14	1	1	0.2622	0.2597	1.0197	0.3126					
CSPMAJ14	2	1	0.1016	0.1858	0.2990	0.5845					
CSPMAJ14	3	1	-0.0896	0.0974	0.8451	0.3579					
CSPMAJ14	4	1	0.0132	0.0848	0.0241	0.8766					
CSPMAJ14	6	1	0.1017	0.0812	1.5703	0.2102					
CSPMAJ14	7	1	-0.0608	0.1167	0.2715	0.6023					
CSPMAJ14	8	1	0.3075	0.1193	6.6433	0.0100					

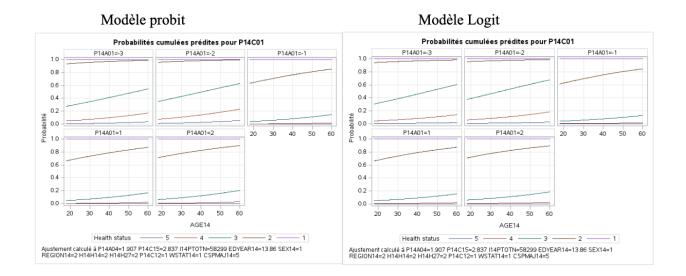


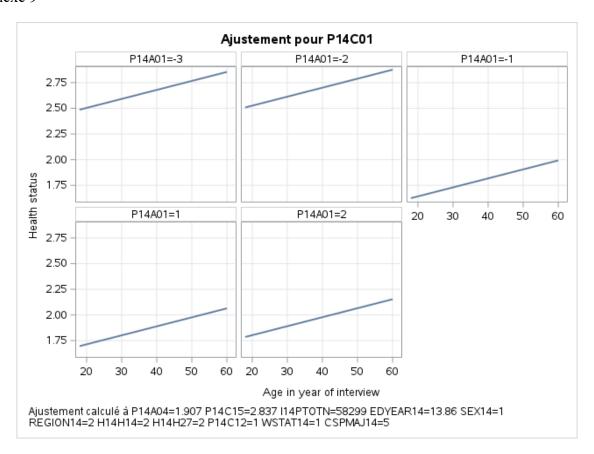
Es	stimation du rapport de	cotes	
Effet	Estimation du point		de confiance ald à95%
AGE14	1.030	1.026	1.034
SEX14 2 vs 1	0.956	0.862	1.060
REGION14 1 vs 2	0.749	0.647	0.867
REGION14 3 vs 2	1.099	0.940	1.284
REGION14 4 vs 2	1.102	0.947	1.281
REGION14 5 vs 2	0.969	0.824	1.140
REGION14 6 vs 2	0.836	0.706	0.990
REGION14 7 vs 2	1.314	1.015	1.701
H14H14 0 vs 2	0.371	0.073	1.890
H14H14 1 vs 2	1.562	1.209	2.018
H14H14 3 vs 2	0.964	0.859	1.082
H14H27 0 vs 2	0.930	0.200	4.324
H14H27 1 vs 2	1.253	1.053	1.492
P14A01 -3 vs 1	8.153	0.166	401.426
P14A01 -2 vs 1	11.158	0.271	458.835
P14A01 -1 vs 1	0.812	0.039	16.887
P14A01 2 vs 1	1.235	0.999	1.528
P14A04	0.937	0.911	0.964
P14C12 -2 vs 1	2.005	0.027	146.469
P14C12 -1 vs 1	4.213	1.772	10.017
P14C12 2 vs 1	0.762	0.677	0.857
P14C15	1.067	1.059	1.074
I14PTOTN	1.000	1.000	1.000
EDYEAR14	0.979	0.957	1.002
WSTAT14 -3 vs 1	2.334	0.128	42.527
WSTAT14 2 vs 1	1.881	1.282	2.760
WSTAT14 3 vs 1	1.445	1.113	1.876
CSPMAJ14 -3 vs 5	0.951	0.755	1.197
CSPMAJ14 1 vs 5	1.300	0.781	2.162
CSPMAJ14 2 vs 5	1.107	0.769	1.593
CSPMAJ14 3 vs 5	0.914	0.755	1.107
CSPMAJ14 4 vs 5	1.013	0.858	1.196
CSPMAJ14 6 vs 5	1.107	0.944	1.298
CSPMAJ14 7 vs 5	0.941	0.749	1.183
CSPMAJ14 8 vs 5	1.360	1.076	1.718



Analyse des valeurs estimées du maximum de vraisemblance										
Paramètre		DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2				
Intercept	5	1	-3.7156	0.1478	631.9242	<.0001				
Intercept	4	1	-2.8605	0.1218	551.8436	<.0001				
Intercept	3	1	-1.8009	0.1162	240.1313	<.0001				
Intercept	2	1	0.2957	0.1146	6.6518	0.0099				
AGE14		1	0.0169	0.00116	209.7724	<.0001				
SEX14	2	1	-0.0208	0.0295	0.4952	0.4816				
REGION14	1	1	-0.1493	0.0417	12.7938	0.0003				
REGION14	3	1	0.0473	0.0441	1.1512	0.2833				
REGION14	4	1	0.0619	0.0428	2.0872	0.1485				
REGION14	5	1	-0.0121	0.0463	0.0684	0.7937				
REGION14	6	1	-0.1051	0.0486	4.6830	0.0305				
REGION14	7	1	0.1450	0.0725	4.0020	0.0454				
H14H14	0	1	-0.5599	0.4813	1.3536	0.2446				
H14H14	1	1	0.2693	0.0722	13.9054	0.0002				
H14H14	3	1	-0.0282	0.0329	0.7329	0.3920				
H14H27	0	1	-0.0420	0.4491	0.0087	0.9255				
H14H27	1	1	0.1351	0.0491	7.5661	0.0059				
P14A01	-3	1	1.0767	1.1183	0.9269	0.3357				
P14A01	-2	1	1.2904	1.0790	1.4303	0.2317				
P14A01	-1	1	-0.0800	0.8261	0.0094	0.9229				
P14A01	2	1	0.1372	0.0603	5.1784	0.0229				
P14A04		1	-0.0324	0.00809	15.9859	<.0001				
P14C12	-3	0	0							
P14C12	-2	1	0.3016	1.1670	0.0668	0.7961				
P14C12	-1	1	0.6413	0.2454	6.8309	0.0090				
P14C12	2	1	-0.2470	0.0336	53.9359	<.0001				
P14C15		1	0.0245	0.00183	180.2581	<.0001				
I14PTOTN		1	-9.59E-7	2.953E-7	10.5547	0.0012				
EDYEAR14		1	-0.00935	0.00645	2.1041	0.1469				
WSTAT14	-3	1	0.4657	0.8219	0.3211	0.5710				
WSTAT14	2	1	0.3841	0.1080	12.6551	0.0004				
WSTAT14	3	1	0.2250	0.0743	9.1595	0.0025				
CSPMAJ14	-3	1	-0.0324	0.0658	0.2426	0.6223				
CSPMAJ14	1	1	0.1215	0.1439	0.7129	0.3985				
CSPMAJ14	2	1	0.0488	0.1043	0.2190	0.6398				
CSPMAJ14	3	1	-0.0562	0.0545	1.0642	0.3023				
CSPMAJ14	4	1	-0.00523	0.0475	0.0121	0.9124				
CSPMAJ14	6	1	0.0551	0.0452	1.4853	0.2229				
CSPMAJ14	7	1	-0.0422	0.0656	0.4136	0.5201				
CSPMAJ14	8	1	0.1813	0.0663	7.4855	0.0062				









Critères	Logit	MPL	Probit
AIC	13,497.568	14,059.4726	13,527.928
-2 Log L	13,419.568	-6,992.7363	13,449.928
R ²	0.1334	-	0.1301
Max-rescaled R ²	0.1581	-	0.1541
Concordance	68.2 %	-	67.8 %
Somers' D	0.376	-	0.366
c-statistic (AUC)	0.688	-	0.683
Likelihood Ratio Chi-Sq	1120.6979	-	1090.3374
Meilleur ajustement global	Oui	Bon	Oui

