



FACTEURS EXPLICATIFS DE LA CONSOMMATION DU TABAC EN FRANCE

Pâquarse Mahouvi & Ronan Gonnet



PROJET ECONOMETRIE
UNIVERSITE DE RENNES 1
Avril 2022

Tables des matières

Tables des matières.....	1
Liste des tableaux	2
Liste des graphiques	2
Introduction, Enoncé du problème	3
I. Présentation des données	3
II. Statistiques descriptives.....	4
1. Analyse univariée	4
1.1 Consommation de tabac.....	4
1.2 Présentation des variables explicatives.....	4
2. Analyse Bivariée	6
2.1. V de cramer entre consommation de tabac et les variables explicatives	6
2.2. Répartition de la consommation de selon quelques variables explicatives	6
2.3. Matrice de corrélation entre les variables explicatives	7
III. Estimation des modèles	7
3.1. Création de l'échantillon	7
3.2. Approche Statistique de sélection.....	7
4. Choix du modèle.....	8
4.1. La qualité d'ajustement	8
4.2. Courbe ROC	8
4.3. Matrice de confusion et erreur de classification	8
4.4. Déviance du modèle	9
4.5. Résumé de la sélection	9
5. Tests d'adéquation du modèle	9
5.1. Test de Hosmer-Lemeshow	9
5.2. Analyse des résidus	10
6. Présentation et interprétation du modèle (Odds ratios)	10
6.1 Le modèle, coefficients et odds ratios	10
6.2. Interprétation du modèle.....	11
7. Scoring.....	12
7.1 Table des scores.....	12
7.2 Seuils des scores	13
Conclusion.....	13
Annexes	14
Webographie.....	15

Liste des tableaux

Tableau 1 Les variables explicatives	6
Tableau 2 : Les modèles retenus par les méthodes pas à pas.	8
Tableau 3 : Pseudo R2	8
Tableau 4 : Erreur de Classification	9
Tableau 5 : Déviance	9
Tableau 6 : Résumé de la sélection	9
Tableau 7 : Le modèle, coefficients, odds ratios	11
Tableau 8 : Table des scores.....	12

Liste des graphiques

Figure 1 Répartition selon la consommation de tabac.....	4
Figure 2 : V de cramer entre la consommation de tabac et les variables explicatives.....	6
Figure 3 : Répartition de la consommation de tabac selon la situation et le logement.....	6
Figure 4 : Matrice de corrélation	7
Figure 5 : Courbe ROC.....	8
Figure 6 : Résidu de Pearson	10
Figure 7 : Seuil de score	13
Figure 8 : Sélection des variables pour le modèle	14
Figure 9 : Variable Importance Plot	14

Introduction, Enoncé du problème

Le tabagisme est l'intoxication aiguë ou chronique de nature physiologique et psychique provoquée par l'abus du tabac. Dans le monde, le tabagisme demeure une menace. Le tabac tue plus de gens que l'alcool, le SIDA, les accidents de la circulation, les drogues illégales, les meurtres et les suicides réunis. Pire, des milliers de personnes supplémentaires meurent de causes liées au tabac telles que l'exposition à la fumée de tabac environnementale (appelée tabagisme passif), la consommation de tabac sans fumée et les incendies causés par des cigarettes (AU, 2013). Son expansion semble exponentielle et se renouvelle à chaque génération. La grande dépendance que crée le tabac auprès de ses consommateurs, rend son abandon difficile. Outre la dépendance, le tabagisme est responsable de nombreuses maladies dont plus d'une dizaine de cancers différents et des maladies cardiovasculaires, qu'il soit actif ou passif. Selon l'Organisation Mondiale de la Santé (OMS), le tabac est la principale cause de décès évitable dans le monde et selon ses tendances actuelles, le tabac tuerait plus d'un milliard de personnes au cours du 21^{ème} siècle.

En Europe, le tabac est également le facteur de risque de mortalité évitable le plus important. Avec 29% de fumeurs chez les adultes, l'Europe est la région du monde où la fréquence du tabagisme est la plus élevée. La part de la mortalité totale due au tabac parmi les adultes de plus de 30 ans est de 16% en Europe, contre 11% en moyenne dans le monde. Le tabac est également une cause de maladies chroniques telles que les cancers, dont les plus fréquents sont ceux des poumons, de la bouche, du sein et de la vessie. Il favorise aussi les cancers de la peau, du larynx, de l'œsophage, de l'estomac, du foie, du pancréas, de l'intestin, du rein, du côlon, du rectum et de la prostate. De plus, le tabagisme est aussi à l'origine des maladies cardiovasculaires comme l'hypertension, l'infarctus du myocarde couramment appelé « crise cardiaque » et les accidents vasculaires cérébraux paralysant la majorité de ses victimes.

En France, le tabac et ses dérivés sont consommés sur tout l'étendue du territoire national. En 2020, selon un sondage réalisé par Santé Publique France mené en métropole sur deux vagues (janvier-mars, juin-juillet) auprès des 18-75 ans, plus d'un quart des français, exactement 25,5% déclarent fumer tous les jours. Bien que les tendances actuelles du tabagisme chez les jeunes soient en baisse depuis 2019, on estime, que chaque année, environ 200.000 jeunes commencent à fumer ; l'âge d'entrée dans le tabagisme est estimé à 13-14 ans. Selon l'enquête ESCAPAD réalisée chez les jeunes de 17 ans, plus d'un quart des jeunes (25,1%) de 17 ans sont des fumeurs quotidiens et près de six jeunes Français sur 10 ont déjà essayé de fumer.

Cette situation peu reluisante de la population française, en particulier de sa jeunesse, mérite donc une réflexion plus approfondie. Dans cette option, l'analyse des facteurs amenant les individus à la consommation du tabac permettra d'identifier plus facilement ces derniers mais aussi de rendre plus efficaces les diverses politiques gouvernementales visant à éradiquer ce fléau.

I. Présentation des données

L'intégralité des données utilisées dans cette enquête proviennent des résultats de l'Enquête Annuelle sur les conditions de vie des français en 2015. Toutes les variables de cette base ne présentent pas d'intérêt pour notre étude, une première sélection de variables ayant un lien de près ou de loin avec la consommation de tabac a été effectuée. Une seconde étape de sélection via un test de chi2 a été effectué afin de sélectionner les

variables pouvant éventuellement entrer dans le modèle. Dans le cas de notre étude, la variable dépendante ‘consommation’ de tabac fait partie de la catégorie des variables qualitatives binaires. En effet, soit l’individu est consommateur de tabac soit il ne l’est pas. On cherche à modéliser la probabilité d’être un consommateur de tabac au regard de certaines caractéristiques sociodémographiques et économiques. On a

$$Y_i = \begin{cases} 1 & \text{si l'individu fume} \\ 0 & \text{sinon} \end{cases}$$

II. Statistiques descriptives

1. Analyse univariée

1.1 Consommation de tabac

La répartition des individus suivant la consommation de tabac révèle une prédominance des non consommateurs de tabac. Environ 33% fument contre 67% qui ne fument pas.

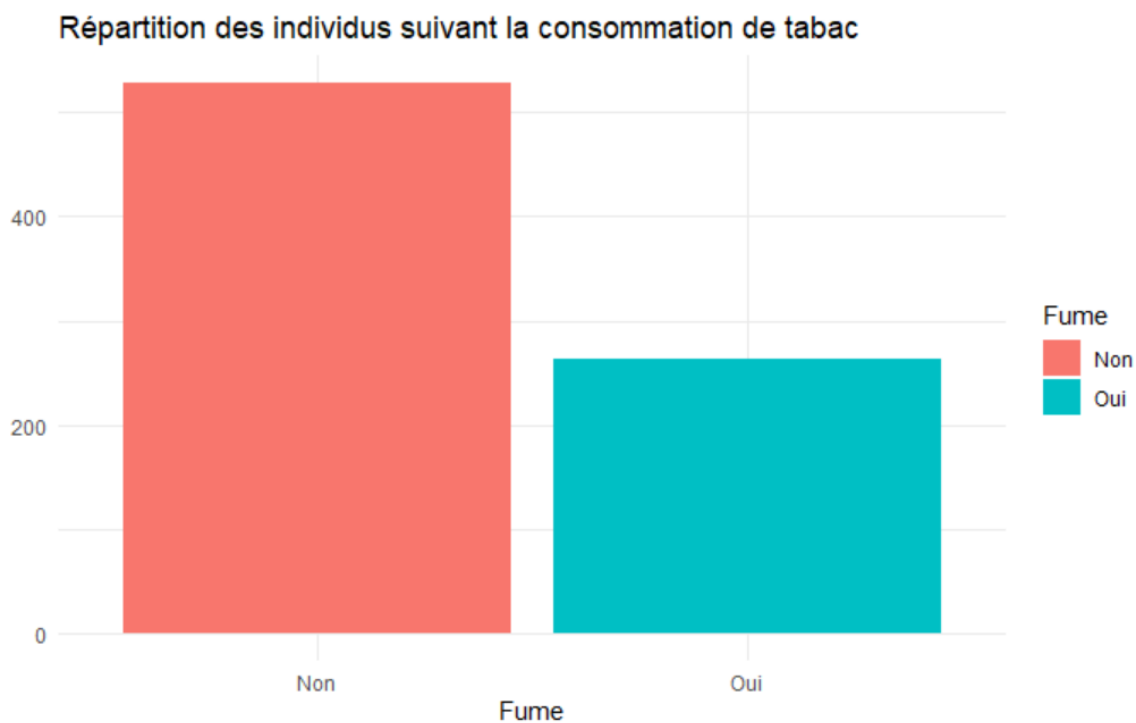


Figure 1 Répartition selon la consommation de tabac

1.2 Présentation des variables explicatives

Variables	Modalités	Hypothèses Influence	Description
Sexe	1- Homme	Référence	Indique le sexe de l'individu
	2- Femme	-	
Situation Matrimoniale	1- Célibataire	Référence	Indique la situation matrimoniale
	2- Marié	-	
	3- Divorcé	+	
	4- Veuf(ve)	+	
Situation du couple	1- Conjoint en logement	Référence	Indique la situation du couple
	2- Conjoint hors logement	+	
	3- Pas en coupe	+	
Situation travail	1- Emploi	Référence	Indique la situation de travail
	2- Apprenti(e)/Stage rénuméré	+	

	3- Etudiant(e)/Stage non rénuméré	-	
	4- Chômeur	+	
	5- Retraité	-	
	6- Femme/Homme au foyer	-	
	7- Handicap/Incapacité	-	
Situation Financière	1- S'améliorer	Référence	Indique la situation financière
	2- Sans grand changement	+	
	3- S'aggraver	-	
Religion	1- Aucune	Référence	Indique la religion
	2- Catholique	-	
	3- Protestante	-	
	4- Musulmane	-	
	5- Autre religion	-	
Situation du Ménage	1- Impossible sans dette	Référence	Indique la situation financière du ménage
	2- Difficile	-	
	3- C'est juste	-	
	4- ça va	+	
	5- Plutôt à l'aise	+	
	6- Très à l'aise	+	
Revenu	1- Moins de 1200€	Référence	Revenu de l'individu
	2- 1200€ - 2500€	-	
	3- 2500€ - 6000€	+	
	4- Plus de 6000€	+	
Alcool	1- Très souvent	Référence	Indique la fréquence à laquelle la personne boit de l'alcool
	2- Très peu	-	
	3- Jamais	-	
Activité Sportive	1- Tous les jours ou presque	Référence	Indique la fréquence à laquelle la personne fait du sport
	2- Plusieurs fois par semaine	+	
	3- Une fois par semaine	+	
	4- Une à trois fois par mois	+	
	5- Moins souvent	+	
Statut logement	1- Propriétaire/Co-propriétaire	Référence	Indique le statut de l'individu pour le logement
	2- Accédant à la propriété	-	
	3- Locataire/Sous-locataire	-	
	4- Gratuitement	-	
Enfant	1- Oui	Référence	Indique si les individus ont des enfants ou pas
	2- Non	+	
Age	Entre 18 et 75 ans		Indique l'âge de l'individu
Personne logement	1- Une seule personne	Référence	Indique le nombre d'habitant du logement
	Le nombre d'habitant : 2,3,4,5,6 ou +		
Rencontre d'ami	1- Au moins 1fois/mois	Référence	Indique la fréquence de voir des amis
	2- (1-2-3) fois/mois	+	
	3- Plusieurs fois dans l'année	-	
	4- A des occasions exceptionnelles	-	
	5- Jamais	-	
Amis proches	1- Aucun ami proche	Référence	Indique le nombre d'amis proche de l'individu.
	2- 1 ou 2 personnes	+	
	3- Entre 3 et 5 personnes	+	
	4- 6 personnes ou plus	+	
	5- Ne sait pas		
Maladie Chronique	1- Oui	Référence	Indique si la personne possède une maladie chronique
	2- Non	+	
Nervosité	1- En permanence	Référence	Indique la fréquence de nervosité de la personne
	2- Souvent	-	
	3- Quelques fois	-	

	4- Rarement	-	
	5- Jamais	-	
Poids	Varie entre 55kg et 95kg (Intervalle de 5kg)		Indique le poids de la personne
Diplôme	1- Aucun diplôme	Référence	
	2- Primaire	-	
	3- BEPC	-	Indique le niveau d'instruction de la personne
	4- CAP	-	
	5- BAC	-	
	6- Universitaire	-	
Zone géographique	1- Région parisienne	Référence	
	Les autres régions		Indique la zone géographique

Tableau 1 Les variables explicatives

2. Analyse Bivariée

2.1. V de cramer entre consommation de tabac et les variables explicatives

Le V de Cramer est une statistique utilisée pour mesurer la force de l'association entre deux variables nominales, et il prend des valeurs de 0 à 1. Les valeurs proches de 0 indiquent une association faible entre les variables et les valeurs proches de 1 indiquent une forte association entre les variables.

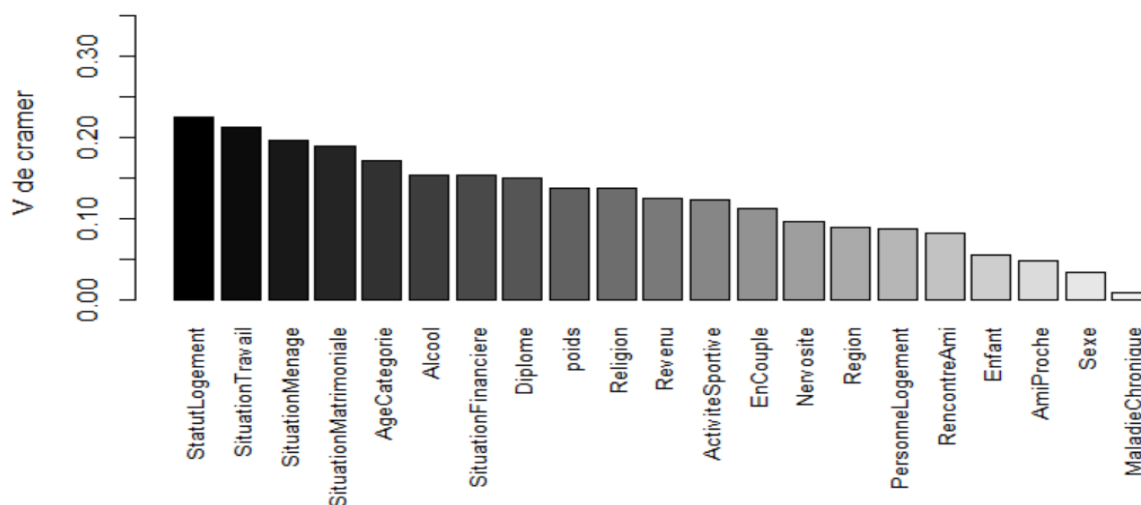


Figure 2 : V de cramer entre la consommation de tabac et les variables explicatives

2.2. Répartition de la consommation de selon quelques variables explicatives

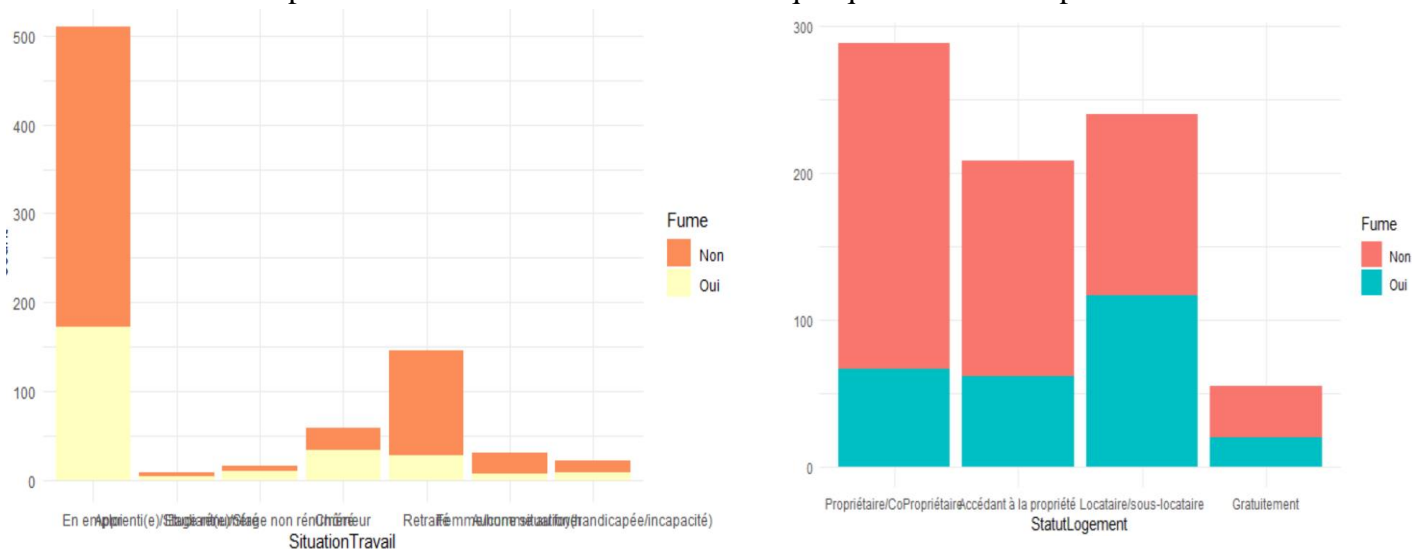


Figure 3 : Répartition de la consommation de tabac selon la situation et le logement

Les graphiques de la figure 3 montrent la répartition des consommateurs de tabac selon la situation vis-à-vis de la situation de travail et de logement (les deux variables ayant le plus grand V de cramer). Il révèle que les personnes qui fument le plus sont les apprentis et les locataires. Ainsi, le fait de devoir payer un loyer chaque mois et le fait d'être un apprenti peuvent être soupçonnés d'être de potentielles causes de consommation de tabac.

2.3. Matrice de corrélation entre les variables explicatives.

Les variables explicatives étant très nombreuses ($p = 21$), la matrice de corrélation est très peu lisible. Les valeurs pour chaque corrélation sont contenues dans la base cramer2 contenu dans le code Rmd. On a aucune corrélation entre variable qui est au-dessus de 0.8. On n'a pas de problème de forte colinéarité entre variable explicative.

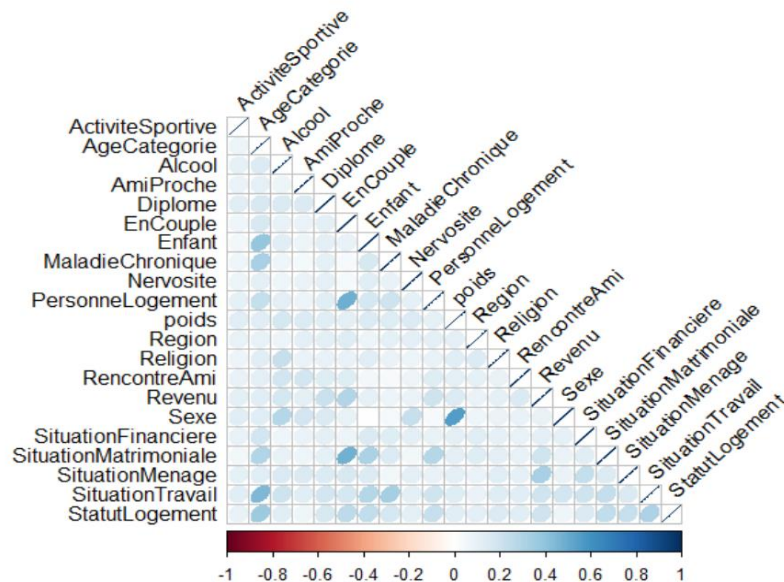


Figure 4 : Matrice de corrélation

III. Estimation des modèles

3.1. Création de l'échantillon

Dans le but de trouver le meilleur modèle possible, parmi ceux proposés par les méthodes de sélection automatique (Backward, Forward), on a créé un échantillonnage d'apprentissage pour entraîner le modèle et estimer les erreurs sur l'échantillon test. Cet rééchantillonnage a été effectué cinq (05) fois dans le but de stabiliser nos résultats.

```
set.seed(1234)
Echantillon <- mc_cv(Base, prop = 3/4, times = 5)
index.app <- Echantillon$splits[[1]]$in_id
head(index.app)
Dapp <- Base[index.app,]
Dtest <- Base[-index.app,]
```

3.2. Approche Statistique de sélection

Dans cette partie nous avons fait une approche statistique de sélection avec la méthode "Backward" et "Forward" pour les deux critères AIC et BIC afin de choisir notre modèle. Cette méthode permet de trouver le meilleur sous-groupe de variables explicatives parmi les variables explicatives possibles. Les modèles retenus par ses variables se présentent comme suit :

Logit AIC Forward	Fume ~ StatutLogement + Alcool + SituationTravail + Diplome + SituationMatrimoniale + ActiviteSportive + SituationFinanciere + SituationMenage
Logit BIC Forward	Fume ~ StatutLogement + Alcool
Logit AIC Backward	Fume ~ SituationMatrimoniale + SituationFinanciere + SituationMenage + Alcool + ActiviteSportive + StatutLogement + Diplome
Logit BIC Backward	Fume ~ Alcool + StatutLogement

Tableau 2 : Les modèles retenus par les méthodes pas à pas.

On remarque que les modèles Logit BIC Forward et BIC Backward, un seul des deux sera donc utilisé dans l'étape de sélection.

4. Choix du modèle

Dans le but de choisir un modèle parmi les trois candidats

4.1. La qualité d'ajustement

Le pseudo R2 permet de mesurer la qualité de l'ajustement.

Modèles	Pseudo R2
AIC Forward	0.1704285
AIC Backward	0.1577601
BIC	0.07389761

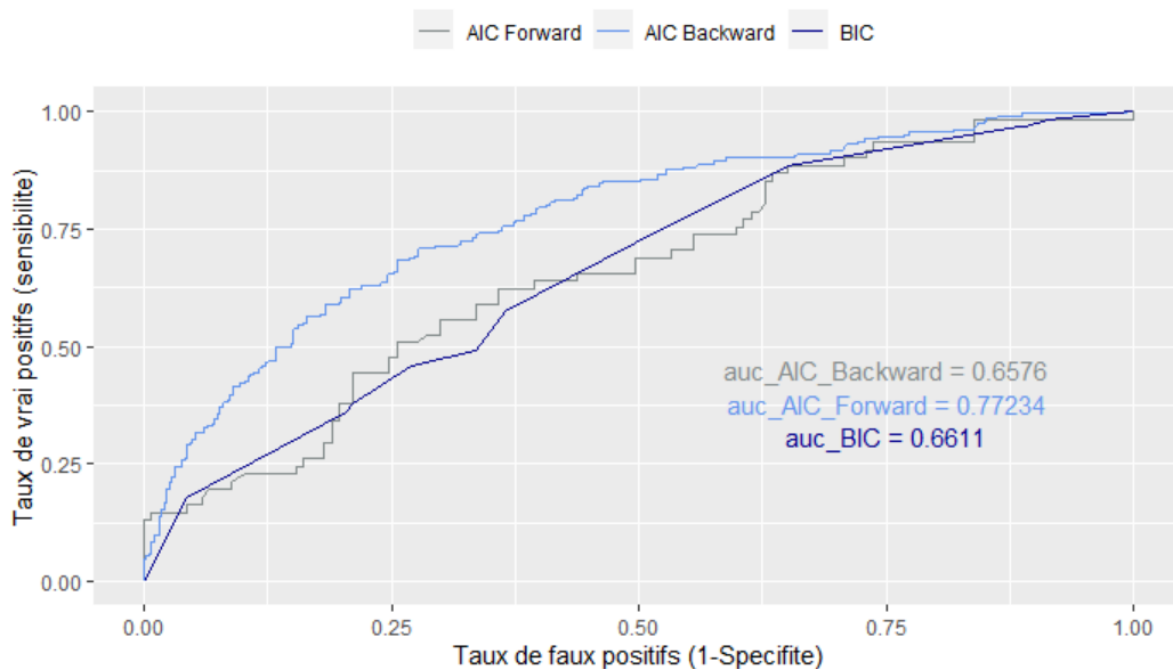
Tableau 3 : Pseudo R2

On peut voir que c'est le modèle AIC Forward qui possède la meilleure qualité de l'ajustement avec 0.17.

4.2. Courbe ROC

L'aire en dessous de la courbe ROC permet de mesurer la performance de classification binaire des modèles. L'aire en dessous de la courbe ROC obtenue pour le modèle AIC Forward est 0.77234. En effet, plus la valeur de l'AUC est proche de 1, plus le modèle a un bon pouvoir discriminatoire. Ainsi, au vu de ce critère, le modèle AIC Forward est le meilleur.

Courbes ROC



4.3. Matrice de confusion et erreur de classification

L'application des différents modèles à l'échantillon test permet d'apprécier le pouvoir prédictif des modèles au travers de l'erreur de classification

```

Prévision AIC Forward
Valeur Non Oui
Non 130 7
Oui 51 10
Prévision AIC Backward
Valeur Non Oui

```

```

Non 131 6
Oui 51 10
Prévision BIC
Valeur Non Oui
Non 137 0
Oui 61 0

```

Modèles	Erreur de classification
AIC Forward	0.2929293
AIC Backward	0.2878788
BIC	0.3080808

Tableau 4 : Erreur de Classification

On peut voir que les erreurs de classification sont très proches pour les 2 modèles AIC et plus élevées pour le modèle BIC. Ce n'est pas ce critère qui va nous permettre de sélectionner un modèle.

4.4. Déviance du modèle

La déviance d'un modèle M est définie par : $DM = 2(L_{sat} - L_n(\hat{\beta}_n))$. C'est une quantité positive. Plus elle est faible, meilleur est le modèle.

Modèles	Déviance
AIC Forward	631.1140
AIC Backward	640.7518
BIC	704.5519

Tableau 5 : Déviance

4.5. Résumé de la sélection

Modèle	Qualité d'ajustement	AUC	Erreur de prévision	Déviance
AIC Forward	***	***	**	***
AIC Backward	**	**	***	**
BIC	*	*	*	*

Tableau 6 : Résumé de la sélection

On retient donc le modèle AIC Forward. Soit le modèle

```

Fume ~ StatutLogement + Alcool + SituationTravail + Diplome + SituationMatrimoniale
+ ActiviteSportive + SituationFinanciere + SituationMenage

```

5. Tests d'adéquation du modèle

5.1. Test de Hosmer-Lemeshow

On va maintenant effectuer le test de Hosmer-Lemeshow. Les hypothèses de ce test sont :

$$\begin{cases} H_0 : \text{le modèle logistique est adéquat} \\ H_1 : \text{le modèle logistique n'est pas adéquat} \end{cases}$$

```

Hosmer and Lemeshow test (binary model)
data: Base$Fume, fitted(Modele)
X-squared = 5.3986, df = 8, p-value = 0.7143

```

On accepte H_0 car p-value supérieur à 5%. Donc le test valide le modèle logiste.

5.2. Analyse des résidus

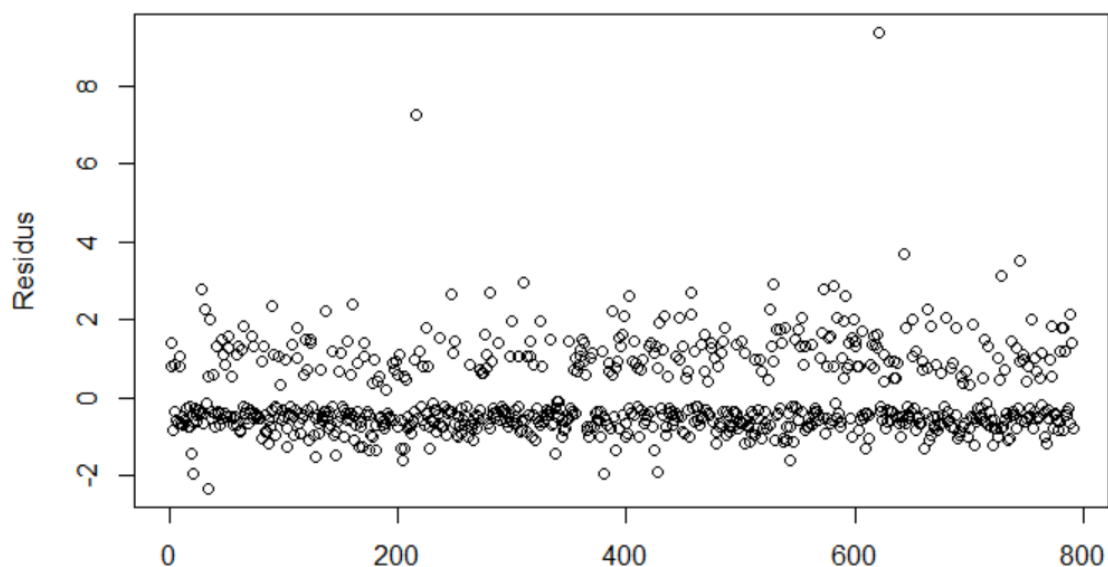


Figure 6 : Résidu de Pearson

On remarque que notre modèle permet de bien différencier les deux groupes d'individus.

6. Présentation et interprétation du modèle (Odds ratios)

6.1 Le modèle, coefficients et odds ratios

Variables	Modalités	Coef	Odds ratio	P > z	Seuil
	Constante	2.00560	7.430515	0.000594	***
Statut logement	1- Propriétaire/Co-propiétaire	Référence			
	2- Accédant à la propriété	0.06934	1.071802	0.778180	
	3- Locataire/Sous-locataire	0.71519	2.044569	0.003311	**
	4- Gratuitement	0.04167	1.042552	0.909700	
Alcool	1- Très souvent	Référence			
	2- Très peu	-0.79022	0.453745	3.80e-05	***
	3- Jamais	-1.77164	0.170055	1.42e-08	**
Situation travail	1- Emploi	Référence			
	2- Apprenti(e)/Stage rénuméré	0.21539	1.240351	0.785553	
	3- Etudiant(e)/Stage non rénuméré	0.88486	2.422633	0.138522	
	4- Chômeur	0.49663	1.643173	0.126042	
	5- Retraité	-0.52552	0.591250	0.059017	.
	6- Femme/Homme au foyer	-0.42535	0.653540	0.385589	
	7- Handicap/Incapacité	0.22792	1.255984	0.651613	
Diplôme	1- Aucun diplôme	Référence			
	2- Primaire	-1.71844	0.179346	0.010649	*
	3- BEPC	-0.30412	0.737776	0.524837	
	4- CAP	-0.66270	0.515456	0.109239	
	5- BAC	-0.26747	0.765311	0.455766	
	6- Universitaire	-0.72899	0.482397	0.044296	*
Situation Matrimoniale	1- Célibataire	Référence			
	2- Marié	-0.51070	0.600074	0.013080	*
	3- Divorcé	-0.25471	0.775138	0.339193	
	4- Veuf(ve)	0.04159	1.042471	0.931497	
Activité Sportive	1- Tous les jours ou presque	Référence			
	2- Plusieurs fois par semaine	-0.75097	0.471908	0.003185	**
	3- Une fois par semaine	-0.86436	0.421319	0.002200	**
	4- Une à trois fois par mois	-0.21720	0.804771	0.485269	

	5- Moins souvent	-0.60774	0.544580	0.013521	*
Situation financière	1- S'améliorer	Référence			
	2- Sans grand changement	-0.60172	0.547867	0.021499	*
	3- S'aggraver	-0.74014	0.477046	0.021752	*
Situation du ménage	1- Impossible sans dette	Référence			
	2- Difficile	0.08488	1.088584	0.813966	
	3- C'est juste	-0.60153	0.547974	0.079101	.
	4- ça va	-0.61972	0.538097	0.087254	.
	5- Plutôt à l'aise	-0.52497	0.591575	0.198936	
	6- Très à l'aise	-1.19020	0.304159	0.097885	.

Tableau 7 : Le modèle, coefficients, odds ratios

***0.1% ; ** 1% ; * 5% ; .10%

6.2. Interprétation du modèle

✚ Statut logement

On peut en sortir du modèle estimé pour la variable statut logement que les locataires ont 2 fois plus de chance d'être fumeur qu'un propriétaire. Cela peut s'expliquer vu que le logement ne leur appartient pas. Ils se sentent plus libres pour fumer chez eux. Aussi on peut aussi expliquer ça par le stress du loyer à payer à chaque fin de mois.

✚ Consommation d'alcool

Il ressort de nos analyses que les consommateurs d'alcool occasionnels ou ceux qui n'en consomment jamais ont respectivement 0,79 et 1,77 fois plus de chance d'être de non consommateur de tabac. Cela pourrait s'expliquer par le fait que c'est très facile de joindre l'alcool au tabac.

✚ Situation au travail

On remarque pour la situation au travail que ce sont les retraités qui ressortent de notre analyse. Ils ont 0.59 moins de chance de fumer que les personnes ayant un emploi. Aussi on remarque que les personnes au foyer fument aussi moins. Cela peut s'expliquer que pour le cas des personnes retraités, elles doivent faire attention à leur santé. Des personnes en apprentis ou étudiant auront tendance à plus fumer qu'une personne ayant un emploi. Cela s'explique que les personnes en apprentissage ou étudiant sont des jeunes. Et les jeunes aujourd'hui ont tendance à plus fumer que les générations d'avant.

✚ Niveau d'instruction

Des analyses, il ressort que le niveau d'instruction a un effet négatif sur la consommation de tabac. En effet, plus la personne est diplômée, moins elle est susceptible de fumer. Cela pourrait s'expliquer par le fait que les personnes les plus instruites sont plus informées des conséquences liées au tabagisme, actif ou passif.

✚ Situation Matrimoniale

Les analyses montrent que les mariées sont moins susceptibles de fumer comparer aux célibataires. Aussi les veuf(ve) sont 1,04 fois plus susceptibles d'être des consommateurs de produits tabagique comparé aux célibataires et que les mariés sont ceux qui fument le moins. Cela pourrait s'expliquer par la solitude que récent généralement les célibataire et veuf(ve).

✚ Activité sportive

Ici on peut voir avec notre modèle que plus une personne fait du sport quotidiennement fumera plus qu'une autre personne. Cela peut s'expliquer par le stress de l'exigence du sport de haut niveau.

✚ Situation financière

Notre modèle montre que plus on a une situation financière aisée, plus la personne fumera. Ce qui paraît logique, une personne avec plus de moyens financiers pourra se permettre plus facilement de se payer des paquets de cigarette qu'une personne avec moins de finance. Une situation financière favorable aura donc plus tendance à fumer que les autres

✚ Situation du ménage

On peut en sortir du modèle que pour la situation financière du ménage, plus la situation est compliquée plus la personne fumera. Cela peut s'expliquer par le stress financier de devoir ramener de l'argent pour combler les besoins du ménage.

7. Scoring

7.1 Table des scores.

Le tableau suivant présente la table des scores

Variables	Modalités	Score
Activité Sportive	1- Tous les jours ou presque	10
	2- Plusieurs fois par semaine	1
	3- Une fois par semaine	0
	4- Une à trois fois par mois	7
	5- Moins souvent	3
Alcool	1- Très souvent	20
	2- Très peu	11
	3- Jamais	0
Diplôme	1- Aucun diplôme	19
	2- Primaire	0
	3- BEPC	16
	4- CAP	12
	5- BAC	16
	6- Universitaire	11
Situation Financière	1- S'améliorer	8
	2- Sans grand changement	2
	3- S'aggraver	0
Situation Matrimoniale	1- Célibataire	6
	2- Marié	0
	3- Divorcé	3
	4- Veuf(ve)	6
Situation du Ménage	1- Impossible sans dette	13
	2- Difficile	14
	3- C'est juste	7
	4- ça va	6
	5- Plutôt à l'aise	7
	6- Très à l'aise	0
Situation du travail	1- Emploi	6
	2- Apprenti(e)/Stage rénuméré	8
	3- Etudiant(e)/Stage non rénuméré	16
	4- Chômeur	11
	5-Retraité	0
	6- Femme/Homme au foyer	1
	7- Handicap/Incapacité	8
Statut logement	1- Propriétaire/Co-propriétaire	0
	2- Accédant à la propriété	1
	3- Locataire/Sous-locataire	8
	4- Gratuitement	0

Tableau 8 : Table des scores

Ici nous avons le tableau des scores des modalités de variable du modèle. On peut en ressortir que les individus qui fument ont les caractéristiques ci-dessous :

- ⇒ Consommer beaucoup d'alcool ;
- ⇒ Sans diplôme/des étudiants/chômeur ;
- ⇒ Situation financière de ménage compliqué.

7.2 Seuils des scores

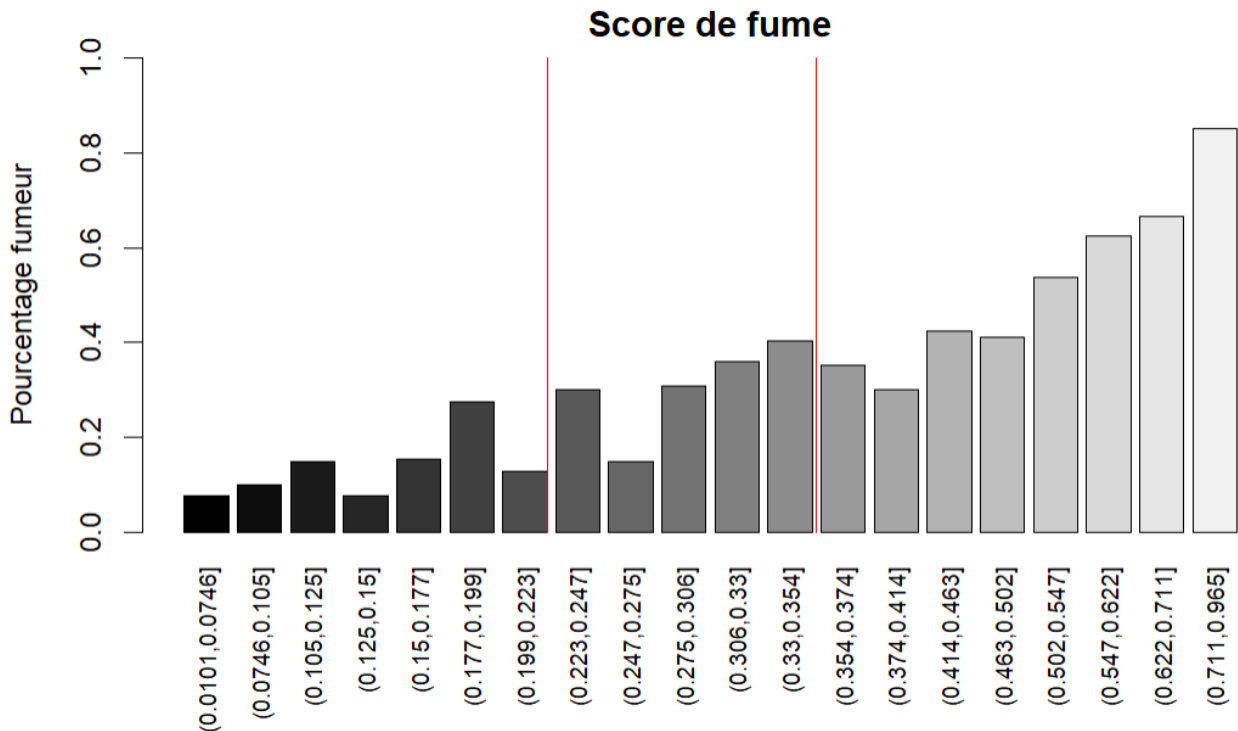


Figure 7 : Seuil de score

Cette figure représente le seuil de score à laquelle la personne fumera. On peut voir qu'une personne qui a un score plus de 0.35 est un individu qui a beaucoup de chance d'être fumeur. Une personne qui a un score en dessous de 0.22 aura tendance à pas fumer.

Conclusion

La consommation de tabac et ses dérivés, malgré les conséquences négatives qui en découlent reste une réalité prenant de l'ampleur dans nos sociétés. La présente étude dont l'objet est d'analyser les facteurs explicatifs de la consommation des produits tabagiques s'est fixé deux principaux objectifs. Le premier consiste à mettre en évidence les facteurs socio-économiques influençant la consommation de tabac en France et le second vise à déterminer le profil le plus susceptible du Français consommateur de tabac. A l'issue des analyses, il ressort que les consommateurs de tabac appartiennent le plus souvent à la catégorie des personnes qui n'ont pas de diplôme. Ce sont des personnes non instruites, consommant de l'alcool et professionnellement non occupées ou en étude ou encore en apprentissage. Ce sont donc plus des profils de jeunes personnes. L'activité sportive quotidienne serait un facteur pour que la personne fume. La vie en solitude a tendance à être le profil type d'un fumeur.

Enfin, nous suggérons aux pouvoirs publics d'accorder une place de choix à la lutte anti-tabac à travers le renforcement des mesures d'éducation en mettant l'accent sur les méfaits liés à la consommation de tabac au cours des leçons d'instruction civique et morale dans les écoles ou encore par de larges campagnes de sensibilisation portant sur les effets nocifs de la consommation de tabac sur tout le territoire national.

Annexes

I. Etapes de sélection du modèle

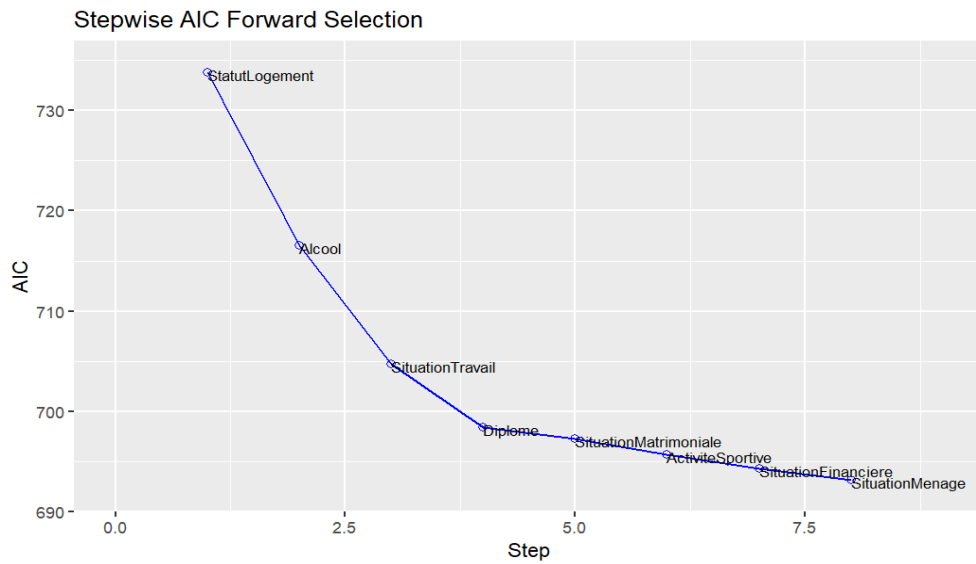


Figure 8 : Sélection des variables pour le modèle

II. Importance des variables dans le modèle

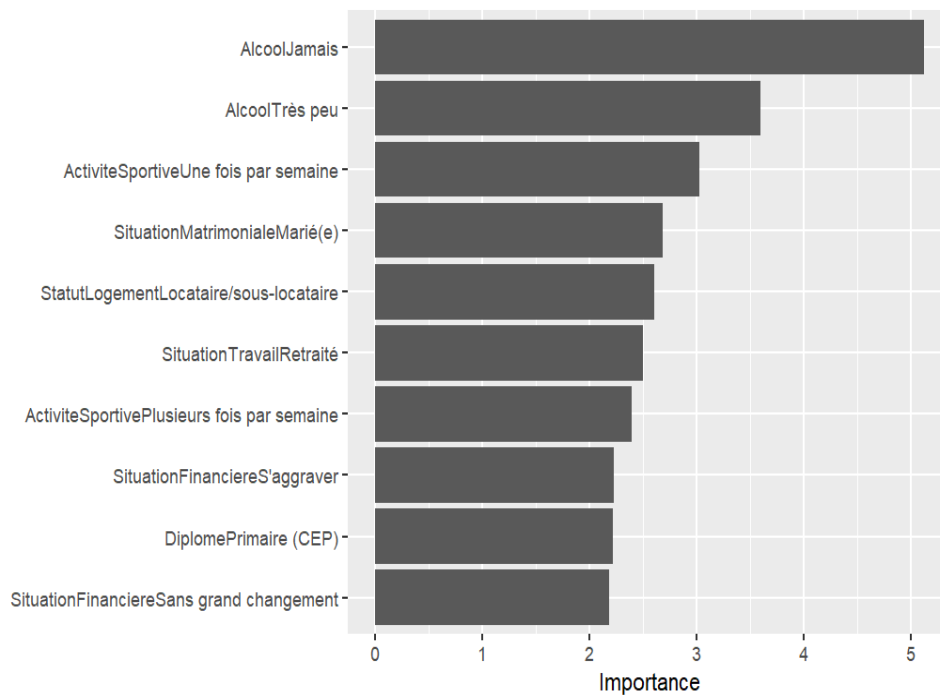
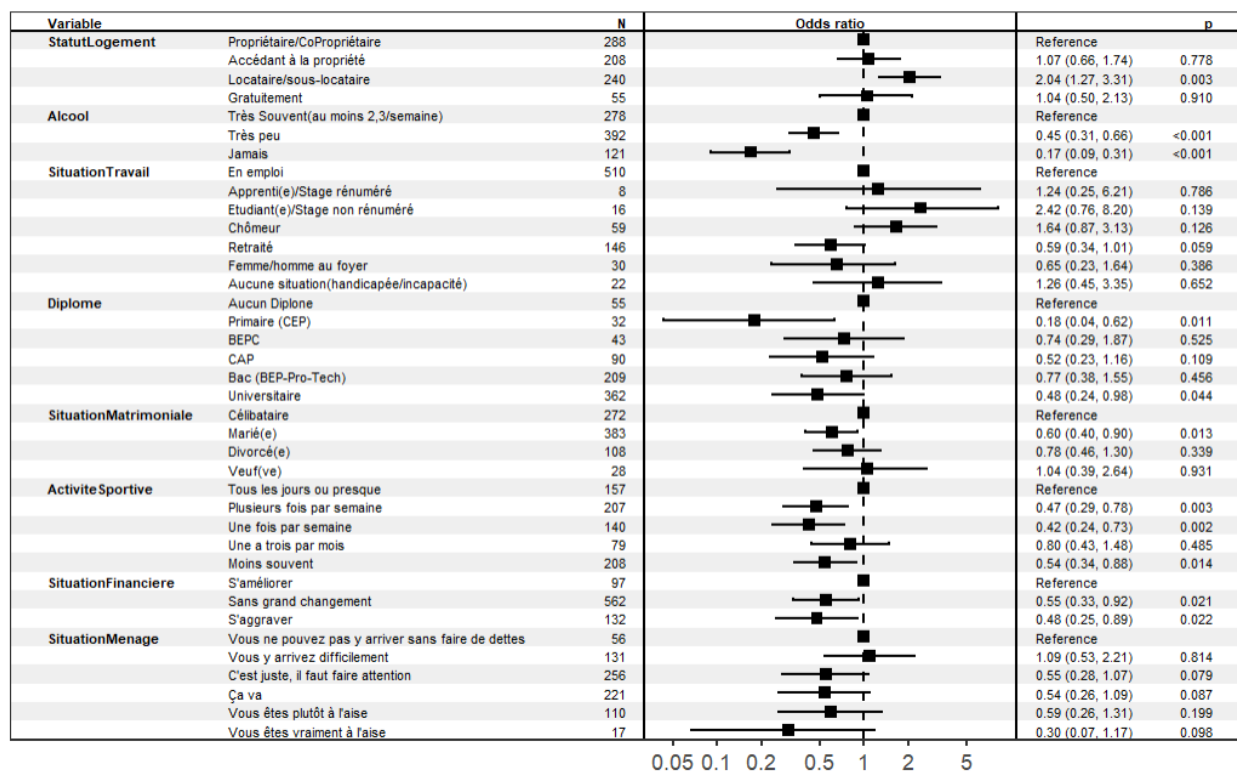


Figure 9 : Variable Importance Plot

III. Graphique odds ratios



Webographie

[https://www.cairn.info/revue-population-et-societes-2019-10-page-](https://www.cairn.info/revue-population-et-societes-2019-10-page-1.htm#:~:text=Avec%2029%20%25%20de%20fumeurs%20chez,dans%20le%20monde%20%5B7%5D.)

[1.htm#:~:text=Avec%2029%20%25%20de%20fumeurs%20chez,dans%20le%20monde%20%5B7%5D.](https://www.cairn.info/revue-population-et-societes-2019-10-page-1.htm#:~:text=Avec%2029%20%25%20de%20fumeurs%20chez,dans%20le%20monde%20%5B7%5D.)