

# L'IMPACT DE LA CONSOMMATION D'ALCOOL SUR LE NIVEAU DE BONHEUR DES PAYS

Dossier d'économétrie



Année universitaire 2020 - 2021

Stefano Bianchini

SCHOTT Victor, LOBET Corentin, RICHARD Mélody

Janvier 2021

# Sommaire

Introduction : description de l'étude et des effets attendus.

I/ Présentation des données

II/ Analyse exploratoire des données : interprétation

- Matrice des corrélations
- Relations linéaires

III/ Régression

- Les résultats de la régression par la méthode des MCO
- Les résultats OLS avec des écarts-types robustes
- Effet de la consommation d'alcool

IV/ Vérification des hypothèses de Gauss-Markov

Conclusion : résumé, présentation du diagnostic obtenu et limites

## Introduction

L'objectif de notre étude est d'analyser l'impact de la consommation d'alcool sur le niveau de bonheur des pays. Pour cela, nous avons utilisé une variable « Happiness score » (HS) pour effectuer notre étude, elle provient d'une enquête historique sur l'état du bonheur dans le monde. Nous avons donc choisi de former une base de données qui recense l'état de bonheur de 106 pays en se concentrant sur les données de 2016.

A des fins de contexte, le rapport sur le bonheur mondial de 2017 a été pour sa part publié aux Nations unies lors d'un événement célébrant la Journée internationale du bonheur le 20 mars. Il classe 155 pays en fonction de leur niveau de bonheur. C'est un événement qui a permis de faire gagner en notoriété cette étude permettant de théoriser une notion très subjective de « bonheur » en un score significatif par pays. Les gouvernements utilisent alors de plus en plus cet indicateur dans leur choix de décisions politiques. En effet, c'est une notion qui permet de lier tant les domaines de l'économie, de la sociologie, des statistiques ou encore de la psychologie. Grâce à cette diversité, l'indicateur (HS) permet de préciser les mesures du bien-être dans une population donnée et également de pouvoir évaluer les progrès d'un pays dans son objectif de bien-être.

Pour ces raisons, nous avons trouvé très intéressant de pouvoir utiliser la « science du bonheur » afin de pouvoir expliquer des différences entre pays dans des notions plus communes et théoriques que sont le niveau de croissance d'un pays.

Pour aller plus loin, une manière intéressante d'expliquer le bonheur d'une population serait de le lier à sa consommation d'alcool. En effet sociologiquement, l'alcool est un facteur d'échanges et d'interactions sociales important, alors qu'il peut aussi être très néfaste s'il n'est pas consommé avec modération. D'où la question à laquelle nous essayerons de répondre lors de ce rapport : « La consommation d'alcool a-t-elle un impact sur le niveau de bonheur des pays ? »

Nous nous attendons alors à ce que la consommation d'alcool ait un impact positif sur le niveau de bonheur selon une courbe en cloche, c'est-à-dire qu'il serait un facteur permettant d'améliorer le bien-être d'une population jusqu'à un certain stade critique où il serait trop présent et important, ce qui réduirait alors le bonheur total.

## I. Présentation des données

La variable « Happiness Score » (HS) utilisée est construite de la manière suivante : selon un sondage accordé à un échantillon donné, elle mesure les 6 facteurs étant la production économique ; le soutien social ; l'espérance de vie ; la liberté ; l'absence de corruption ; et la générosité. Tout ceci comparé à un niveau de « dystopie » représenté par un pays hypothétique ayant des valeurs égales aux moyennes nationales des pays les plus basses du monde. Au final, ceci forme des niveaux permettant de classer les pays.

Nous avons alors choisi nos propres déterminants de ce niveau de bonheur (HS) afin d'en évaluer la significativité :

- La consommation d'alcool exprimée en doses bar (AC)
- L'indice de Développement Humain (HDI)
- Le PIB par tête (GPD)
- L'indice de Gini (GI)
- La richesse des régions
  - pays développées (Rich) : Europe de l'ouest ; Amérique du Nord ; Australie et Nouvelle-Zélande
  - pays sous-développés (Poor) : Afrique Sub-saharienne ; Asie du Sud-Est
  - pays en développement sinon : Afrique du Nord et Moyen Orient ; Amérique Latine et Caraïbes ; Asie de l'Est ; Europe centrale et de l'est

Les données ont été manipulées à partir des données suivantes :

- [Gini index \(World Bank estimate\)](#)
- [Happiness and Alcohol Consumption \(Kaggle\)](#)

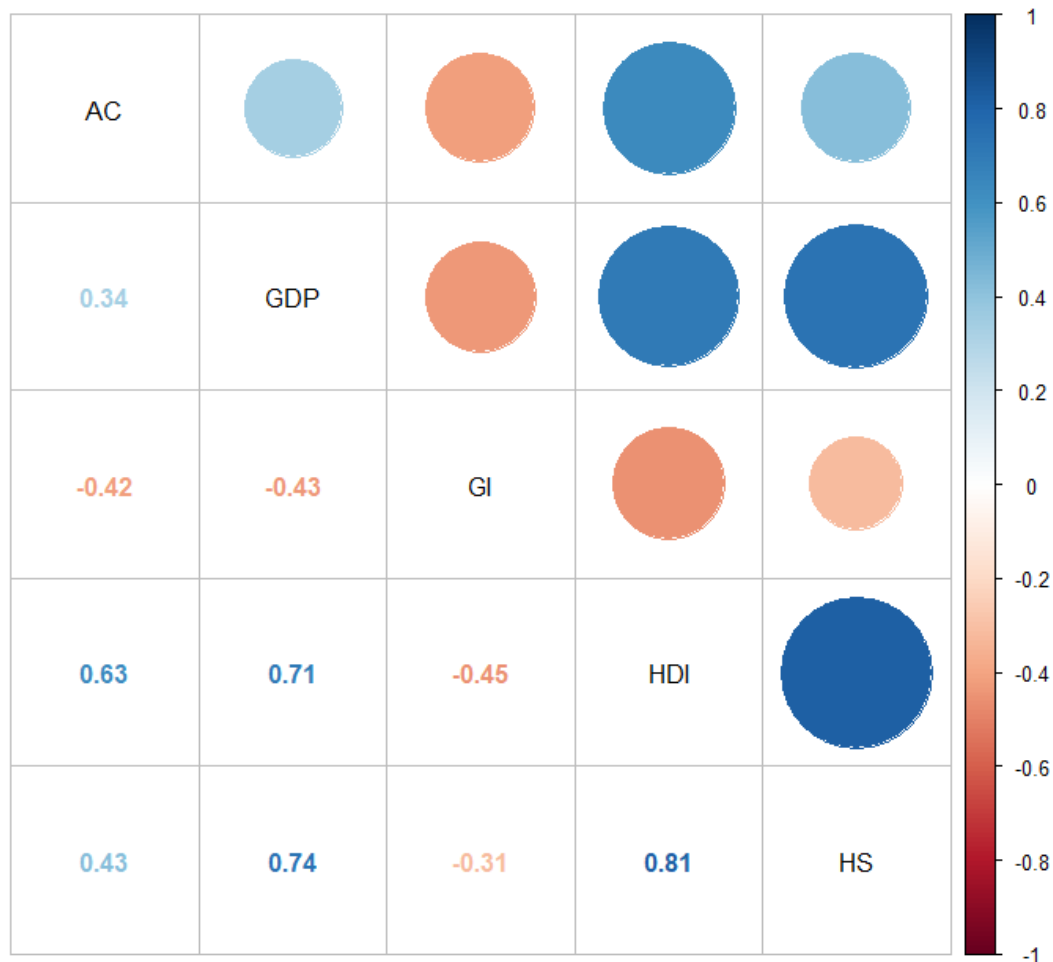
La valeur de l'indice de Gini est représenté par la moyenne de 2010-2019 afin d'obtenir des informations manquantes sur les données de 2016. La richesse des régions est déterminée par la moyenne des PIB des pays qui les composent. Ce choix a été fait pour éviter les problèmes des multicollinéarité qui surviendraient si l'on se contentait de faire de chaque région du monde une variable binaire.

À des fins de simplifications, seuls les groupes « pays développés » et « pays pauvres » seront étudiés car ce sont ceux qui prennent le dessus sur les autres, en effet ils correspondent à 2 extrêmes intéressants à comparer. Ces régions sont alors exprimées en tant que variables binaires. C'est-à-dire que dans la colonne des régions, une valeur de 1 pour un pays signifie qu'il appartient à cette région géographique ; et au contraire, une valeur de 0 signifie qu'il n'y appartient pas. L'effet pour un pays appartenant à une région intermédiaire sera compris dans la constante.

Toutes nos variables continues sont exprimées en logarithme afin de pouvoir interpréter les résultats en termes d'élasticités.

## II. Analyse exploratoire des données

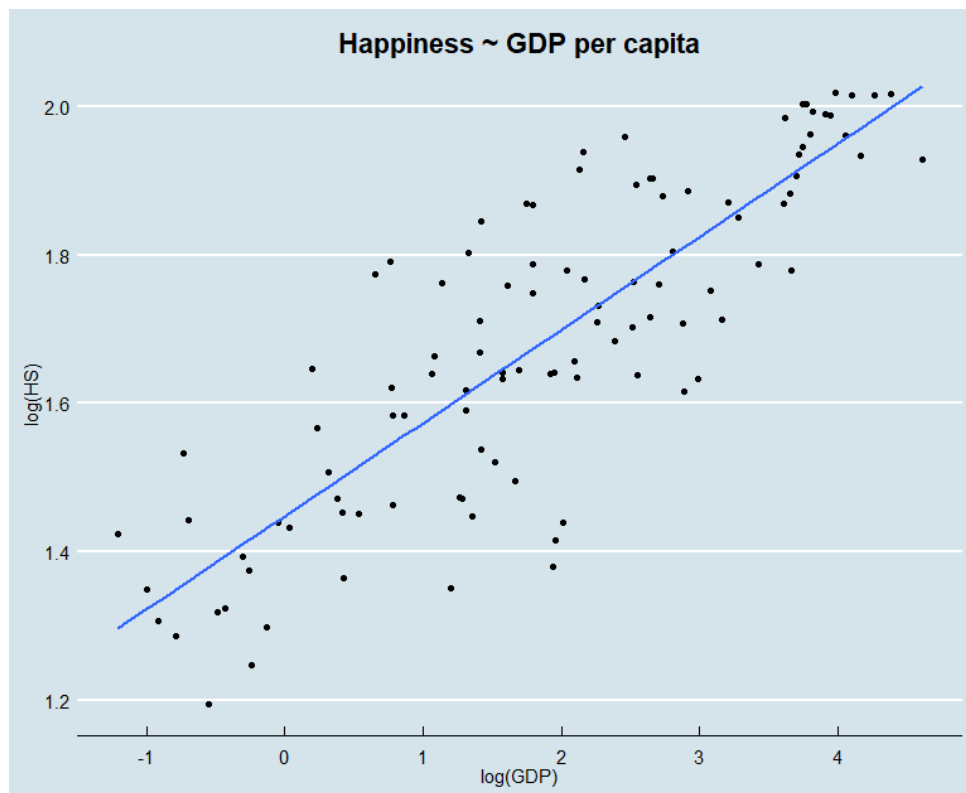
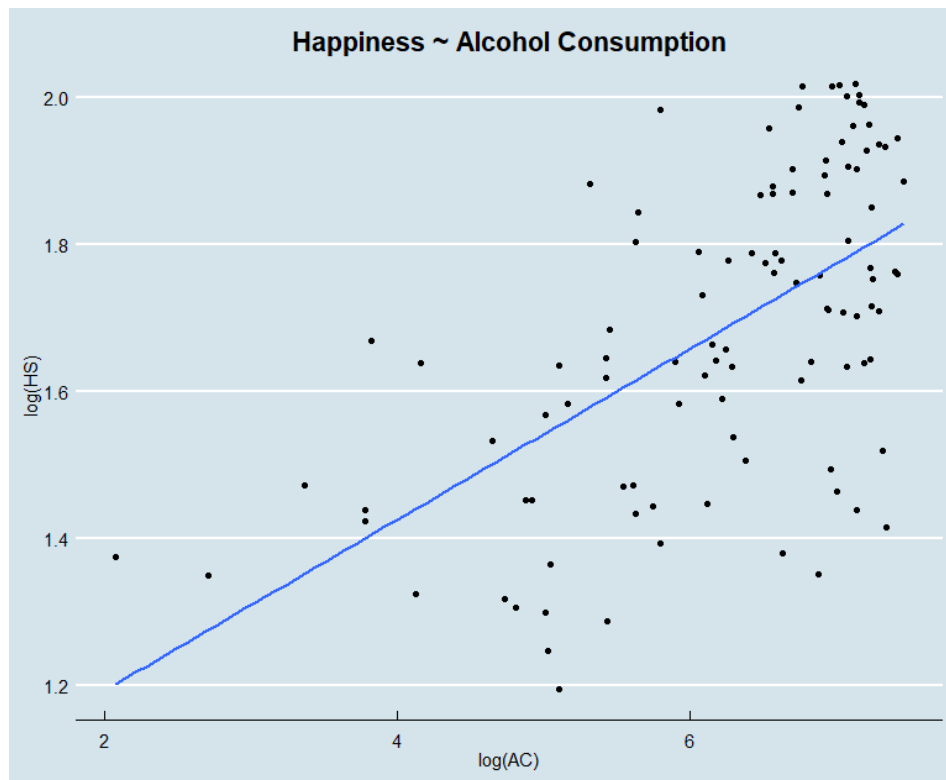
### A. Matrice des corrélations

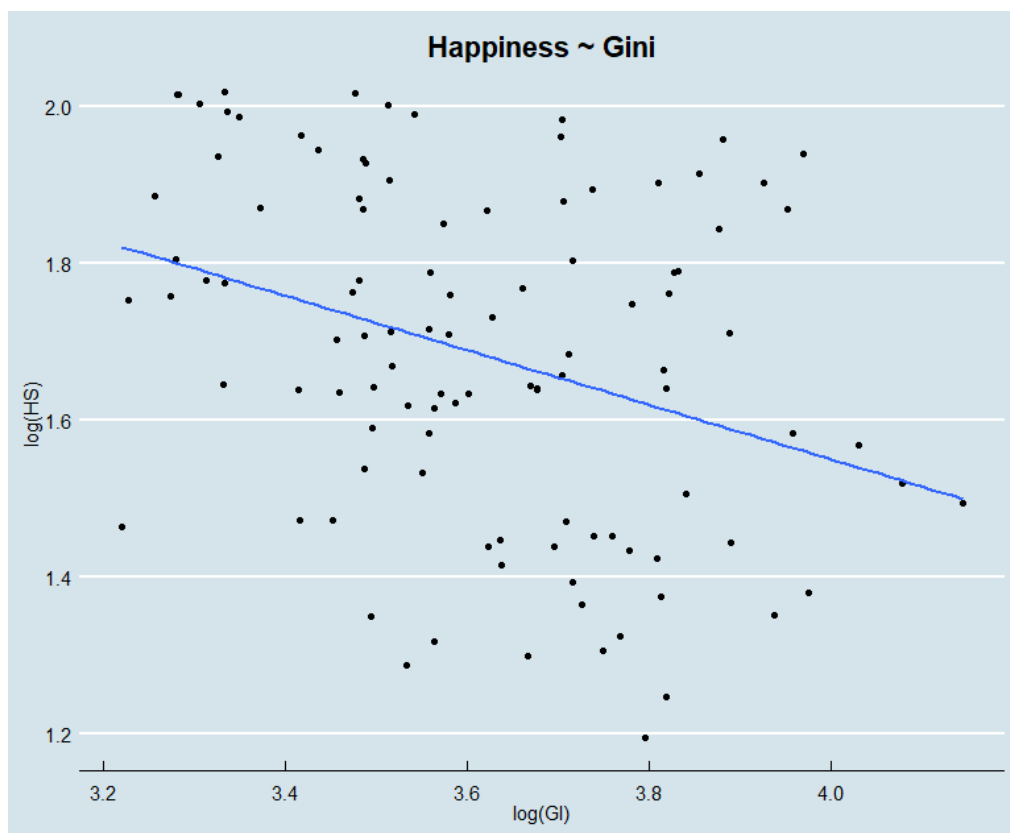
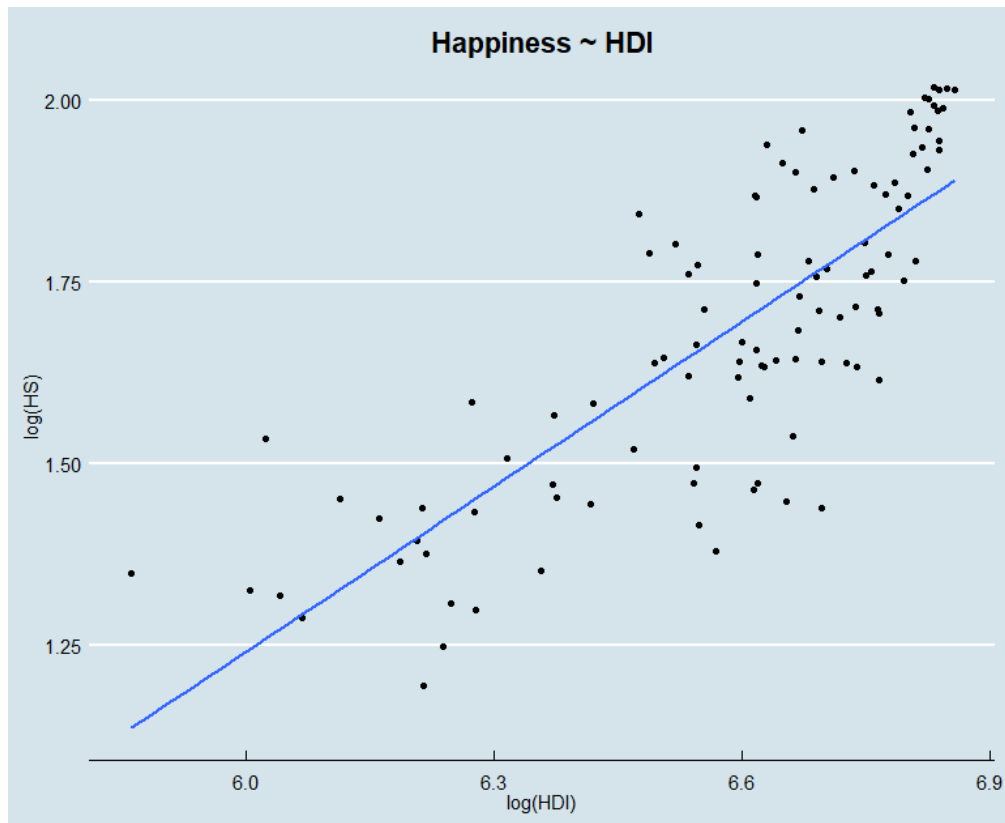


Les corrélations se révèlent être les plus fortes pour le PIB et l'IDH, tous deux fortement corrélés puisque le calcul de l'IDH tient compte du PIB. Nous remarquons aussi une corrélation positive de 0.43 avec la consommation d'alcool et une corrélation négative de -0.31 avec l'indice de Gini, un résultat attendu puisque le bonheur agrégé ne peut qu'être plus faible dans un pays souffrant de fortes inégalités de revenus.

### B. Relations linéaires

Comme le montrent les graphiques ci-dessous, toutes les spécifications log-log entre le score de bonheur et les variables explicatives continues présentent une linéarité non négligeable.





### III. Régression

Ainsi la forme complète de notre modèle est la suivante :

$$\ln(HS) = a + \beta_1 \ln(AC) + \beta_2 \ln(GDP) + \beta_3 \ln(HDI) + \beta_4 \ln(GI) + \beta_5 Rich + \beta_6 Poor + \varepsilon$$

#### A. Les résultats de la régression par la méthode des MCO

En utilisant la méthode OLS pour notre modèle, on trouve les résultats suivants pour nos différentes variables :

	Béta	Ecart-types	t-test	P-value	Significativité
Constante	1.084	1.035	1	2.977489 e-01	
AC	-0.005	0.016	-0.3	7.687891 e-01	
GDP	0.1	0.024	4.2	6.342716 e-05	***
HDI	0.027	0.158	0.2	8.662658 e-01	
GI	0.081	0.063	1.3	2.033055 e-01	
Rich	0.023	0.041	0.6	5.698659 e-01	
Poor	-0.188	0.04	-3	3.725445 e-03	***

#### Qualité de l'ajustement

R <sup>2</sup>	$\overline{R^2}$	Fisher	P-value
0.740099	0.7243474	46.98571	7.792555 e-27

On peut alors remarquer que les variables ajoutées au modèle sont pertinentes car elles l'expliquent plutôt bien étant donné que le R<sup>2</sup> est élevé pour une valeur de 74%. D'un autre côté, elles ne sont pas toutes significatives selon cette régression. En effet on peut remarquer que seul le PIB et le fait d'être dans un continent pauvre ont un effet significatif. Ainsi, une hausse de 1% du revenu par tête augmenterait le bonheur agrégé de 0.1 %. Si le pays est situé dans une région du monde pauvre, son bonheur agrégé sera 0.118% inférieur à un pays situé dans une région en développement.

#### Test d'hétéroscédasticité

R <sup>2</sup> auxiliaire	N	Breush-Pagan	P-value
0.08162079	106	8.651803	0.194128

Le test de Breush-Pagan n'aboutit pas au rejet de l'hypothèse d'homoscedasticité mais le graphique de la relation entre bonheur et alcool montre qu'il pourrait y avoir de l'hétéroscédasticité.



## B. Les résultats OLS avec des écarts-types robustes

En présence d'hétéroscédasticité, il convient de modifier le modèle pour résoudre ce problème. On introduit alors des écarts-types robustes nous permettant d'avoir sensiblement plus confiance dans les résultats.

	Béta	Ecart-type	t-test	P-value	Significativité
<i>Constante</i>	0.99	0.868	1.1	2.567096 e-01	
<i>AC</i>	-0.01	0.015	-0.7	4.999201 e-01	
<i>GDP</i>	0.098	0.022	4.5	2.193676 e-05	***
<i>HDI</i>	0.051	0.135	0.4	7.078789 e-01	
<i>GI</i>	0.079	0.063	1.2	2.151014 e-01	
<i>Rich</i>	0.021	0.038	0.6	5.788865 e-01	
<i>Poor</i>	-0.125	0.053	-2.3	2.090308 e-02	**

Les résultats sont sensiblement similaires à l'exception que la significativité de la variable Poor a un peu diminuée.

## C. Effet de la consommation d'alcool

Si les variables GI, HDI et AC ne sont pas significatives c'est probablement qu'elles sont captées par le PIB. En effet, la matrice des corrélations révèle des relations importantes entre le PIB et les autres variables. Souhaitant tout de même avoir une idée de l'effet de l'alcool sur la population, nous proposons ci-dessus des régressions n'incluant pas les variables de contrôle.

$$\ln(HS) = a + \beta_1 \ln(AC) + \varepsilon$$

### Estimation OLS

	Béta	Ecart-type	t-test	P-value	Significativité
<i>Constante</i>	0.961	0.097	9.9	1.057902 e-16	***
<i>AC</i>	0.116	0.015	7.6	1.670636 e-11	***

### Qualité de l'ajustement

$R^2$	$\overline{R^2}$	Fisher	P-value
0.3545597	0.3483536	57.13032	1.670636 e-11

### Test d'hétéroscédasticité

$R^2$ auxiliaire	N	Breush-Pagan	P-value
0.001167861	106	0.1237932	0.7249563

Conclusion : il n'y a pas d'hétéroscédasticité.

### Estimation OLS avec des écarts-types robustes

	Béta	Ecart-type	t-test	P-value	Significativité
<i>Constante</i>	0.961	0.09	10.7	1.765932 e-18	***
<i>AC</i>	0.116	0.014	8.1	1.024755 e-12	***

Avec ou sans correction du biais d'hétéroscédasticité, l'effet de l'alcool sur le bonheur agrégé est statistiquement significatif et révèle qu'une hausse de 1% de la consommation d'alcool induit une hausse de 0.116% du bonheur agrégé. La qualité de l'ajustement est relativement intéressante pour une régression linéaire simple (35%).

En complément, puisqu'il est considéré que la consommation d'alcool présente un effet opposé lorsqu'elle est abusive, nous analysons une régression quadratique :

$$HS = a + \beta_1 AC + \beta_2 AC^2 + \varepsilon$$

#### Le test de Student des coefficients

	Estimation	Erreur standard	t-test	P-value	Significativité
Constante	1.3996 e+00	3.8908 e-02	35.9732	< 2.2 e-16	***
AC	6.76 e-04	1.2307 e-04	5.4928	2.869 e-07	***
AC <sup>2</sup>	-2.7434 e-07	7.5977 e-08	-3.6108	0.0004734	***

#### Qualité de l'ajustement

R <sup>2</sup>	$\overline{R^2}$	Fisher	P-value
0.3892986	0.3774403	37.82927	9.336586 e-12

L'effet à la moyenne de la consommation d'alcool par personne (750 doses bar) d'un verre supplémentaire est une hausse de  $0.000676 - 2 \times 0.000000274 \times 450 = 0.00043$  du score de bonheur. En France l'effet d'un verre supplémentaire serait de 0.00011 sur le bonheur, bien plus faible car les français consomment deux fois plus d'alcool que la moyenne mondiale. Les pays les plus friands d'alcool comme la République Tchèque et l'Allemagne verraient plutôt une baisse de leur bien-être s'ils augmentaient leur consommation.

## IV. Vérification des hypothèses de Gauss-Markov

Les hypothèses de Gauss-Markov nous permettent d'utiliser la méthode OLS avec un estimateur non biaisé. Respecter ces conditions permet de vérifier que nos coefficients sont proches de la réalité car en moyenne l'estimateur converge vers sa vraie valeur. Et cet estimateur est BLUE donc il représente le meilleur estimateur non biaisé.

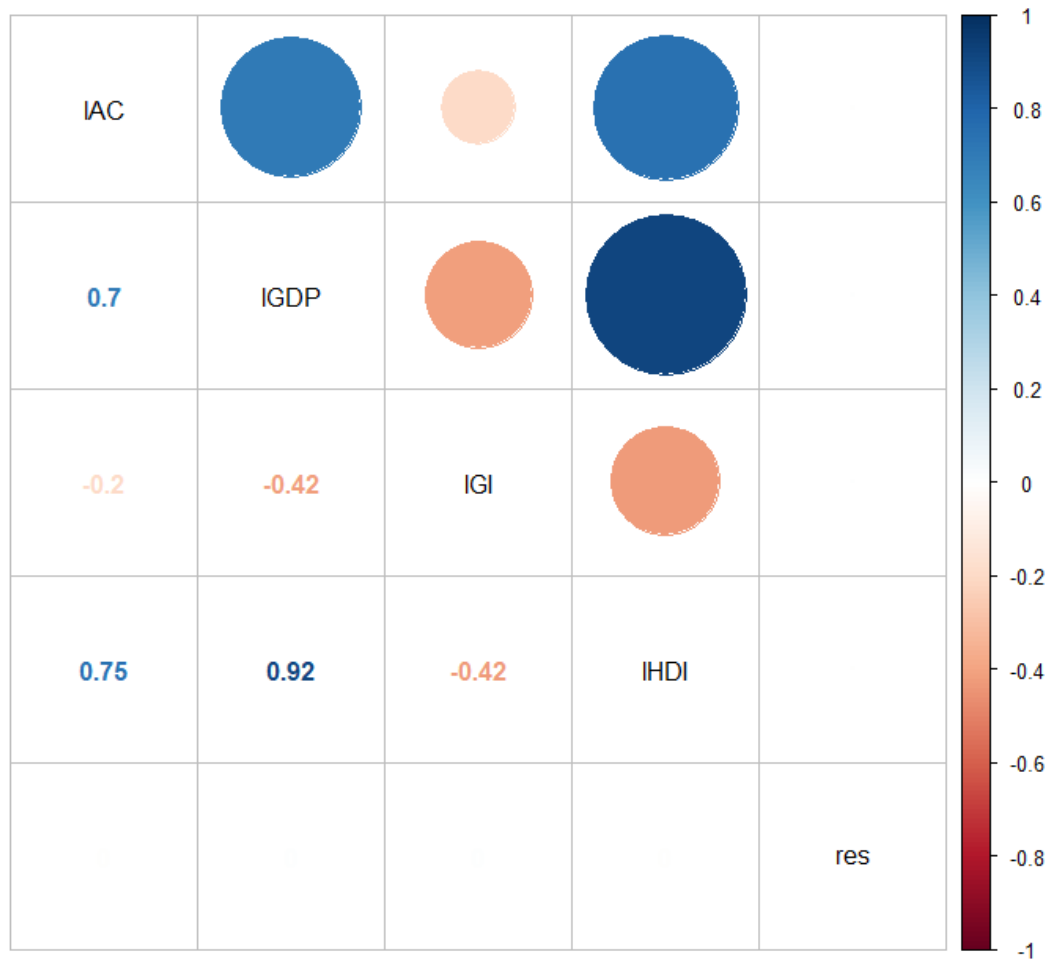
A1 : L'espérance du terme d'erreur est nulle

A2 : Les termes d'erreurs du modèle sont indépendants des variables (exogénéité) c'est-à-dire que :  $(x_1, \dots, x_i)$  est indépendant de  $(e_1, \dots, e_i)$

A3 : La variance des termes d'erreurs est constante, c'est le principe d'homoscédasticité

A4 : La covariance des termes d'erreurs est nulle donc on a absence d'autocorrélation

Les potentiels problèmes d'hétéroscédasticité ont déjà été traités dans la partie précédente. La présence d'autocorrélation est incohérente dans ces données individuelles non ordonnées. Nous pouvons encore regarder s'il y a un biais d'endogénéité dans notre modèle. La matrice de corrélations ci-dessous révèle que les coefficients de corrélation entre les variables et les résidus du modèle tendent vers 0.



## Conclusion

Pour répondre à notre problématique, nous avons effectivement pu montrer qu'il existe une relation positive entre la consommation d'alcool et le niveau de bonheur d'un pays. En plus de cela, selon les pays, si cette consommation d'alcool est trop forte et atteint un certain niveau, le bien-être de la population peut réduire.

Nous pouvons avoir confiance en ces résultats car notre estimateur OLS vérifie bien les conditions de Gauss-Markov et il est BLUE.

Pour faire un parallèle avec l'actualité, nous pouvons nous demander si le bien-être des pays ne serait-il pas à la baisse étant donné les mesures sanitaires actuelles conduisant à des confinements, couvre-feu et fermeture de lieux culturels tels que les bars-restaurants. Il s'agirait alors de voir si cette fermeture a conduit à une baisse de la consommation d'alcool et de ce fait, à une baisse du niveau de bonheur des pays.

Limite : on trouve un effet de l'alcool sur le bonheur mais lorsque le PIB est inclus dans le modèle, il constitue la variable la plus corrélée à l'indicateur du bonheur donc son introduction va rendre les autres variables non significatives.