Techniques quantitatives – feuille d’exercice 3

La base de données utilisée comprend des informations sur la consommation d’alcool dans différents pays du monde, ainsi que sur la mortalité associée à l’alcool.

Les données sont issues du projet « *Our world in Data »*, soutenu par l’université d’Oxford (<https://ourworldindata.org/alcohol-consumption>).

On cherchera à connaître la relation entre les différentes variables et à tenter de construire une explication de la mortalité associée à l’alcool.

# Statistiques bivariées – une variable qualitative et une variable quantitative

## Statistiques par groupe

Calculez la moyenne, la médiane et l’écart-type de la consommation de vin, de bière et de spiritueux par continent.

Représentez ces informations sur un ou plusieurs graphiques.

Qu’observe-t-on ?

# Statistiques bivariées – deux variables quantitatives

## Covariance

Calculez la covariance entre la consommation de vin et celle de bière. Que pouvez-vous en dire ?

Comparez-la à la covariance entre la consommation de vin et celle de spiritueux.

## Corrélation

Calculez la corrélation entre toutes les variables de consommation (vin, bière, spiritueux), deux à deux. Qu’observez-vous ?

Calculez la corrélation entre les variables de consommation et la mortalité. Que pouvez-vous en dire ?

Calculez le reste des corrélations entre les variables. Qu’observez-vous ? Commentez un fait intéressant.

## Nuage

Faites le graphique du nuage de points entre la consommation de vin et la part des femmes qui boivent.

Idem pour la consommation de bière.

Tracez les droites de corrélation et ajoutez les R².

Qu’observez-vous ? Il y-t-il des valeurs extrêmes à exclure ?

Faites l’analyse uniquement sur les pays européens.

## Concentration

Calculez l’indice de Gini de la consommation de vin. Peut-on légitimement dire que certains pays consomment pour les autres ?

# Régression

## Linéaire simple

Calculez la pente (*a*) et l’ordonnée (*b*) de l’équation :

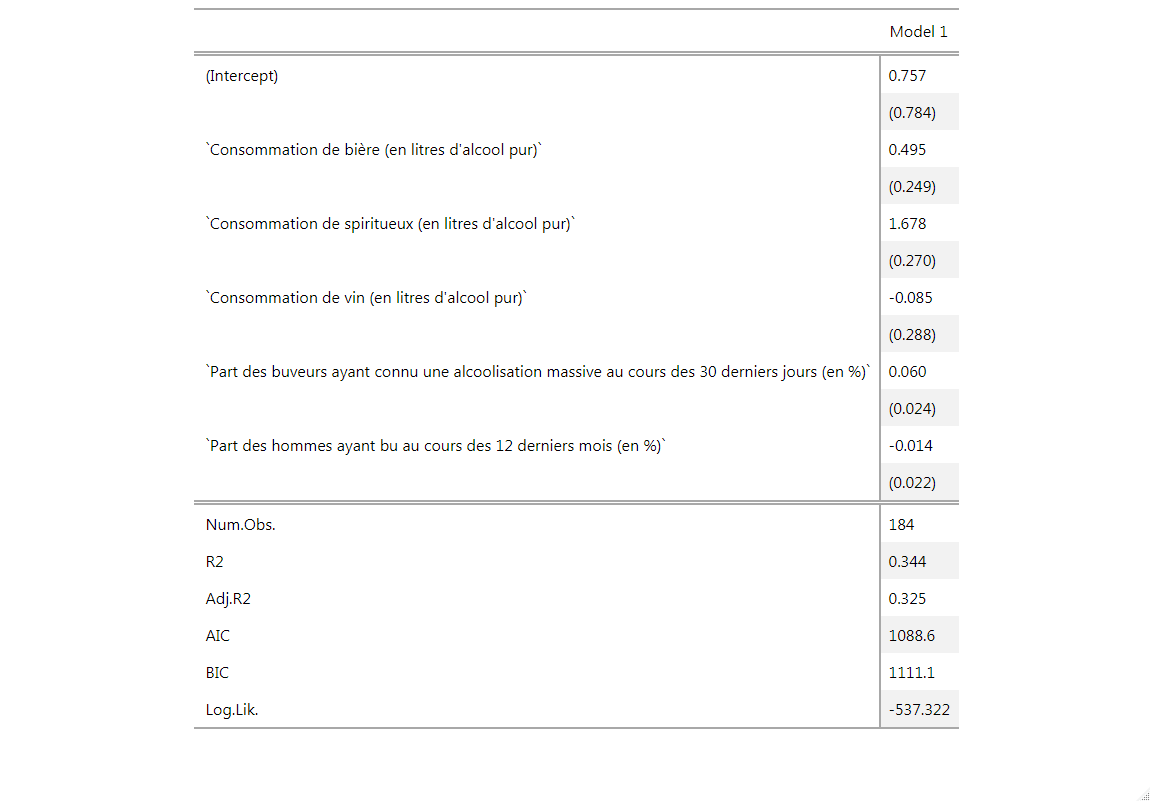
Mortalité = *a* \* alcoolisation massive + *b*

Calculez le coefficient de détermination.

## Linéaire multiple

Le tableau suivant présente les résultats d’une régression linéaire multiple à partir de la formule suivante :

Mortalité = *a1*\* alcoolisation massive + *a2* \* part des hommes qui boivent + *a3* \* consommation de vin + *a4* \* consommation de bière + *a5* \* consommation de spiritueux + *b*



Commentez ce tableau. Comment expliquez-vous la mortalité due à l’alcool ?