

Fiche de TD1b : Boîte à outils du microéconomiste 2

Exercice 1 : Toutes choses étant égales par ailleurs

Un groupe de psychologues de l'éducation a étudié le comportement de 84 étudiants de la Florida State University afin d'identifier les facteurs influençant leur performance académique.

Les deux tableaux ci-dessous montrent comment la moyenne générale (notée de 0 à 4) varie avec le temps d'étude. Dans le premier tableau, les étudiants sont répartis en deux groupes selon leur temps d'étude. Dans le second tableau, ils sont répartis en quatre groupes selon leur temps d'étude *et* leur environnement d'étude.

Tableau 1 : Temps d'étude et performance

Temps d'étude	Long (42 étudiants)	Court (42 étudiants)
Moyenne générale	3.43	3.36

Tableau 2 : Temps d'étude, environnement et performance

Temps d'étude	Long (42 étudiants)	Court (42 étudiants)
Bon environnement	3.63 (11 étudiants)	3.43 (31 étudiants)
Mauvais environnement	3.36 (31 étudiants)	3.17 (11 étudiants)

Questions :

1. En comparant les informations des deux tableaux, expliquez pourquoi il est important de réfléchir aux variables qui sont *retenues constantes* (ou supposées identiques entre les étudiants) lorsqu'on étudie la relation entre les heures d'étude et les performances aux examens.
2. En plus de l'environnement d'étude, quels autres facteurs pensez-vous qu'il faudrait idéalement inclure dans l'analyse de la relation entre les heures d'étude et la moyenne générale ?

Exercice 2 : Prix des loyers dans les villes universitaires

On cherche à estimer le prix moyen des loyers dans les villes universitaires américaines. Pour cela, on estime la régression suivante :

$$\log(\text{loyer}) = \beta_0 + \beta_1 \log(\text{pop}) + \beta_2 \log(\text{revmoyen}) + \beta_3 \text{pctetudiants} + \varepsilon$$

où *loyer* désigne le loyer mensuel moyen des appartements aux États-Unis, *pop* représente la population totale de la ville considérée, *revmoyen* désigne le revenu moyen par habitant, *pctetudiants* renvoie au pourcentage d'étudiants parmi la population totale de la ville, par exemple, 10%. *Rappel* : Le logarithme permet d'interpréter directement les coefficients en termes de pourcentage de variation.

1. Interprétez les coefficients de la régression et indiquez les signes attendus.

On cherche à savoir si le pourcentage d'étudiants présents dans la ville influe sur le niveau des loyers. À partir des données de 1990 sur 64 villes universitaires américaines, on obtient l'estimation suivante :

$$\log(\text{loyer}) = 0.043 + 0.066 \times \log(\text{pop}) + 0.057 \times \log(\text{revmoyen}) + 0.0056 \times \text{pctetudiants} + \varepsilon$$

2. Si la taille de la population augmente de 10%, de combien de pour cent augmente le prix des loyers ?
3. Si le pourcentage d'étudiants passe de 10% à 30%, de combien de pour cent augmente en moyenne le prix des loyers ?

Exercice 3 : Établir la causalité

Vous décidez de mener une expérience. Vous invitez 50 amis à une fête. Vous en sélectionnez 25 au hasard et vous leur annoncez que la nourriture sera gratuite ; la plupart d'entre eux viennent à votre fête. Les 25 autres amis ne sont pas informés que la nourriture sera gratuite ; aucun d'entre eux ne se présente. En vous fondant sur la corrélation observée, vous concluez que la nourriture gratuite incite les gens à venir aux fêtes. Un ami vous prévient toutefois : « Attention, corrélation n'implique pas causalité. » Comment devez-vous répondre ?

Exercice 4 : Corrélation ou causalité

Vous tombez sur une étude qui a mis au jour une corrélation entre la lecture et l'espérance de vie : les grands lecteurs vivent plus longtemps.

1. Identifiez au moins une explication plausible pour que cette corrélation existe même en l'absence de lien de causalité directe.