

Séquence 4 : Tableaux croisés sous Excel et Rstudio

Cet énoncé s'appuie sur les données de l'article suivant :

Gaële Henri-Panabière, 2010, « Élèves en difficultés de parents fortement diplômés », *Sociologie*, N°4, vol. 1, 457-477, <http://journals.openedition.org/sociologie/652>

[A] Contexte et nature de l'étude

A partir de l'article TexteTD4.pdf, répondez rapidement aux questions suivantes :

- 1) Quel est le champ de l'étude ?
- 2) Quelle est l'unité statistique ici étudiée ?
- 3) Quelle est l'année de la collecte ?
- 4) Qui a collecté les données ?
- 5) Quelle est la taille de l'échantillon ?

[B] Prise en main de la base de données

- 1) Ouvrez le document DicVarTD4.docx et lisez attentivement les premières questions. Pour l'entête 1, quelle est la variable étudiée ? Quelles sont les modalités ? De quel type de variable s'agit-il ?
- 2) Vérifiez la disponibilité de Microsoft Excel ou à défaut de LibreOffice ou d'un autre tableur gérant les tableaux croisés dynamiques sur votre machine. Le cas échéant, rendez-vous sur <https://www.microsoft.com/fr-fr/education/products/office> en y indiquant votre adresse de courrier électronique @etu.univ-paris1.fr. Vous devriez pouvoir vous connecter sur le site microsoft.com et avec vos identifiants de Paris 1 et y utiliser une version en ligne du logiciel Office365. Alternativement, rendez-vous sur le site <https://fr.libreoffice.org/> afin de télécharger et installer la version de libreoffice compatible avec votre plate-forme. Répondez si ce n'est déjà fait au sondage <https://cours.univ-paris1.fr/mod/feedback/view.php?id=507502>
- 3) Ouvrez le classeur Excel DonneesTD4.xlsx. Si vous êtes sous LibreOffice, ne tenez pas compte du message d'erreur vous indiquant que les données ne peuvent pas être complètement affichées. À quoi correspondent les colonnes ? Et les lignes ? Les données vous paraissent-elles analysables en l'état ?

[C] Tableau croisé dynamique

- 1) Insérez un tableau croisé dynamique. Il y a deux variables de difficultés scolaires : vérifiez qu'elles sont bien identiques.
- 2) Faites un tri croisé des données entre les difficultés scolaires et le rang de naissance. Copiez les données du tableau (valeurs uniquement) dans un nouvel onglet.
- 3) Procédez au calcul des pourcentages en ligne dans un nouveau tableau
- 4) Évaluez la force du lien puis testez sa significativité entre le rang de naissance et les difficultés scolaires.

[D] Graphique des écarts à l'indépendance

- 1) Tracez dans le tableur les graphiques des écarts à l'indépendance.
- 2) Commentez rapidement les résultats.