

## Laboratoire 4 : Distribution des moyennes d'échantillon et vérification des hypothèses

Au cours des trois laboratoires précédents, vous avez pu vous familiariser avec les données du Programme pancanadien d'évaluation (PPCE) 2013 élaboré par le Conseil des ministres de l'Éducation (Canada). Vous avez réalisé des statistiques descriptives sur les données démographiques et de rendement des élèves, en plus de standardiser les données de rendement pour mieux comprendre ces scores. Dans ce laboratoire, nous nous concentrerons sur l'utilisation des données pour répondre à des questions de recherche!

### 1. Préparation

Les méthodes quantitatives sont normalement utilisées pour les analyses de confirmation. Ainsi, une fois qu'une question de recherche est définie, la prochaine étape est de préparer l'hypothèse (en général, cette hypothèse est basée sur les travaux de recherche).

Certains travaux de recherche donnent à croire que le statut socio-économique des élèves est lié à leur rendement scolaire. Plus précisément, ils semblent indiquer qu'un statut socio-économique plus élevé est étroitement lié à un rendement scolaire plus élevé. Nous allons maintenant tester cette théorie.

Même si, dans le cadre du PPCE 2013, les élèves n'étaient pas explicitement invités à indiquer leur statut socio-économique, il y avait des questions (p. ex., SQ08. Environ combien de livres y a-t-il chez toi?) pouvant être interprétées comme une mesure observable du statut socio-économique. La question de recherche à laquelle vous allez répondre est la suivante : les élèves ayant un statut socio-économique plus élevé, tel que mesuré par le nombre de livres chez eux, réussissent-ils mieux à un test de rendement en sciences? (Remarque : il s'agit d'une question directionnelle.)

### Questions :

1. Veuillez énoncer, sous forme de phrase courte, l'hypothèse nulle et l'hypothèse alternative.

$H_0$  : \_\_\_\_\_

$H_1$  : \_\_\_\_\_

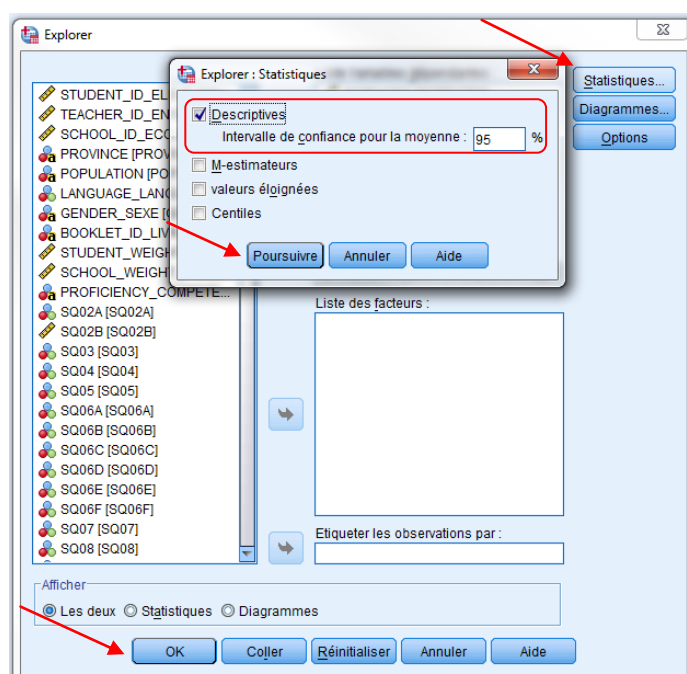
2. Veuillez indiquer vos variables dépendantes et indépendantes.

### 2. Compréhension des données

Il existe de nombreuses façons d'analyser statistiquement les données pour repérer des différences. Dans ce laboratoire, nous nous concentrerons sur l'utilisation des bornes

inférieure et supérieure de l'intervalle de confiance à 95 % pour les moyennes. La première étape est de comprendre les données en fonction des groupes de statut socio-économique. Veuillez créer un tableau des statistiques descriptives (c.-à-d., taille d'échantillon, moyenne, variance, intervalle de confiance à 95 % [borne inférieure et supérieure], minimum, maximum, plage, intervalle interquartile) pour chaque groupe de statut socio-économique (p. ex., 0 à 10 livres, 11 à 25 livres). Pour ce faire, veuillez procéder comme suit : Analyse → Statistiques descriptives → Explorer → déplacez la variable dépendante vers la boîte « Liste Variables dépendantes » et déplacez la variable indépendante vers la boîte « Liste des facteurs » → cliquez sur Statistiques → veillez à ce que la case « Descriptives » soit cochée → veillez à ce que l'intervalle de confiance pour la moyenne soit réglé à 95 % → cliquez sur Poursuivre → cliquez sur OK.

**\* Veuillez vous assurer que les pondérations des élèves sont utilisées tout au long des analyses.\***



Le tableau ci-dessous pourrait vous être utile lors de la création de votre tableau.

Tableau 1.

*Titre (veuillez créer un titre pour ce tableau)*

Livres	<i>n</i>	Moyenne	Variance	Minimum	Maximum	Plage	Intervalle interquartile	Intervalle de confiance à 95 % pour la moyenne	
								Inf.	Sup.
0 à 10									
11 à 25									
26 à 100									
Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.

### Questions :

3. Que représente l'intervalle de confiance à 95 % pour la moyenne? (Vous aurez peut-être besoin d'utiliser Google ou votre manuel pour trouver la réponse.)
4. Veuillez fournir le diagramme de quartiles. Dans le diagramme, veuillez étiqueter l'intervalle interquartile, le maximum, le minimum et les éventuelles erreurs aberrantes. (Vous aurez peut-être besoin d'utiliser Google pour déterminer les différences entre le minimum et les valeurs aberrantes.)

### 3. La réponse à la question de recherche

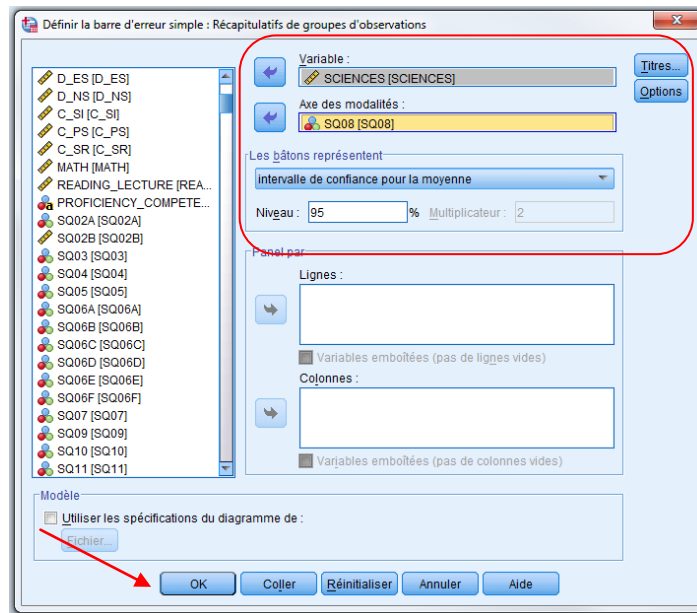
Même si le diagramme ci-dessus donne une analyse descriptive des scores de rendement des élèves en fonction de leur statut socio-économique, il n'indique toutefois pas s'il existe des différences statistiquement significatives entre les groupes. Pour cerner la différence, il faut analyser les bornes supérieure et inférieure de l'intervalle de confiance à 95 %.

### Question :

5. En examinant l'intervalle de confiance à 95 % pour la moyenne, existe-t-il un chevauchement entre l'un ou l'autre de vos intervalles parmi les groupes de statut socio-économique?

La meilleure façon de déterminer s'il y a un chevauchement entre les intervalles, c'est d'examiner les valeurs réelles, comme celles que vous avez présentées dans le Tableau 1. Cependant, une image communique plus facilement les tendances qui peuvent être évidentes. Pour ce faire, procédez comme suit : Graphes → Boîtes de dialogue ancienne version →

Barre d'erreur → sélectionnez Simple → cliquez sur Définir → déplacez Sciences dans la boîte « Variable » et SQ08 dans l'« Axe des modalités » → veillez à ce que « Intervalle de confiance pour la moyenne » soit réglé à 95 % → cliquez sur OK.



### Questions :

6. Votre nouveau diagramme confirme-t-il votre réponse à la question 5?
7. Quelle tendance avez-vous dégagée de ce graphe d'intervalle de confiance à 95 %?
8. Veuillez répondre à la question de recherche initiale et confirmer votre réponse avec des données probantes à l'appui (p. ex., des références à des statistiques tirées des tableaux et/ou diagrammes). La réponse doit tenir en trois phrases **maximum**.

### 4. À votre tour

Vous lisez des travaux de recherche qui semblent indiquer que la première langue des élèves est liée à leur rendement dans des évaluations en lecture. Vous souhaitez tester cette théorie à l'aide des données du PPCE 2013, car l'ensemble de données contient de l'information sur la première langue des élèves (c.-à-d., SQ04) et le rendement en lecture (c.-à-d.

READING\_LECTURE). La question de recherche qui oriente cette enquête est la suivante : les élèves ayant de premières langues différentes obtiennent-ils des résultats différents à un test de rendement en lecture? (Remarque : il s'agit d'une question non directionnelle.)

### Questions :

Veuillez répondre aux questions 1, 2, 5, 7 et 8, en plus d'inclure un tableau des statistiques descriptives des sections 2 et 3 ci-dessus. (Votre réponse à la question 8 peut être **un peu** plus longue.)

Vous avez maintenant terminé le laboratoire.