

Laboratoire 6 – Test T pour deux échantillons appariés

Dans le laboratoire précédent, nous avons travaillé sur les données du Programme pancanadien d'évaluation (PPCE) 2013 à l'aide de techniques d'analyse inférentielle qui ciblent les différences entre deux groupes d'élèves. Dans le présent laboratoire, nous approfondirons des techniques d'analyse inférentielle pour deux groupes **dépendants** d'élèves (par exemple, une mesure préintervention et une mesure postintervention portant sur les mêmes élèves). ****Les données additionnelles utilisées dans ce laboratoire sont simulées et visent à vous permettre de travailler avec des données préintervention et postintervention. Puisqu'elles ne sont qu'hypothétiques, elles ne peuvent pas être utilisées en dehors de ce laboratoire.****

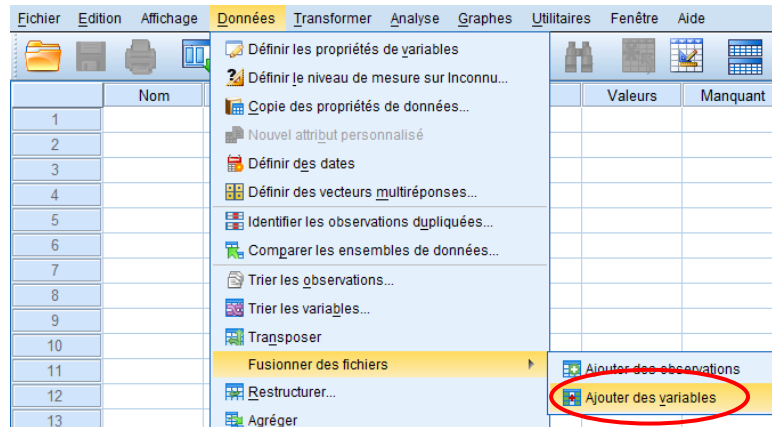
Scénario – Une fois la première évaluation du PPCE terminée, le personnel administratif des écoles s'est rendu compte, à la lumière des résultats, que l'intimidation était problématique en Alberta. Une campagne intensive de sensibilisation a donc été lancée pour lutter contre l'intimidation à l'école. Les élèves qui avaient initialement répondu aux items du PPCE 2013 sur l'intimidation ont de nouveau répondu à ces mêmes items après la campagne.

1. Fusion de deux ensembles de données

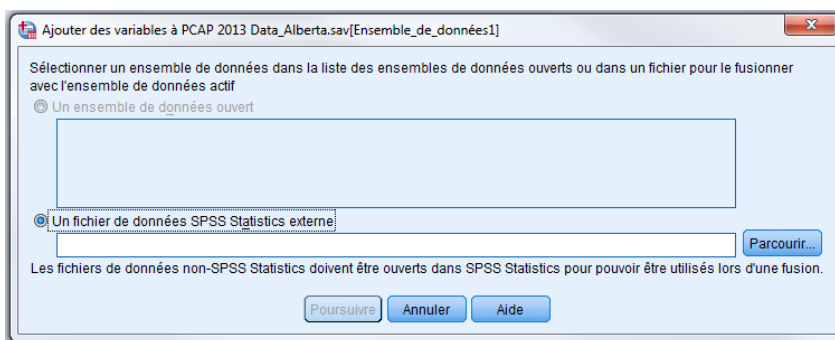
Les mêmes élèves ont répondu aux mêmes items après l'évaluation du PPCE 2013. Les données ont été collectées par le personnel administratif scolaire local, mais chaque ensemble de données a été sauvegardé dans son propre fichier SPSS, soit `additional_lab6_data.sav`. Les deux ensembles de données présentent la même variable « `STUDENT_ID_ELEVE` ». Téléchargez et ouvrez l'ensemble supplémentaire de données SPSS pour ce laboratoire.

STUDENT_ID_ELEVE	SCIENCE_Post	MATH_Post	READING_Post	SQ13A_Post	SQ13B_Post	SQ13C_Post	SQ
54538	618.64	547.00	495.52	2.00	3.00	1.00	
54539	524.21	439.95	507.71	2.00	3.00	2.00	
54540	455.47	475.91	423.67	1.00	2.00	3.00	
54541	593.75	283.14	602.08	2.00	3.00	1.00	
54542	441.56	412.27	520.09	1.00	3.00	1.00	
54543	545.34	514.16	574.72	1.00	2.00	1.00	
54545	537.53	511.40	464.06	2.00	3.00	1.00	
54546	694.28	437.46	495.15	2.00	2.00	3.00	

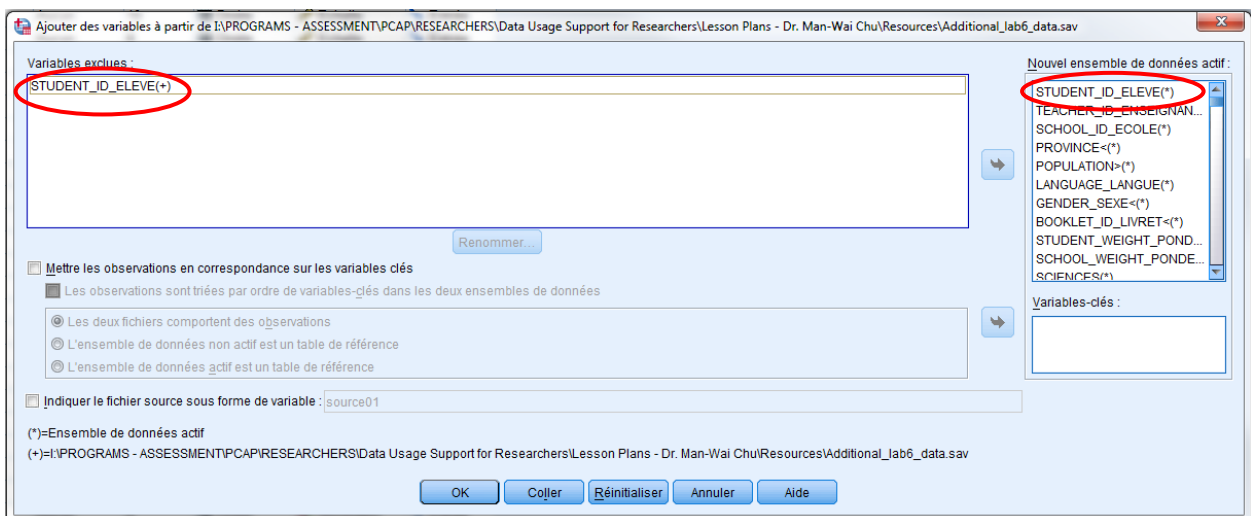
Ouvrez votre ensemble initial de données SPSS du PPCE 2013. Pour fusionner le nouveau fichier (`additional_lab6_data.sav`) au fichier initial (`PCAP2013_Data_Alberta.sav`), veuillez sélectionner vos données PPCE 2013 initiales et cliquer sur **Données → Fusionner des fichiers → Ajouter des variables**.



Maintenant, cherchez votre nouveau fichier à l'aide du bouton « Parcourir ».



Toutes les variables des deux ensembles de données sont énumérées sous « Nouvel ensemble de données actif ». Pour fusionner les deux ensembles de données, vous devrez mettre en correspondance les numéros d'identification des élèves. Cliquez sur la variable STUDENT_ID_ELEVE dans la boîte « Variables exclues ». Puis cliquez sur « Renommer » et assurez-vous que le « Nouveau nom » (sensible à la casse) correspond **EXACTEMENT** au nom de la variable dans vos deux fichiers SPSS.



Cliquez maintenant sur « OK ». Vos ensembles de données devraient maintenant être fusionnés. Les nouvelles données se trouvent à la fin de votre fichier SPSS initial (vous devrez faire défiler le fichier jusqu'à la fin pour voir les données nouvellement fusionnées). Veuillez vérifier quelques cases pour vous assurer que SPSS a correctement fusionné vos deux fichiers.

2. Interprétation de l'ensemble de données

Vous souhaitez maintenant évaluer la campagne contre l'intimidation. La première étape consiste à fournir quelques analyses descriptives de vos données.

Question :

1. Présentez vos résultats dans un tableau qui indique a) la moyenne, b) l'écart type et c) l'erreur standard moyenne pour les données préintervention et postintervention de chaque item sur l'intimidation. Pour la production de votre tableau, veuillez tenir compte des règles de la 6^e édition de l'APA.

3. Test T pour deux échantillons appariés – Campagne contre l'intimidation

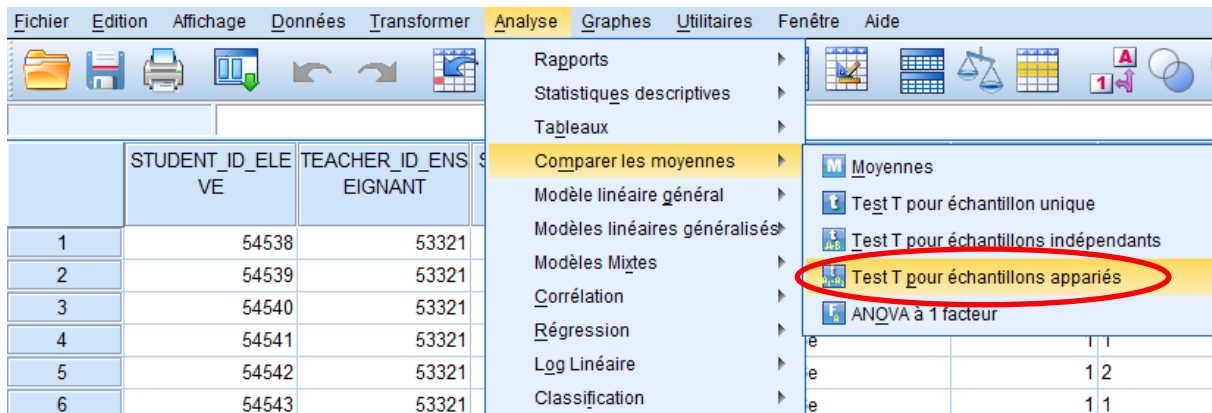
Puisque vous cherchez à évaluer la campagne contre l'intimidation, la question de recherche à laquelle vous voulez répondre est : « La campagne contre l'intimidation a-t-elle diminué l'intimidation dans les écoles? » Puisque nous avons deux ensembles de données pour chaque élève et pour chaque item sur l'intimidation, nous utiliserons un test T pour échantillons appariés. Toutefois, avant d'amorcer nos analyses inférentielles, nous devons nous assurer que nos données respectent quelques hypothèses de base. Dans une analyse d'échantillons appariés, l'hypothèse d'indépendance ne tient pas puisque les mêmes élèves ont répondu au questionnaire préintervention et au questionnaire postintervention. Il nous reste toutefois à vérifier les deux autres hypothèses.

Question :

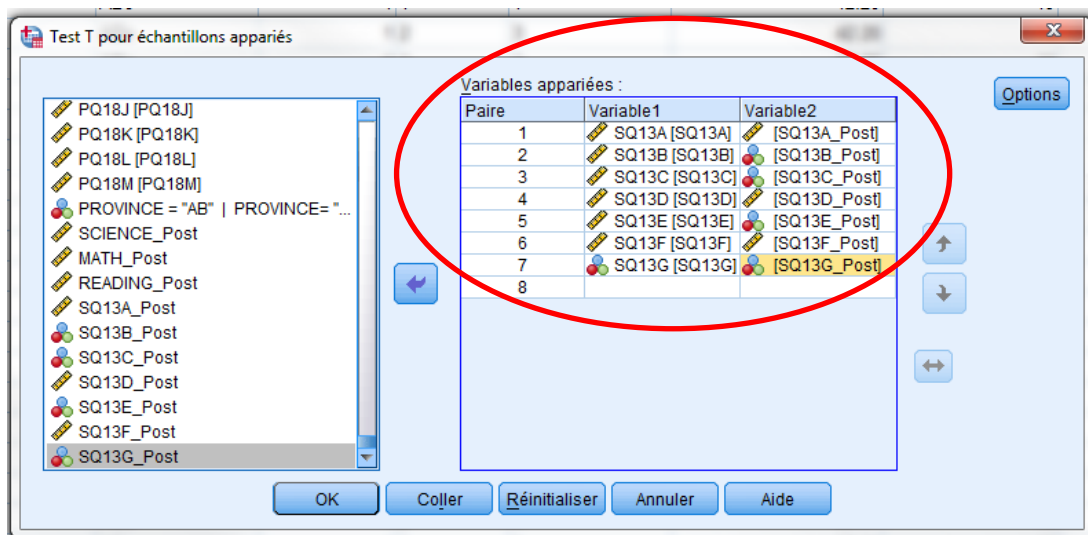
2. Veuillez vérifier les hypothèses suivantes.
 - a) **Normalité** – Votre échantillon viole-t-il l'hypothèse de normalité? (Si vous avez oublié comment vérifier la normalité, veuillez consulter les questions 3, 4 et 5 du laboratoire 5.)
 - b) **Homogénéité de la variance** – Vous n'avez pas à vérifier cette hypothèse dans ce laboratoire. Pour ce faire, il vous faudrait un « codage factice », que je couvrirai dans notre prochain cours. Vous devrez toutefois vérifier cette hypothèse si vous effectuez un test T pour échantillons appariés lors de futurs projets.

Afin de lancer un test T pour échantillons appariés pour cet échantillon d'élèves, cliquez sur

Analyse → Comparer les moyennes → Test T pour échantillons appariés → placez *chacun des items préintervention et postintervention* dans la boîte « Variables appariées » sous « Variable 1 » et « Variable 2 » → OK. Il vous faudra entrer les items un à la fois.



Si vous trouvez un raccourci pour tous les entrer ensemble, veuillez me faire profiter de votre savoir! J'essaie constamment d'acquérir de nouvelles compétences avec SPSS.



Questions :

3. Veuillez formuler l'hypothèse nulle et l'hypothèse alternative pour ce problème de recherche.
4. En consultant les résultats dans la boîte « Test T pour échantillons appariés », veuillez indiquer : la différence moyenne, l'écart type, l'erreur standard moyenne, la valeur t, les degrés de liberté et la valeur p observée. (Veuillez noter que la « différence moyenne » est la différence entre les deux moyennes que vous avez indiquées dans votre tableau de statistiques descriptives. Dans le rapport SPSS, elle est simplement appelée « moyenne », mais dans vos documents, vous devrez indiquer qu'il s'agit de la « différence moyenne » pour montrer la différence entre la moyenne présentée dans

vos analyses descriptives et la différence moyenne présentée dans vos analyses inférentielles du test T.)

5. À la lumière des valeurs p observées, est-ce que chaque différence moyenne se trouve dans la région critique (INDICE : il s'agit ici d'une question de recherche directionnelle)? Si oui, pour quels items? (Vous devrez peut-être utiliser Google pour voir comment calculer la valeur « Sig. » pour un test unilatéral dans SPSS.) **Veillez noter** que la valeur « Sig. » du rapport SPSS et les valeurs p des régions critiques ne correspondent pas aux mêmes concepts.

Si vous procédez à plusieurs analyses inférentielles, le risque est élevé de gonfler l'erreur de type 1. Nous nous protégeons donc contre ce gonflement en exécutant une correction de Bonferroni.

Question :

6. Veuillez expliquer *pourquoi* l'exécution de multiples analyses inférentielles pourrait accroître le risque d'erreur de type 1. (Vous devrez peut-être utiliser Google pour cette question; veuillez limiter votre réponse à *quelques* phrases.)

La correction de Bonferroni compense ce risque accru d'erreur de type 1 en réduisant le niveau alpha à partir duquel nous vérifions la région critique. Le nouveau niveau alpha selon la correction de Bonferroni est :

$$\text{niveau alpha initial} \div \text{nombre de tests inférentiels} = \text{nouveau niveau alpha}$$

Dans notre cas, le nouveau niveau alpha, après correction de Bonferroni, sera ???. Vous utiliserez cette réponse dans la prochaine série de questions.

Questions :

7. Veuillez analyser de nouveau vos résultats et dire quels items présentent une différence statistiquement significative. Veuillez également indiquer ces différences statistiquement significatives dans votre tableau de résultats (produit à la question 3) en insérant un « * » à côté des valeurs « Sig. » qui sont statistiquement significatives.
8. Vous devrez présenter vos résultats pour chaque item avec votre tableau de résultats inférentiels. Veuillez répondre à la question de recherche pour chaque item sur l'intimidation. Répondez à la question en suivant le libellé « La campagne contre l'intimidation a diminué de façon statistiquement [significative/non significative] le taux déclaré par les élèves de [insérez le libellé de l'item], t ([degré de liberté]) =

[valeur t], p [</>]nouveau niveau alpha. » Veuillez suivre les règles de l'APA dans le libellé de cette phrase (par exemple, assurez-vous d'avoir mis en italique les bons symboles t et p). Il importe d'utiliser l'expression « différence statistiquement [significative/non significative] », car votre décision quant à savoir si le résultat est significatif ou non significatif ne devrait pas se baser uniquement sur votre test statistique. Il est donc important de dire à votre lectorat que cette différence « [significative/non significative] » l'est sur le plan statistique.

Une des difficultés dans l'explication de ces résultats est de veiller à inclure tous les éléments de preuve techniques qui appuient vos affirmations et de formuler des phrases grammaticalement cohérentes. Habituellement, dans une publication, seuls les items statistiquement significatifs sont présentés dans une phrase, de sorte que vous puissiez discuter plus à fond de chaque résultat.

Questions :

9. À la lumière des informations obtenues, veuillez calculer le d de Cohen et le pourcentage de variance expliquée. Théoriquement, qu'indiquent ces valeurs? Que révèle l'ampleur de ces valeurs aux gens qui lisent votre étude?

10. Maintenant, veuillez répondre dans un *bref* paragraphe à votre question de recherche initiale : « La campagne contre l'intimidation a-t-elle diminué l'intimidation dans les écoles? » Il vous faudra peut-être utiliser plusieurs énoncés quantitatifs et travailler à partir de chaque item pour ne pas généraliser de façon excessive les résultats de votre analyse. Ce paragraphe étant souvent utilisé dans votre discussion, il ne devrait présenter **aucun résultat technique**.

4. Test T pour deux échantillons appariés – Projet de refonte des programmes d'études

L'Alberta a récemment lancé un projet de refonte des programmes d'études des sciences, des mathématiques et de la lecture pour les actualiser. Vous voulez savoir si cette refonte a eu une incidence sur le rendement des élèves dans ces trois matières, tel qu'il est mesuré par les tests de rendement normalisés. Les élèves ont passé l'évaluation du PPCE 2013 avant la refonte, puis ils ont passé un test similaire à la fin de l'année. Leurs résultats ont été fusionnés à la section 1 de ce laboratoire. Veuillez répondre à la question de recherche suivante : « La refonte des programmes d'études a-t-elle eu une incidence sur le rendement des élèves au cours de la première année de son déploiement? »

Questions :

11. Veuillez procéder aux analyses pour cette question de recherche. Dans votre réponse à cette question, veuillez inclure : a) un tableau descriptif des données des analyses, b) les résultats de la vérification de l'hypothèse de normalité, c) le tableau de statistiques inférentielles des résultats du test T (indiquez les résultats statistiquement significatifs dans votre tableau), d) un *bref* paragraphe expliquant vos résultats (utilisez le format approprié de l'APA pour communiquer vos conclusions) et e) les valeurs d de Cohen pour les variables applicables (insérez ces valeurs dans le paragraphe d précédent).

12. Veuillez répondre à la question de recherche initiale dans un *bref* paragraphe, à utiliser éventuellement dans la section réservée à votre discussion.

Vous avez maintenant terminé le laboratoire.