

Protocole de TP 03 :

Test de comparaison des moyennes

Notre objectif dans cette TP c'est

- 1) Savoir comment faire une comparaison entre deux échantillons indépendants.
- 2) Aussi de savoir comment faire une comparaison des moyennes, mais dans une seule échantillon, (Avant le traitement et après le traitement).
- 3) Ainsi la comparaison entre la moyenne d'une population (la moyenne théorique) et la moyenne d'échantillon (la moyenne observé).

En utilisant pour cette raison les techniques du test de comparaison des moyennes, par utilisation du logiciel SPSS.

I) Comparaison entre deux échantillons indépendants :

On utilise ce type du test lorsqu'on a deux échantillons, et on veut faire étude sur l'homogénéité entre les deux échantillons, et on faire la comparaison entre les moyennes des deux échantillons.

Pour cela on doit prendre un exemple d'explication :

Exemple 1

Dans des études d'anesthésie, voulant comparer l'effet de deux somnifères, on a noté les durées de sommeil qui ont suivi les injections d'une dose bien définie. Les durées étant exprimées en minutes:

Somnifère 01	170	175	187	180	190	165	175	174	173	181		
Somnifère 02	155	160	164	150	160	159	154	156	160	167	153	158

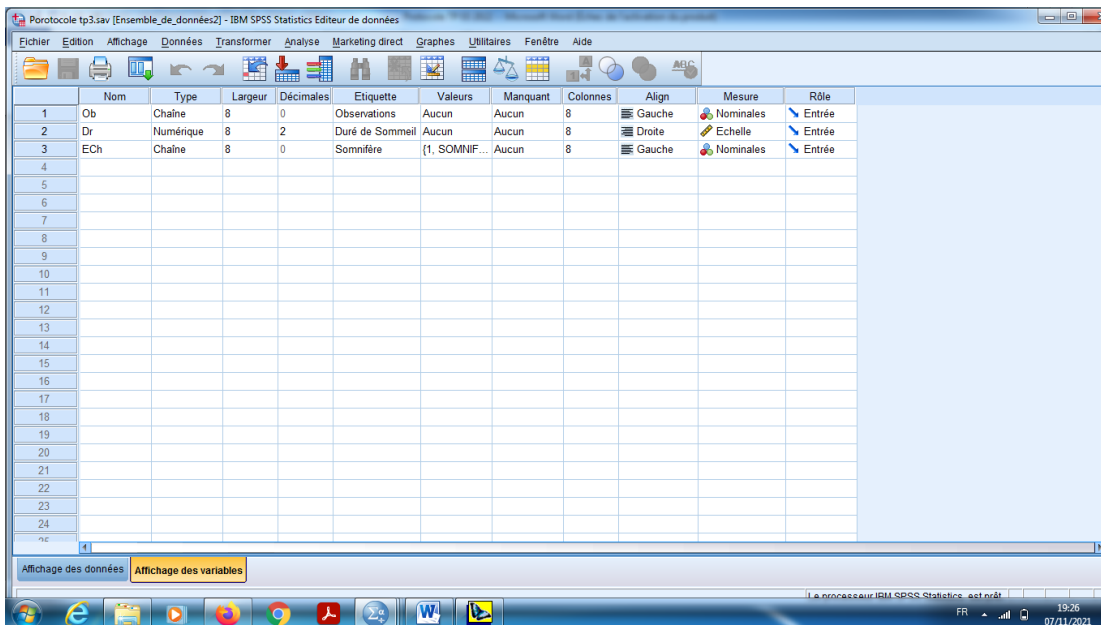
- 1) Déterminer l'objectif pour cette expérience.
- 2) Déterminer la variable qualitative qui exprime les deux échantillons, et la variable quantitative a mesuré.
- 3) Déterminer l'hypothèse nulle et alternative.
- 4) Tracer le tableau de de la statistique descriptive pour l'effectif, moyennes ainsi que l'écart-types.
- 5) Avec un risque de signification de 5%, que peut-on dire pour cette comparaison?

Solution

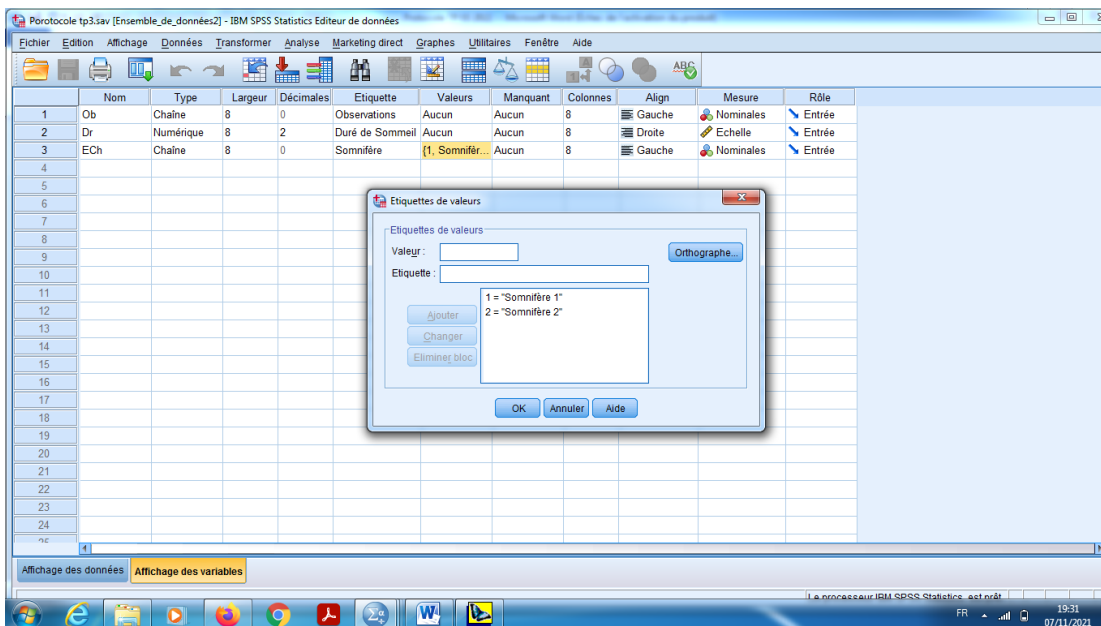
Tout d'abord il faut entrer ces données dans SPSS, on an a déjà vu les techniques dans TP 1.

Nous suivant les étapes suivantes :

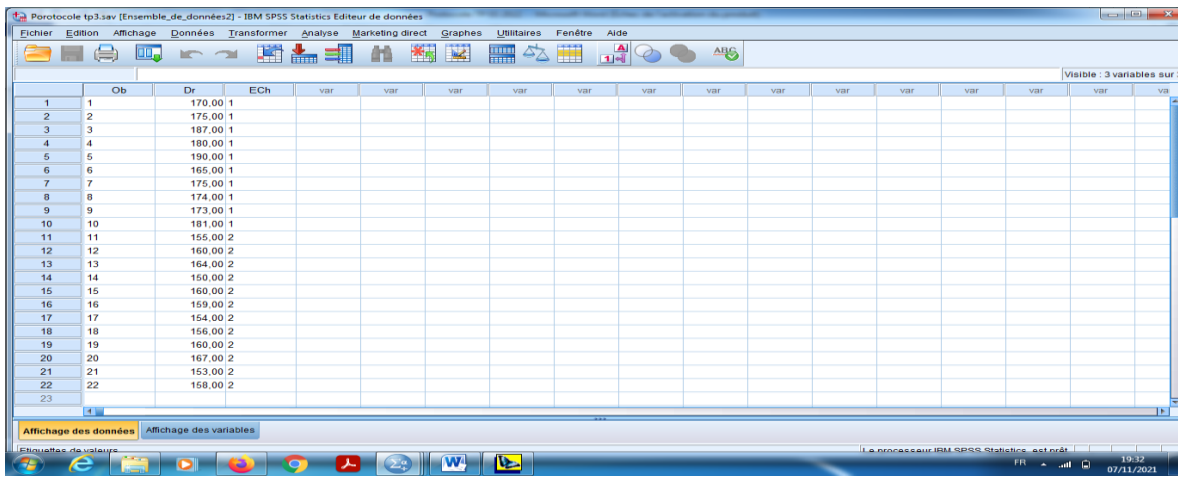
Il faut définir dans la barre en bas « Affichage des variables » : les variables qualitative et quantitative suivantes : observations, la durée de sommeil, et les somnifères.



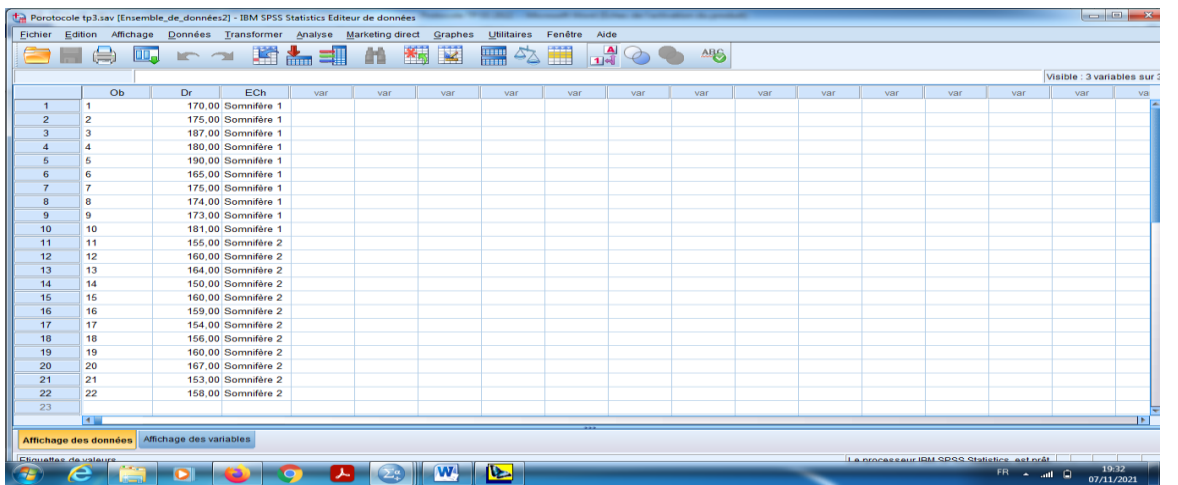
On choisit les modalités pour la variable qualitative qui représente les deux échantillons de somnifères dans « valeurs », on peut prendre comme un exemple pour la variable somnifère (1= pour somnifère 1 et 2= pour somnifère 2).



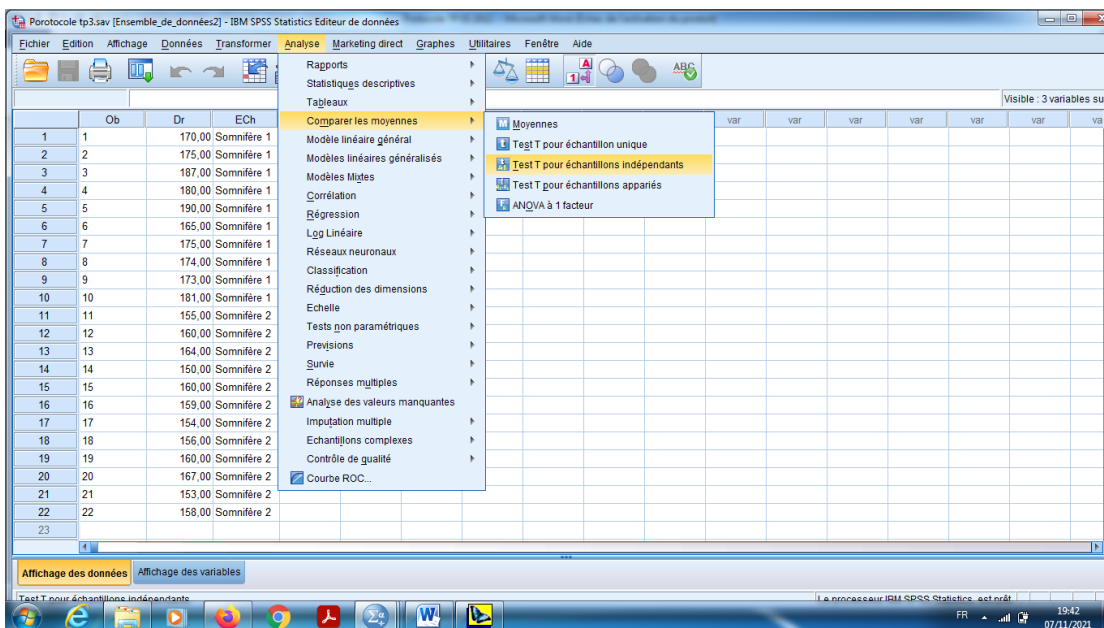
On introduit les données dans la barre « Affichage des données » par SPSS.



En cliquant sur le bouton « Etiquettes des valeurs » pour visualiser le codage des échantillons.

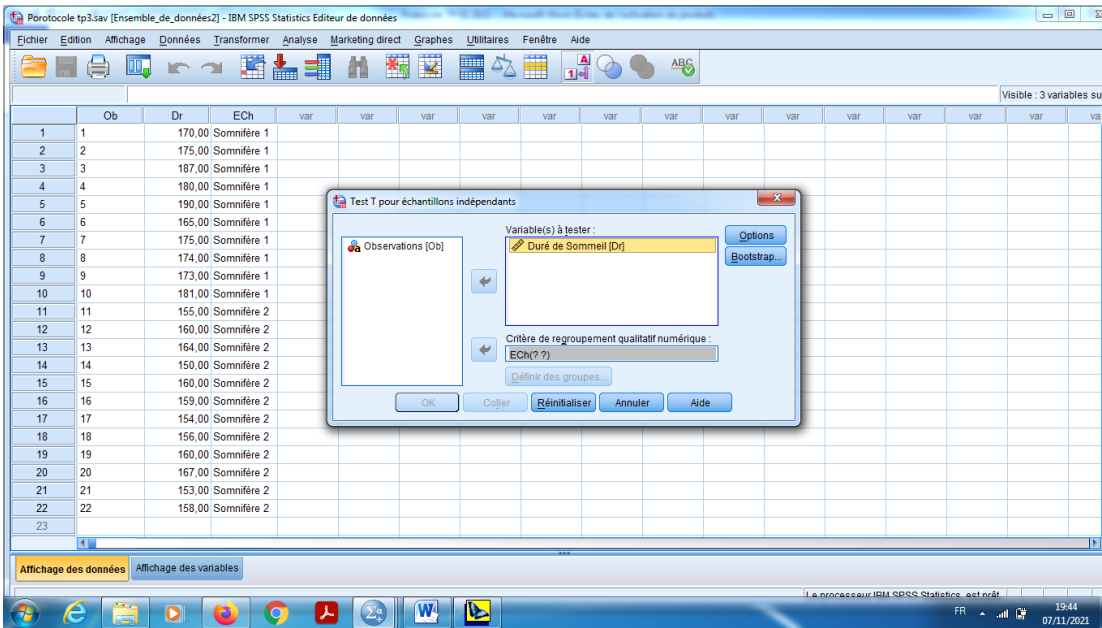


En suit, pour obtenir la comparaison entre deux échantillons de somnifères, en cliquant sur le bouton « Analyse » dans la barre des outils, et choisir « Comparer les moyennes », et puis « Test T pour échantillons indépendants ».

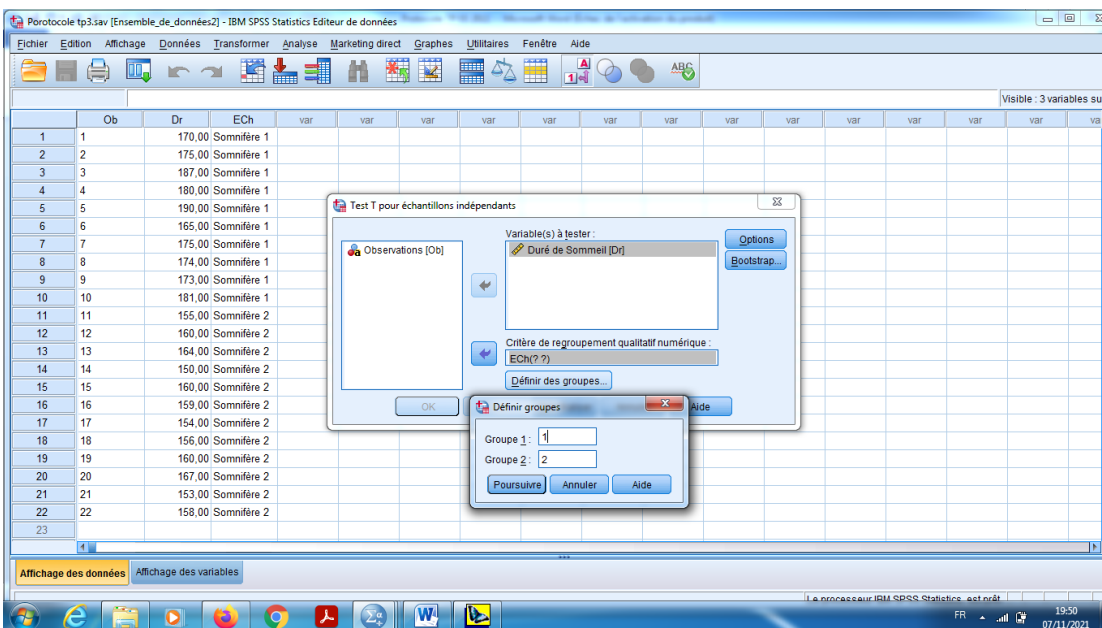


a- On choisit dans « la variable à tester » la variable Durée de Sommeil.

- b- On choisit la variable qualitative (Echantillons=somnifères) dans la barre « Critère des regroupements qualitatif numérique ».
- c- On choisit aussi « Définir les groupes », et on prend pour le somnifère 1 (groupe 1) le code 1, et on prend pour le somnifère 2 (groupe 2) le code 2.



- d- Dans le Bouton « Options » on garde le pourcentage pour l'intervalle de confiance le seuil de signification 95%, en cliquant sur « poursuivre ».



En fin en cliquant sur « Ok ».

Interprétation des résultats

1) Notre objectif est de comparer les deux types des somnifères, pour cela on utilise le teste de comparaison entre les moyennes (Test d'échantillons indépendants).

2) Proposition d'hypothèses :

Hypothèse nulle, H0 : « Il n'y a pas une différence significative entre les deux somnifères », c.-à-d les deux moyennes sont significativement coïncident. $\bar{x}_1 = \bar{x}_2$.

Hypothèse alternative, H1 : « Il y a une différence significative entre les deux somnifères », c.-à-d. les deux moyennes sont significativement différentes. $\bar{x}_1 \neq \bar{x}_2$.

Remarque:

On parle aussi d'effet du facteur « les somnifères (variable qualitative) » sur la variable « quantitative la durée de sommeil ».

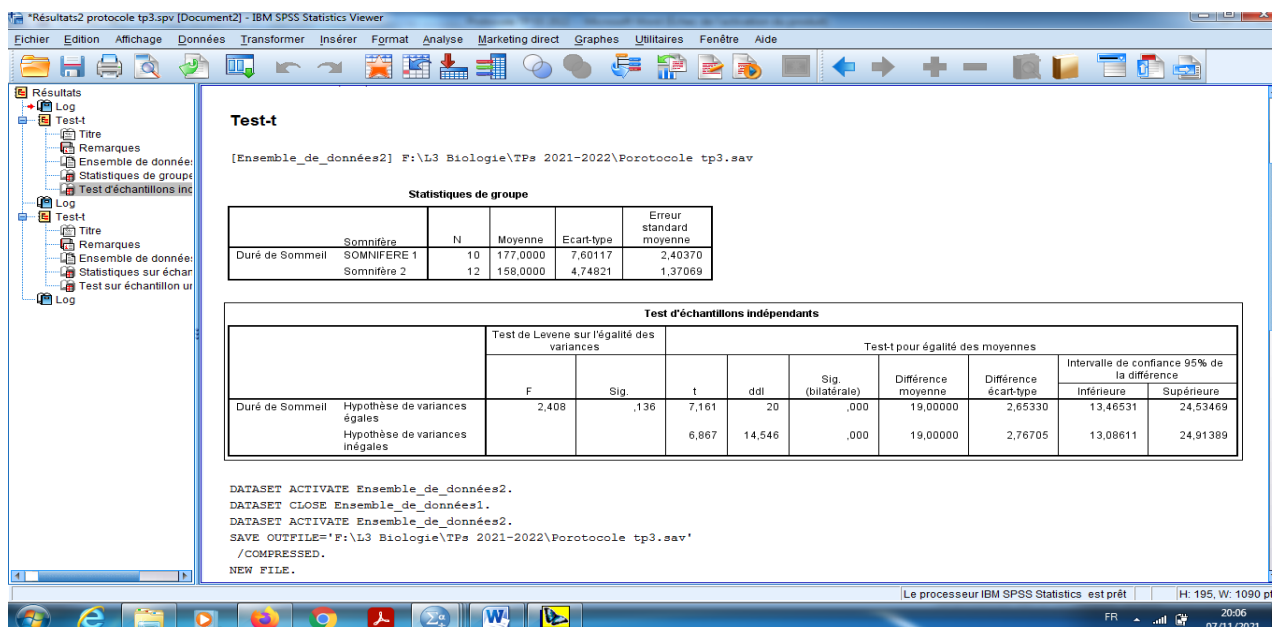
Pour la décision, on utilise souvent la règle suivante

Si Signification (bilatérale) inférieure à 0,05, alors on rejette H0 (Il y a une différence significative).

Si Signification (bilatérale) supérieure à 0,05, alors on accepte H0 (Pas différence significative).

Dans la table 1 «_Statistique de groupe» on remarque qu'il y a 10 observations dans la somnifère 1 et 12 pour la somnifère 2, et pour les moyennes 177 pour le somnifère 1 et 158 pour la somnifère 2, de plus pour l'écart-type on remarque 7,60117 et 4,74821 respectivement pour la somnifère 1 et 2.

Statistiques de groupe					
	Somnifère	N	Moyenne	Ecart-type	Erreur standard moyenne
Duré de Sommeil	SOMNIFERE 1	10	177,0000	7,60117	2,40370
	Somnifère 2	12	158,0000	4,74821	1,37069



On s'intéresse seulement sur (Test T pour l'égalité des moyennes).

Statistique t=7,161.

DDL=n1+n2-2=20.

Signification (bilatérale)=0,000.

Pour cela on peut voir dans les tableaux obtenues dans les résultats de SPSS.

Signification (bilatérale)=0,000<0,05. Donc on rejette H_0 , alors il y a une différence significative entre les somnifères, les deux échantillons ne sont pas homogènes.

D'autre part on peut conclure aussi que la moyenne se trouve dans un intervalle qui s'appelle « Intervalle de confiance de la différence », par la forme suivante

$m \in [13,46; 24,53]$, alors il est clair que les deux bornes **sont positives**, donc on en conclure que (Somnifère 1-Somnifère 2 est positive), c'est – à –dire que avec une taux de 95% la Somnifère 1 est plus efficace que Somnifère 2.

II) Comparaison entre la moyenne observé et la moyenne théorique

On utilise ce type de test dans le cas où on a un seul échantillon, et une seule valeur pour la moyenne théorique qui représente bien sur la population, avec une autre valeur observé qui représente l'échantillon.

Ou bien on pose la question suivante :

Est-ce que l'échantillon représente bien la population ?

Pour cela on doit prendre un exemple d'explication :

Exemple 2

On a prélevé un échantillon de 10 paquets de tabac dans la production d'une machine à paqueter.

La mesure du poids de ces paquets a donné des résultats suivants.

Poids	17	17,5	18,7	18,0	19,0	16,5	17,5	17,4	17,3	18,1		
-------	----	------	------	------	------	------	------	------	------	------	--	--

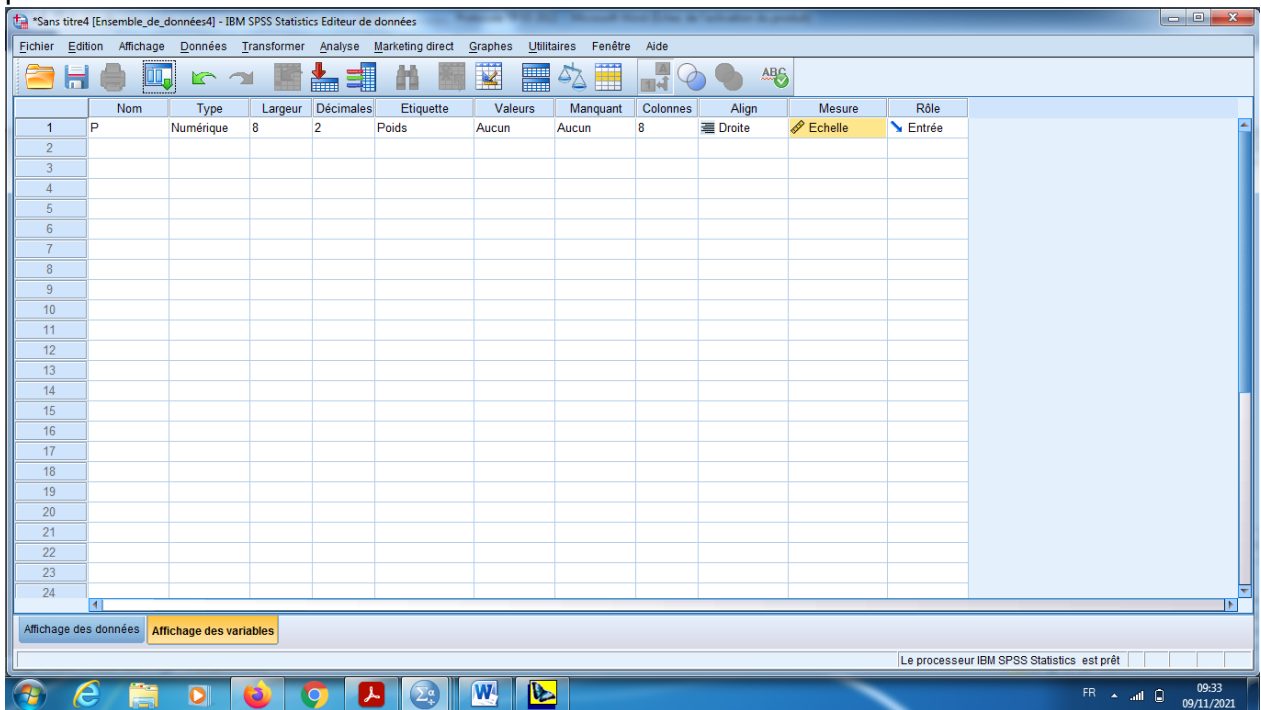
- 1) Déterminer l'objectif pour cette expérience.
- 2) Déterminer la variable qualitative qui exprime les deux échantillons, et la variable quantitative a mesuré.
- 3) Déterminer l'hypothèse nulle et alternative.
- 4) Tracer le tableau de de la statistique descriptive pour l'effectif, moyennes ainsi que l'écart-types.
- 5) On demande si la moyenne observée est compatible avec l'hypothèse que la machine fabrique « en moyenne » des paquets de 18,0 g au risque de signification de 5%.

Solution :

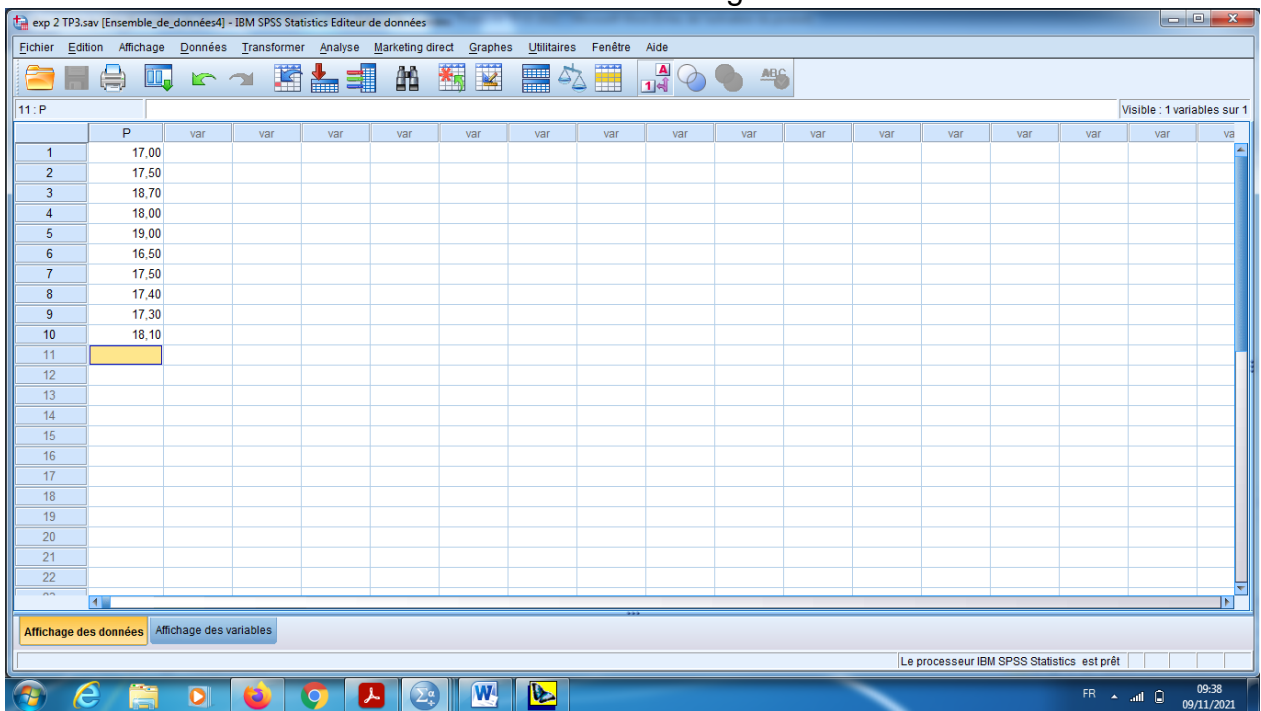
Pour cela on utilise les étapes suivantes :

- a) Il faut définir les variables dans la barre « Affichage des variables ».

- b) On choisit « P » pour le nom de la variable, et « numérique » pour le type, et « échelle » pour sa mesure.

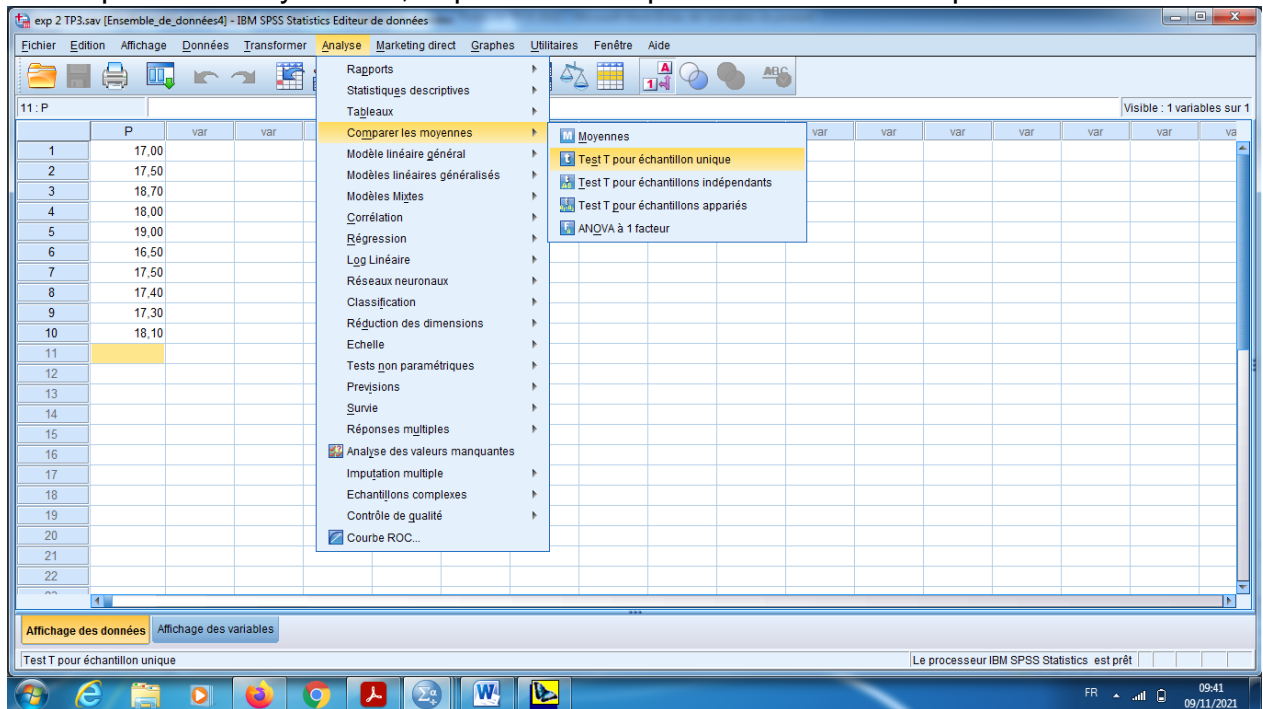


- c) On introduit les données dans la barre de « Affichage des Données ».

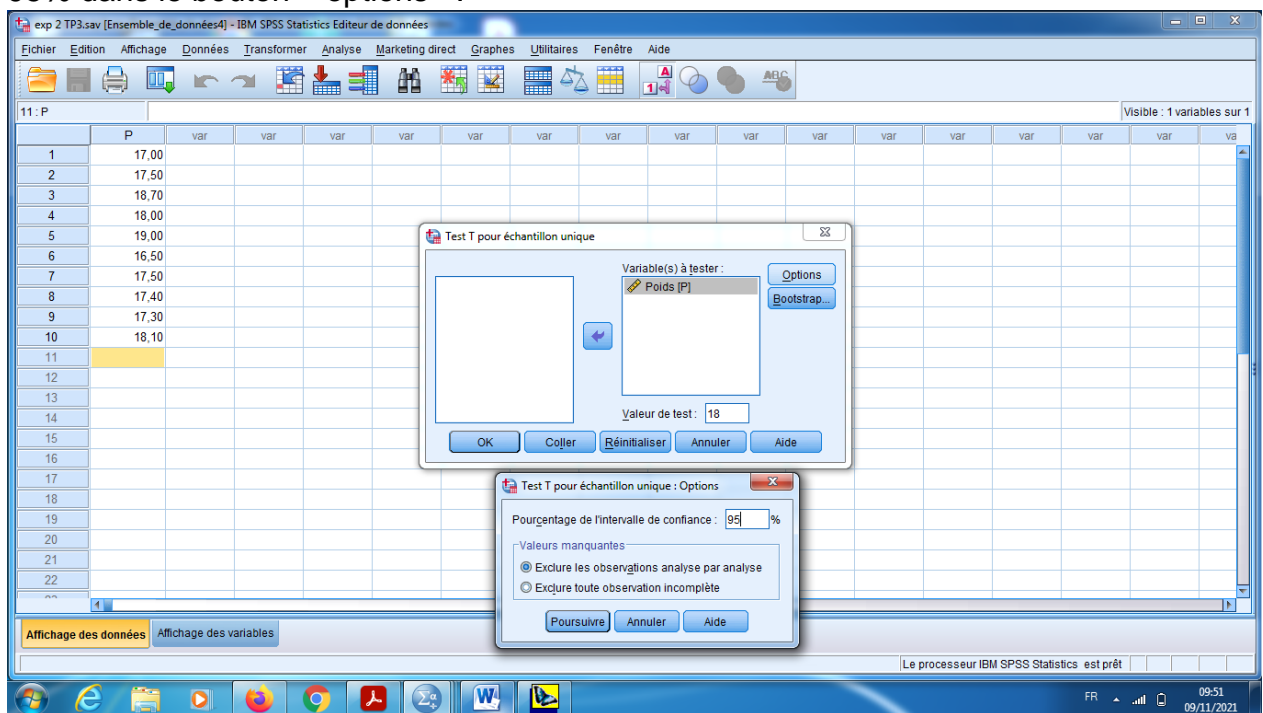


- d) En suit, pour obtenir la comparaison entre la moyenne observé et la moyenne théorique sur un seul échantillon, en cliquant sur le bouton « Analyse » dans le menu, et choisir

« Comparer les moyennes », et puis « Test T pour échantillons unique ».



- e) Sur la boîte de dialogue, on pose la variable « Poids » dans la barre de la « variable à tester », et la valeur de test de la moyenne théorique =18g (c'est la valeur qui se propose par industrie), et on garde la valeur de seuil de signification pour l'intervalle de confiance 95% dans le bouton « options ».



- f) En cliquant sur « poursuivre » et puis sur « OK », pour obtenir les résultats des tests T.

Résultats des tests :

Hypothèses

Hypothèse nulle : H0 (La moyenne observé est confirmé avec la moyenne théorique)

$$\bar{x}_1 = m = 18.$$

Hypothèse nulle : H1 (La moyenne observée est différente que la moyenne théorique).

$$\bar{x} \neq m = 18.$$

Remarque : Il bien noté que la différence entre les moyennes est significative avec un taux de risque c'est 5% (la confiance pour la décision c'est 95%).

Pour la décision, on utilise souvent la règle suivante

Si Signification (bilatérale) inférieure à 0,05, alors on rejette H0 (Il y a une différence significative).

Si Signification (bilatérale) supérieure à 0,05, alors on accepte H0 (La différence n'est pas significative).

Interprétation des résultats :

- a) Dans la table des statistiques descriptives, on remarque ici que la moyenne des 10 paquets c'est « 17,7 », mais l'écart-type c'est « 0,760 ».

Statistiques sur échantillon unique

	N	Moyenn e	Ecart- type	Erreur standard moyenne
Poid s	10	17,7000	0,76012	0,24037

- b) Sur le test T des comparaisons entre les deux moyenne observé er théorique, on remarque : la valeur du test « T=-1,248 », avec ddl=k-1=10-1=9.

Mais ce qui m'intéresse ici c'est la valeur « Sig(bilatéral) =0,243 », alors cette valeur est supérieure à 0,05.

Alors d'après la règle de décision on accepte de H0, c'est-à-dire les deux moyennes sont significativement coïncidentes, alors les résultats obtenus par la fabrication des machines sont sur la norme.

Test sur échantillon unique

	Valeur du test = 18					
	T	Ddl	Sig. (bilatérale)	Différence moyenne	Intervalle de confiance 95% de la différence	
					Inférieure	Supérieure
Poid s	-1,248	9	,243	-,30000	-0,8438	,2438

- c) Finalement la valeur de la différence entre les moyennes ($m - \bar{x}$) est se trouve dans les deux bornes (-0,843 et 0,243),

$$m \in [\bar{x} - 0,843; \bar{x} + 0,243].$$

III) Comparaison entre la moyenne observé apparié

On utilise ce type de test dans le cas où on a un seul échantillon, mais on a deux valeurs (deux observations) effectuées pour le même individu (avant et après le traitement).

Pour cela on doit prendre un exemple d'explication :

Exemple 3

Un groupe de 10 sujets entreprend sur une période de six mois un programme d'enrichissement cognitif destiné à améliorer leurs processus de traitement de l'information. Pour évaluer l'effet du programme, on fait passer aux sujets deux tests de niveaux comparables l'un avant, l'autre après la période d'apprentissage, dont voici les scores des notes sur 10 points:

Scores AVANT	5	2	8	9	5	3	2	5	9	8
Scores APRÈS	6,5	2	9	10	6	1	3	8	7	9

A l'aide d'un test de comparaison, décider au niveau 5% si les sujets ont en moyenne amélioré leur processus de traitement de l'information après le programme d'enrichissement cognitif.

Remarque importante:

On remarque tout d'abord qu'on a un seul échantillon, avec deux mesures pour chaque individu.

Solution :

1) Objectif : Notre objectif dans cette expérience est de savoir l'effet du programme d'amélioration d'apprentissage.

2) Proposition des hypothèses :

Hypothèse nulle : H_0 : (le programme d'enrichissement cognitif ne change rien aux capacités de traitement de l'information.) ($\bar{x}_1 = \bar{x}_2$)

Hypothèse alternative H_1 : (le programme d'enrichissement cognitif améliore les capacités de traitement de l'information.) ($\bar{x}_1 < \bar{x}_2$).

3) Pour la décision, on utilise souvent la règle suivante

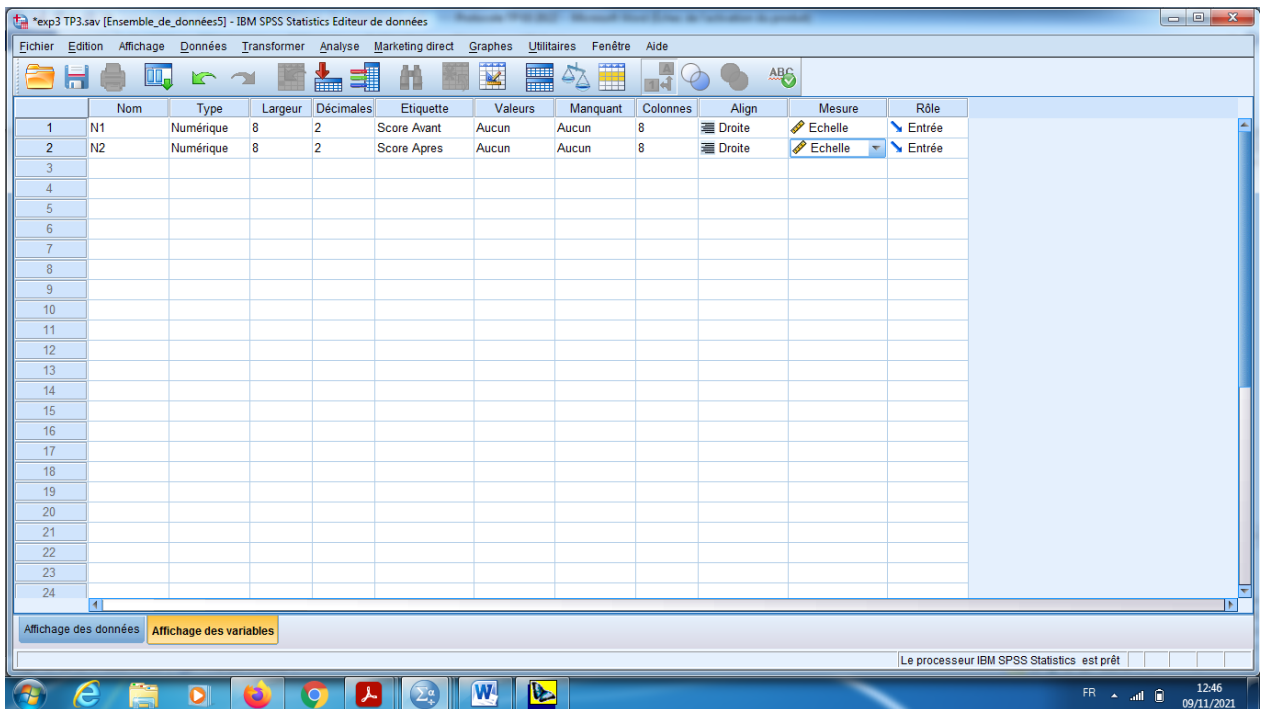
Si Signification (bilatérale) inférieure à 0,05, alors on rejette H_0 (Il y a une différence significative).

Si Signification (bilatérale) supérieure à 0,05, alors on accepte H_0 (La différence n'est pas significative).

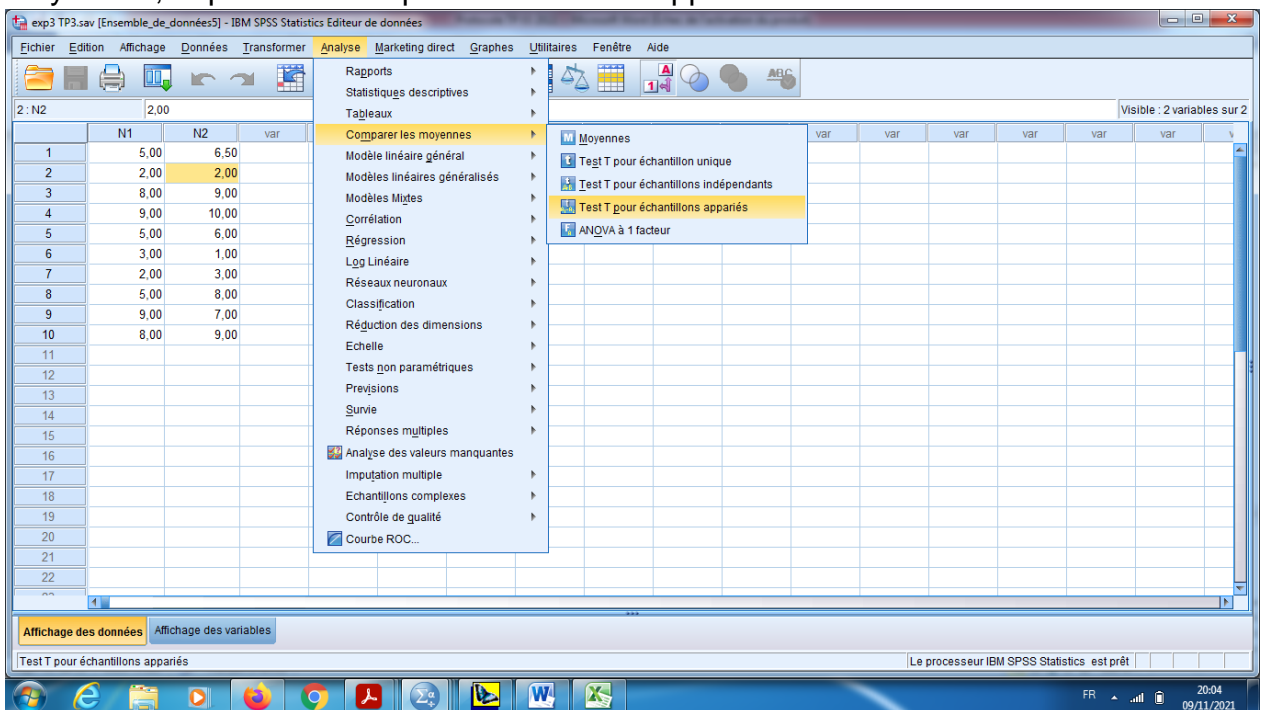
Pour l'introduction des données, en suivant les étapes connues :

a) Il faut définir les variables dans la barre « Affichage des variables ».

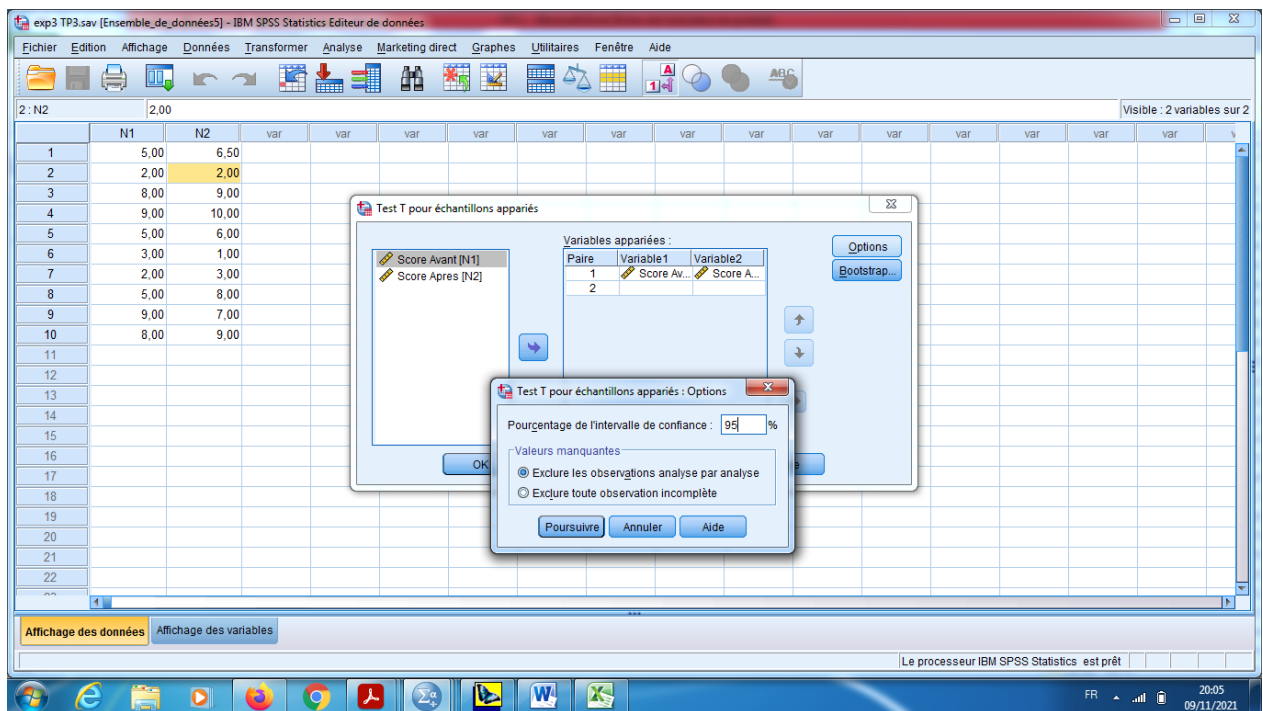
b) On choisit « N1 » pour le nom de la variable score avant, et « numérique » pour le type, et « échelle » pour sa mesure, « N2 » pour le nom de la variable score après, et « numérique » pour le type, et « échelle » pour la mesure.



- c) On introduit les données dans la barre de « Affichage des Données ».
- d) En suit, pour obtenir la comparaison entre les deux moyennes (avant et après) sur un seul échantillon, en cliquant sur le bouton « Analyse » dans le menu, et choisir « Comparer les moyenne », et puis « Test T pour échantillons appariés ».

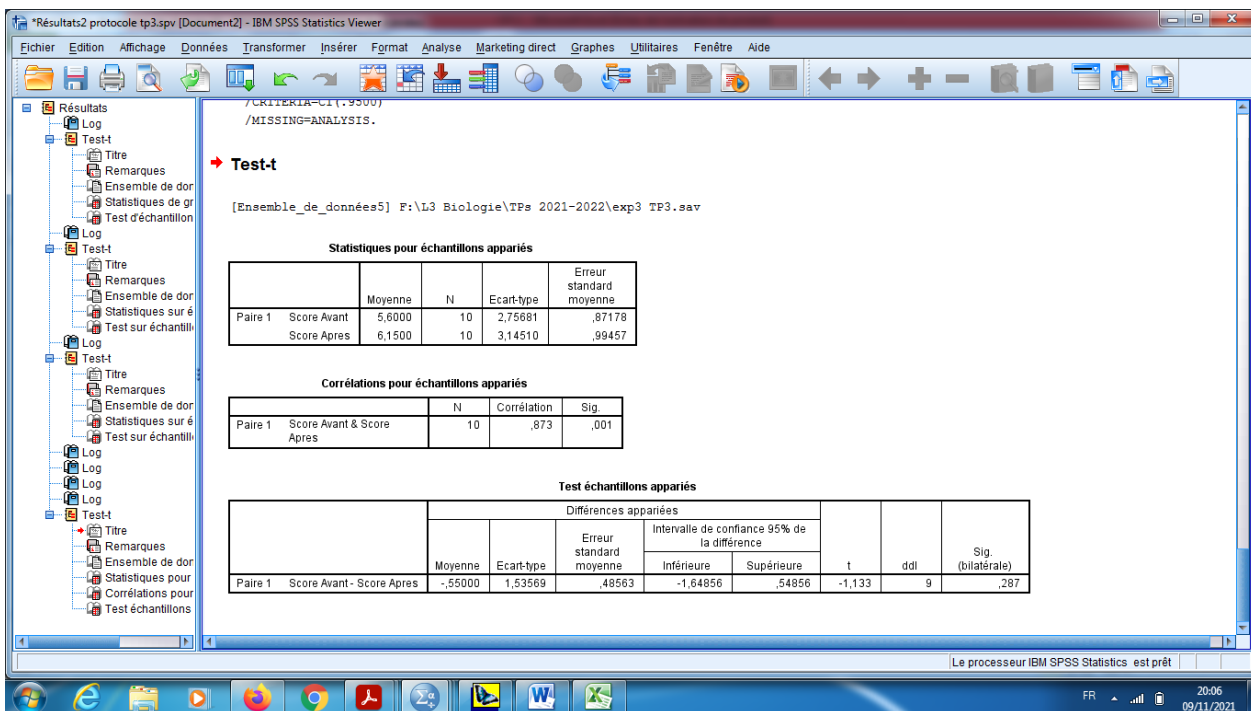


- e) Sur la boîte de dialogue « la variable appariés » on choisit sur la variable 1 (Score avant), et sur la variable 2 (le score Après), et on garde la valeur de seuil de signification pour l'intervalle de confiance 95% dans le bouton « options ».



f) En cliquant sur « poursuivre » et puis sur « OK », pour obtenir les résultats des tests T appariés.

Interprétation des résultats :



1) Pour la statistique descriptive, la moyenne pour la partie avant c'est (5,600), et la moyenne pour le score après c'est (6,150), mais pour l'écart-type avant (2,756) et pour l'écart-type après (3,145).

Statistiques pour échantillons appariés

	Moyenne	N	Ecart-type	Erreur standard moyenne	
Paire 1	Score Avant	5,6000	10	2,75681	,87178
	Score Apres	6,1500	10	3,14510	,99457

2) Dans la table du Test échantillons appariés, on remarque que la différence entre les deux moyennes c'est (-0,550) et que la statistique $T = -1,133$; de plus $DDL = k - 1 = 10 - 1 = 9$.

3) Conclusion

- a) Si $\text{sig (bilatérale)} = 0,287 > 0,05$. Alors on accepte de H_0 . Alors les deux moyennes sont égales, c'est-à-dire que le test d'amélioration ne change rien sur l'apprentissage des sujets. (la différence n'est pas significative).
- b) Comme on a l'acceptation de l'hypothèse nulle, alors on ne peut pas parler sur l'efficacité, c'est pour ça, on trouve la valeur 00 sur l'intervalle de confiance de la différence, « l'intervalle est symétrique ».