

Protocole de TP 01 :

Exécution du test ANOVA

Pour validé test de comparaison entre deux échantillons (test de Student), ou bien test d'ANOVA pour plus que deux échantillons indépendants, il faut que les conditions suivantes soient valides :

- a) Une seule variable quantitative mesurée X, et une variable qualitative A avec deux modalités (pour test de Student) ou bien une variable qualitative avec trois modalités et de plus (pour test d'ANOVA).
- b) La distribution soit gaussienne (suit la loi Normale).
- c) L'échantillon est homogène pour la variance pour la variable quantitative, on ne peut pas trouver des valeurs plus loin que la moyenne, (Il n'y a pas des valeurs extrêmes dans la série statistique).

I) Test ANOVA

On utilise ce type du test lorsqu'on veut faire une comparaison entre plus que trois échantillons.

Pour résoudre, on peut poser les questions suivantes :

Existe-il une influence de la variable qualitative « nominale ou bien ordinale » (qui s'appelle aussi le Facteur A, avec nombre k des modalités) sur la variable quantitative mesuré X ?

D'autre manière, est ce que les k échantillons sont homogènes ou bien non (par rapport la variable mesurable).

Tableau des données

Modalité 1	Modalité 2	Modalité 3	...	Modalité k
$X_{11}$	$X_{12}$	$X_{13}$		$X_{1k}$
$X_{21}$	$X_{22}$	$X_{23}$		$X_{2k}$
..	...	...		...
$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$		$\bar{x}_k$

Pour cela on doit ordonner les réponses par les étapes suivantes

Etape 1 : Proposition d'hypothèses

Hypothèse nulle  $H_0$ : (Il existe homogénéité sur les k échantillons)=(tous les moyennes sont significativement égales), (Il n'a pas une influence du facteur A sur la variable X).

$$(\bar{x}_1 = \bar{x}_2 = \bar{x}_3 =, \dots \bar{x}_k)$$

Hypothèse alternative  $H_1$ : (Il n'y a pas homogénéité sur les k échantillons)=(tous les moyennes sont significativement différentes)=(Il y a une influence du facteur A sur la variable mesuré X).

## Etape 2 : Les calculs (tableau d'ANOVA)

Source des variations	Somme des carrés	Ddl	Moyenne des carrés	F
Inter-groupes	SCE (Inter)	k-1	$CM(Inter) = \frac{SCE(Inter)}{k-1}$	$F = \frac{CM(Inter)}{CM(Intra)}$
Intra-groupes	SCE (Intra)	N-k	$CM(Intra) = \frac{SCE(Intra)}{N-k}$	
Total	SCE (Total)	N-1		

## Etape 3 : Conclusion (la décision)

Pour la décision, on utilise souvent la règle suivante

Si Signification inférieure à  $\alpha\%$ . Alors on rejette  $H_0$ .

Si Signification supérieure à  $\alpha\%$ . Alors on accepte  $H_0$ .

### Exemple:

Pour mettre en évidence l'effet éventuel de l'absorption d'un médicament sur le rythme cardiaque, on forme trois groupes, par tirage au sort parmi les malades traités par ce médicament: Au premier groupe, on n'administre pas le médicament, mais reste un placebo. Au deuxième et troisième groupe on administre le médicament avec différents dosage.

Les données relatives pour cette expérience sont

rythme cardiaque du groupe témoin	170	175	187	180	190	165	175	174	173	181		
rythme cardiaque du groupe 2 traité	155	160	164	150	160	159	154	156	160	167	153	158
rythme cardiaque du groupe 3 traité	140	155	153	167	162	144	148	141	160			

- 1) Déterminer l'objectif pour cette expérience.
- 2) Déterminer la variable qualitative qui exprime les échantillons (Facteur A), et la variable quantitative a mesuré X.
- 3) Déterminer l'hypothèse nulle et alternative pour le test d'ANOVA à un facteur.
- 4) Avec un risque de signification de 4%, que peut-on dire sur l'influence du facteur A sur la variable X?

## Reponse

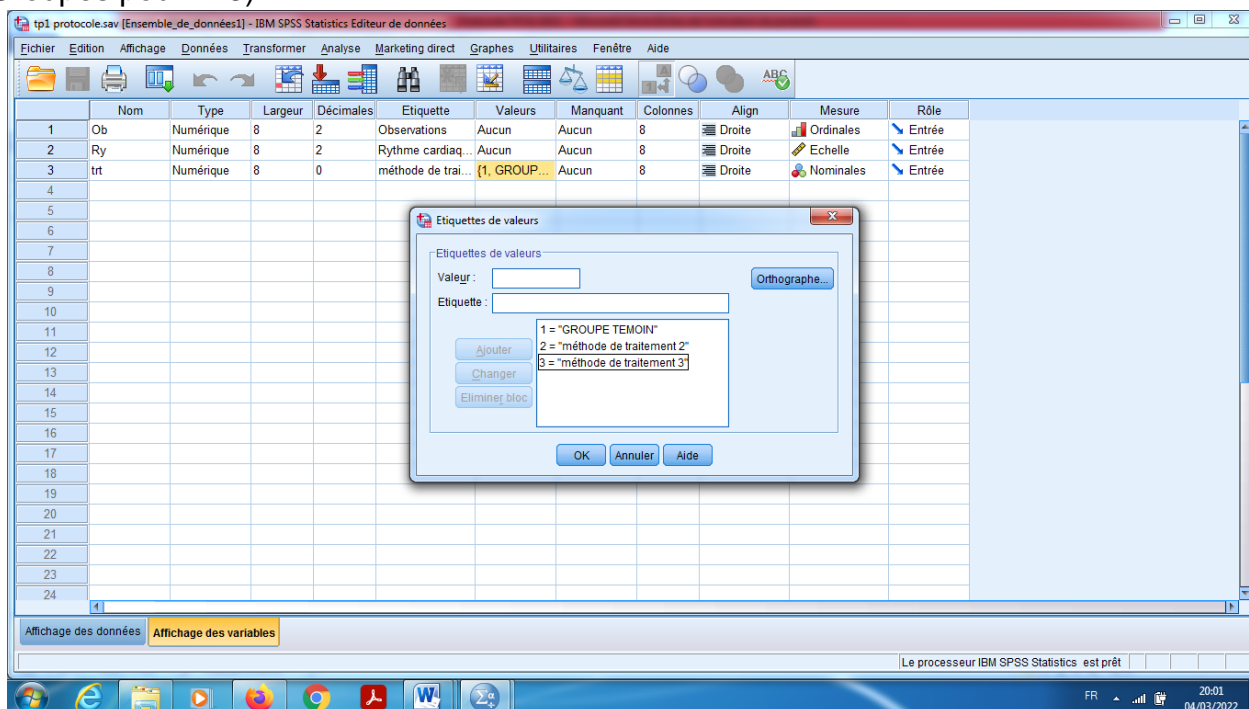
- 1) Notre objectif est de savoir l'existence d'effet éventuel de l'absorption d'un médicament (type de traitement) sur le rythme cardiaque (la variable X), et de plus pour déterminer les groupes qui sont homogènes, ainsi que d'établir qu'elle est la meilleur traitement parmi ses trois types.
- 2) Facteur A « variable qualitative » : la méthode de traitement, dont ses modalités sont groupe témoin, méthode 2 de traitement, méthode 2 de traitement. La variable quantitative mesuré X c'est Rythme cardiaque.
- 3) Hypothèse nulle  $H_0$ : (Il y a homogénéité sur les 3 méthodes de traitement), (tous les moyennes sont significativement égaux), (Il n'a pas une influence du facteur « méthodes de traitement » sur la variable « rythme cardiaque »), ( $\bar{x}_1 = \bar{x}_2 = \bar{x}_3$ ).

Hypothèse alternative  $H_1$ : (Il n'y a pas homogénéité sur les 3 méthodes de traitement), (tous les moyennes sont significativement différentes), (Il n une influence du facteur « méthode de traitement » sur la variable « rythme cardiaque »).

Pour vérifier les propositions d'hypothèses, tout d'abord il faut entrer ces données dans SPSS.

Nous suivons les étapes suivantes :

- a) Il faut définir dans la barre en bas « Affichage des variables » : les variables (qualitative et quantitative) suivantes : observations, Rythme cardiaque, et les Types de traitement (Groupes pour  $k=3$ ).



- b) On choisit les modalités pour la variable qualitative qui représente les échantillons (type de traitement) dans « valeurs », on peut prendre comme un exemple (la valeur 1 pour groupe témoin, et la valeur 2 pour la méthode de traitement 2, et 3 pour la méthode de traitement 3 (aussi pour la méthode de traitement N° :3).
- c) On introduit les données dans la barre « Affichage des données ».

tp 1.sav [Ensemble\_de\_données0] - IBM SPSS Statistics Editeur de données

Visible : 3 variables sur 3

	Ob	Ab	Ry	var	var	var	var	var	var	var	var	var	var	var	var	var
4	4	180,00	groupe Té...													
5	5	190,00	groupe Té...													
6	6	165,00	groupe Té...													
7	7	175,00	groupe Té...													
8	8	174,00	groupe Té...													
9	9	173,00	groupe Té...													
10	10	181,00	groupe Té...													
11	11	155,00	groupe2													
12	12	160,00	groupe2													
13	13	164,00	groupe2													
14	14	150,00	groupe2													
15	15	160,00	groupe2													
16	16	159,00	groupe2													
17	17	154,00	groupe2													
18	18	156,00	groupe2													
19	19	160,00	groupe2													
20	20	167,00	groupe2													
21	21	153,00	groupe2													
22	22	158,00	groupe2													
23	23	140,00	groupe 3													
24	24	155,00	groupe 3													
25	25	153,00	groupe 3													

Affichage des données Affichage des variables

Le processeur IBM SPSS Statistics est prêt

FR 19:44 15/02/2022

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

tp1 protocole.sav [Ensemble\_de\_données1] - IBM SPSS Statistics Editeur de données

Visible : 3 variables sur 3

	Ob	Ry	trt	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1,00	170,00	1													
2	2,00	175,00	1													
3	3,00	187,00	1													
4	4,00	180,00	1													
5	5,00	190,00	1													
6	6,00	165,00	1													
7	7,00	175,00	1													
8	8,00	174,00	1													
9	9,00	173,00	1													
10	10,00	181,00	1													
11	11,00	155,00	2													
12	12,00	160,00	2													
13	13,00	164,00	2													
14	14,00	150,00	2													
15	15,00	160,00	2													
16	16,00	159,00	2													
17	17,00	154,00	2													
18	18,00	156,00	2													
19	19,00	160,00	2													
20	20,00	167,00	2													
21	21,00	153,00	2													
22	22,00	158,00	2													

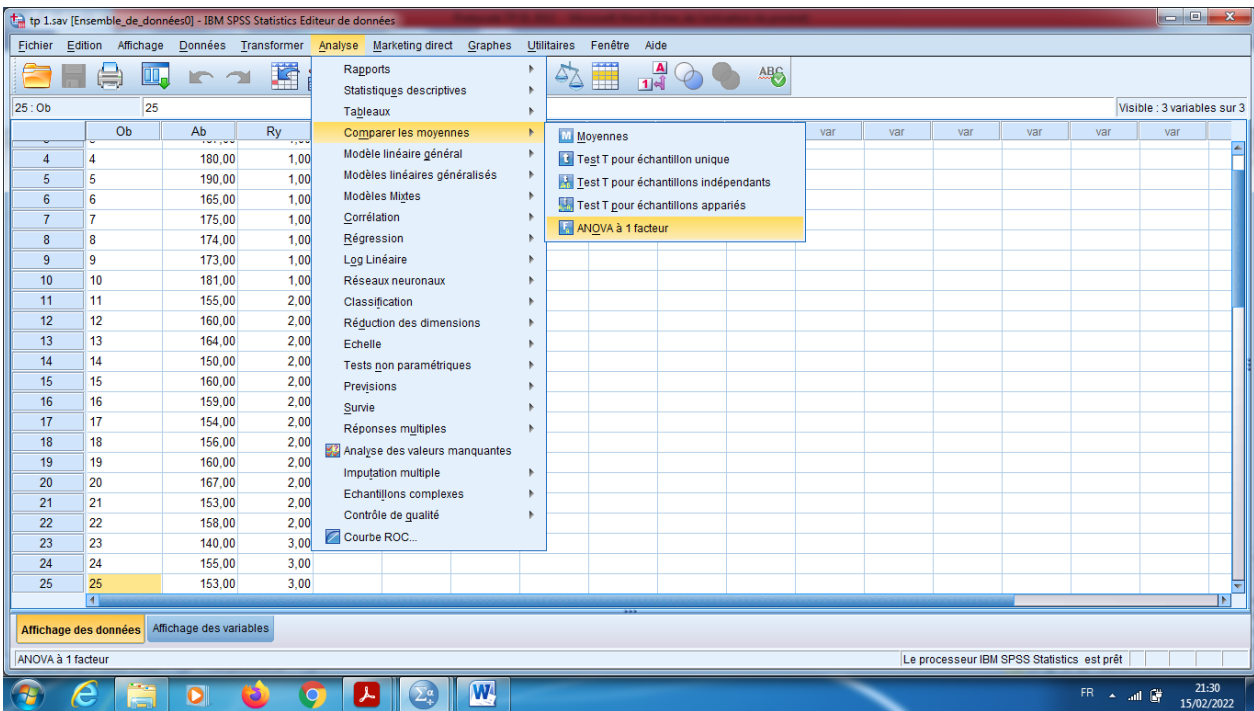
Affichage des données Affichage des variables

Etiquettes de valeurs

Le processeur IBM SPSS Statistics est prêt

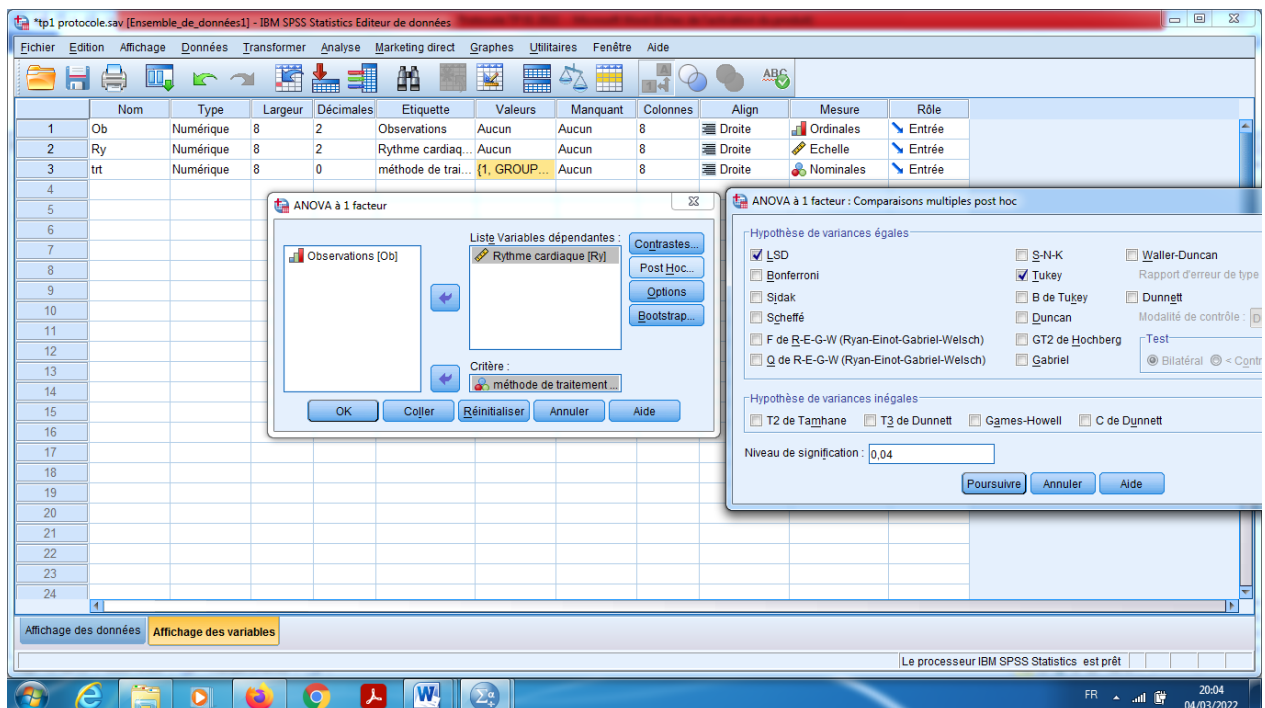
FR 20:38 04/03/2022

Analyse, puis comparer les moyennes, puis ANOVA à 1 facteur.

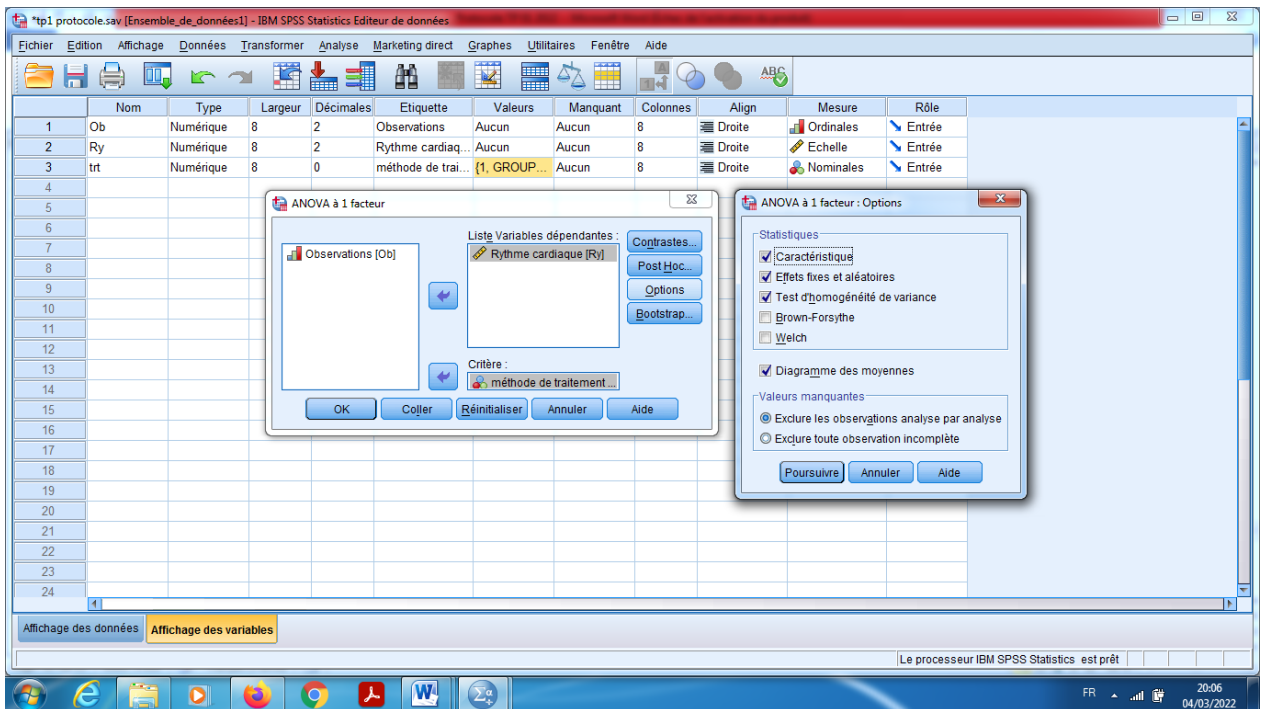


d) On pose la variable X dans le choix « liste des variables dépendantes », et dans le « critères » on pose la variable qui indique le facteur A.

Dans (Post Hoc), on choisit pour hypothèse de variance égales (LSD et Tukey), aussi on prend pour le niveau de signification 4%.



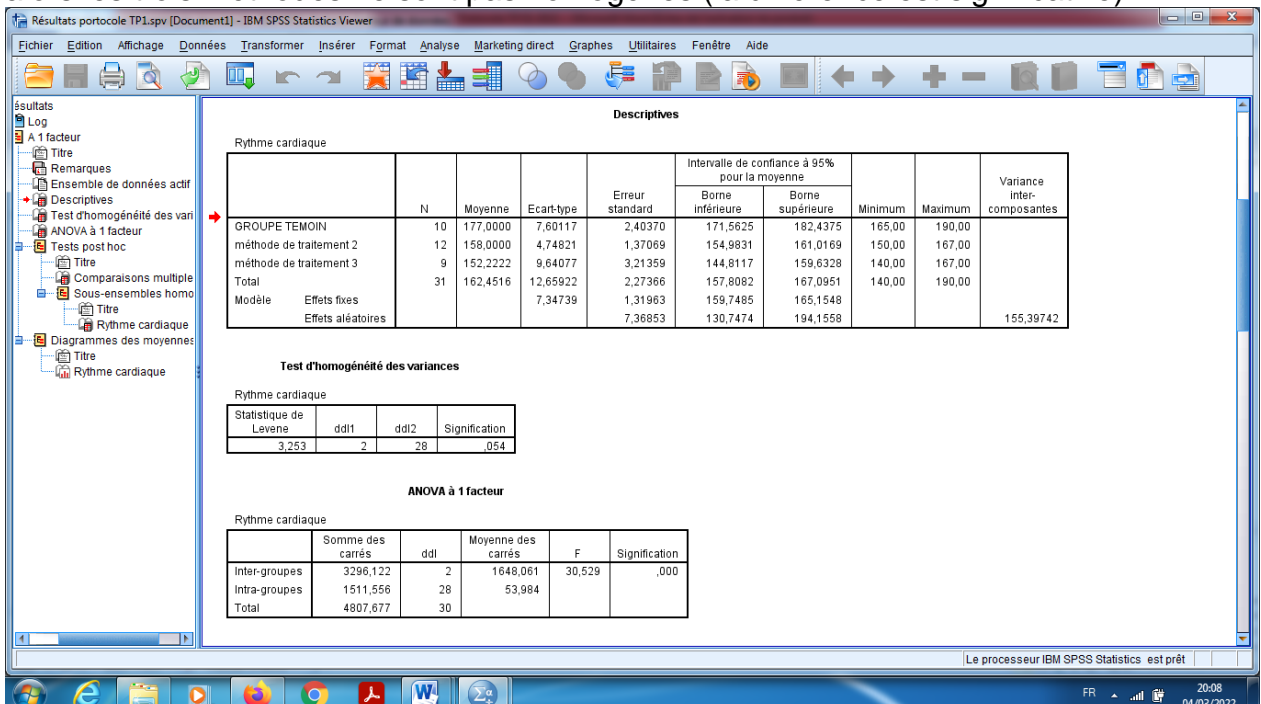
e) Dans le choix « option », on coche sur le caractères, test d'homogénéité de variance, et Diagramme des moyennes.



En fin OK.

### Interprétation des résultats

Pour test d'homogénéité de la variance, c'est le test de Levene, on remarque que  $Sig=0,054 > 0,04$  alors on accepte l'hypothèse de l'homogénéité de la variance. Tableau d'ANOVA qui nous permet d'établir s'il existe un effet du facteur A sur la variable X, ou bien non. On remarque pour cela que ( $Sig=0,000 < 0,04$ ) alors on rejette  $H_0$ , et on accepte de  $H_1$ , c'est-à-dire il y a influence du facteur type de traitement sur la variable rythme cardiaque, alors les trois méthodes ne sont pas homogènes (la différence est significative).



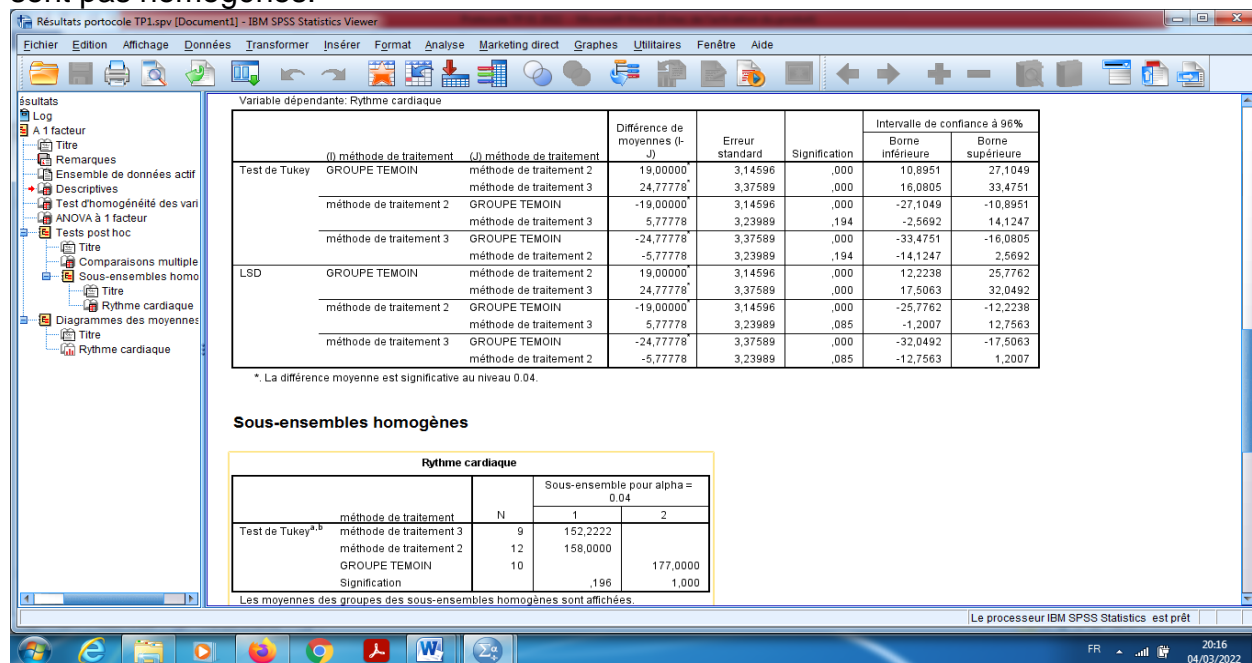
Alors on cherche les groupes qui sont homogènes deux à deux, on effectue le test de LSD ou bien test de Tukey.

On remarque que les groupes 3 et 2 sont homogènes.

En effet : On compare entre (groupe 2 et groupe 3), on remarque que  $Sig=0,194 > 0,04$ . Alors on accepte l'hypothèse  $H_0$  : (méthode 2 et Méthode 3 sont homogène).

De même pour la comparaison entre groupe 3 et groupe 2.

Par contre si on compare entre groupe 1 et 2, on trouve que  $Sig=0,00 < 0,04$ . On accepte  $H_1$  : (méthode 1 et Méthode 2 sont significativement différentes). Alors les deux groupes ne sont pas homogènes.



Finalement pour la meilleur méthode c'est le groupe témoin, et puis le groupe 2 et puis groupe 3.

Alors le groupe témoin est plus efficace par rapport à taux d'absorption, et le plus faible par rapport à taux d'absorption c'est le groupe 3.

