

Master 1 LMD Biologie (Tous les spécialités)

Protocole de TP 02 :

Exécution du test Analyse de la variance

(Execution test of Analysis of Variance with one way)

ANOVA à un facteur

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

- a) Une seule variable quantitative mesurée X, et une seule 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).
- d) Tous les observations sont present au hasard.

I) Test ANOVA

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

Pour cela, on peut poser les questions suivantes :

Existe-il une influence de la variable qualitative A « 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 à 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 H₀: (Il existe un homogénéité sur les k échantillons)=(tous les moyennes sont significativement égales)= (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 H₁: (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 mesurée X).

Etape 2 : Les calculs (tableau d'ANOVA)

Source des variations	Somme des carrés	Ddl	Moyenne des carrés	F
Intergroupes	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 α%. Alors on rejette H₀.

Si Signification supérieure à α%. Alors on accepte H₀.

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 dosages.

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 à mesurer 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'effet du facteur A sur la variable X?

Reponse

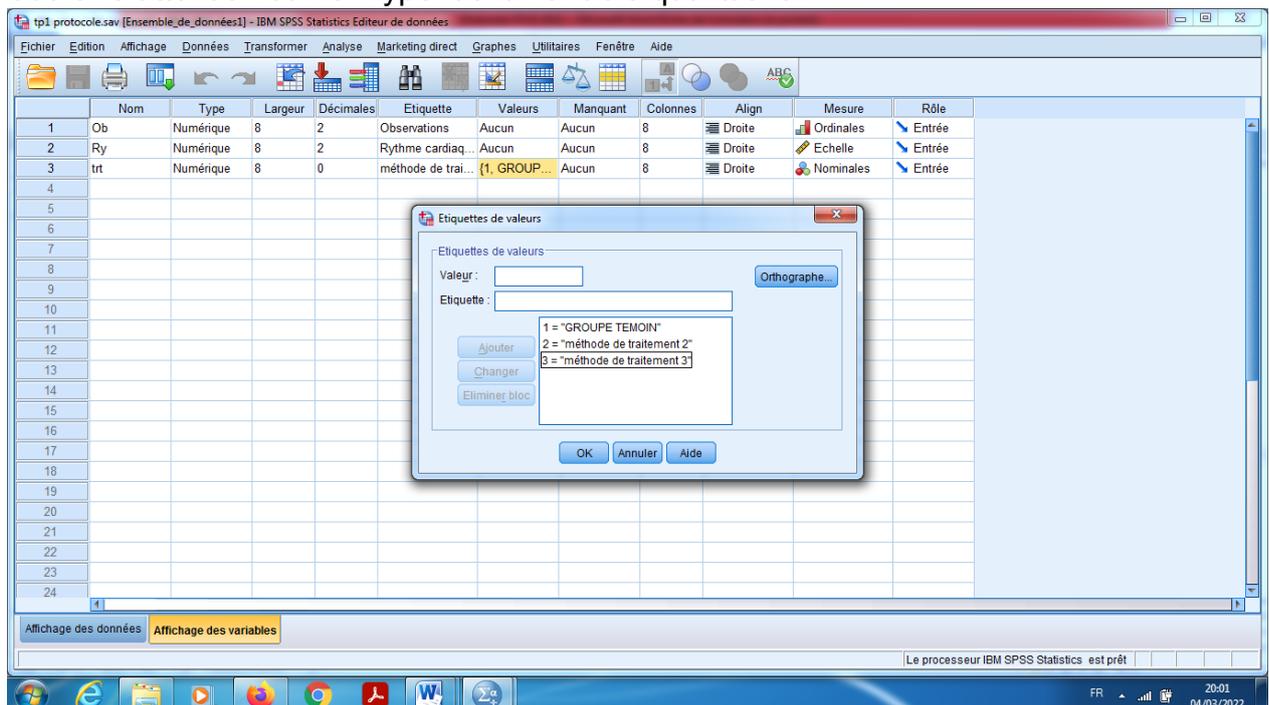
- 1) Notre objectif est de savoir l'existence d'effet éventuel de l'absorption d'un médicament (type de traitement) sur le rythme cardiaque (la variable X), et de plus pour déterminer les groupes qui sont homogènes, ainsi que d'établir qu'elle est la meilleur traitement parmi ses trois types.
- 2) Facteur 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'a pas une influence du facteur « méthode de traitement » sur la variable « rythme cardiaque »).

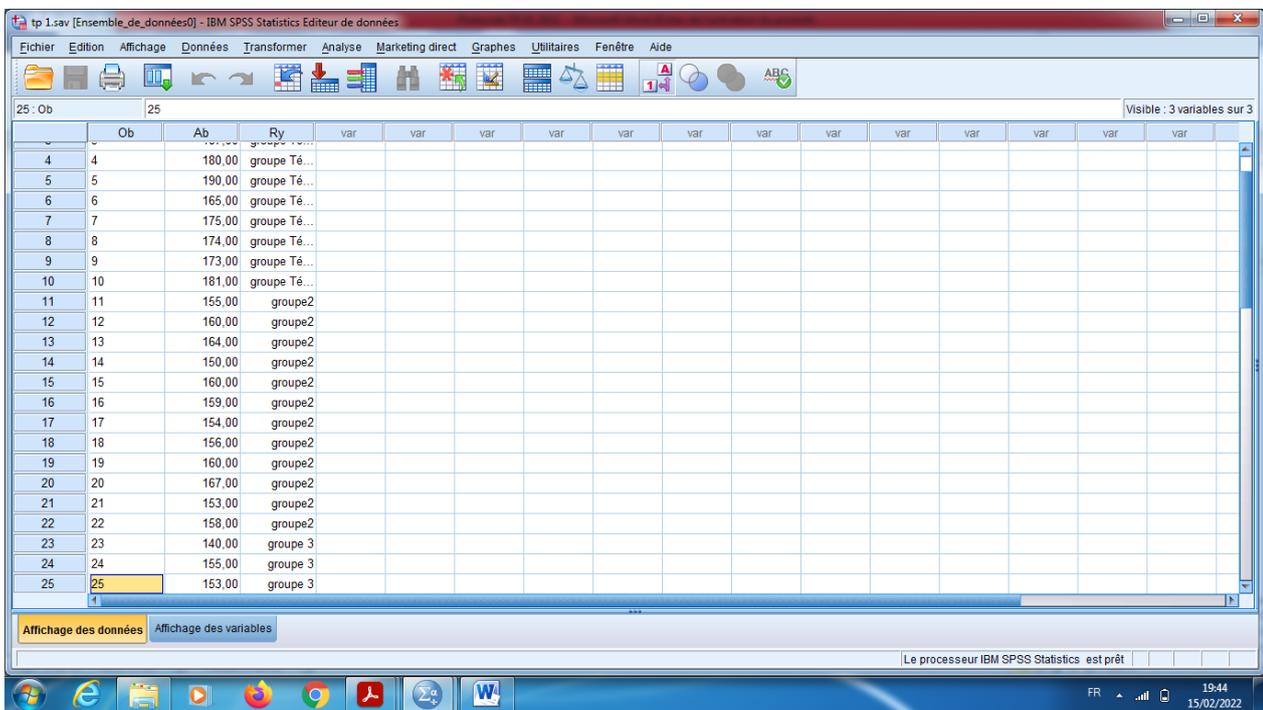
Pour vérifier les propositions d'hypothèses, tout d'abord il faut entrer ces données dans SPSS.

Nous suivons les étapes suivantes :

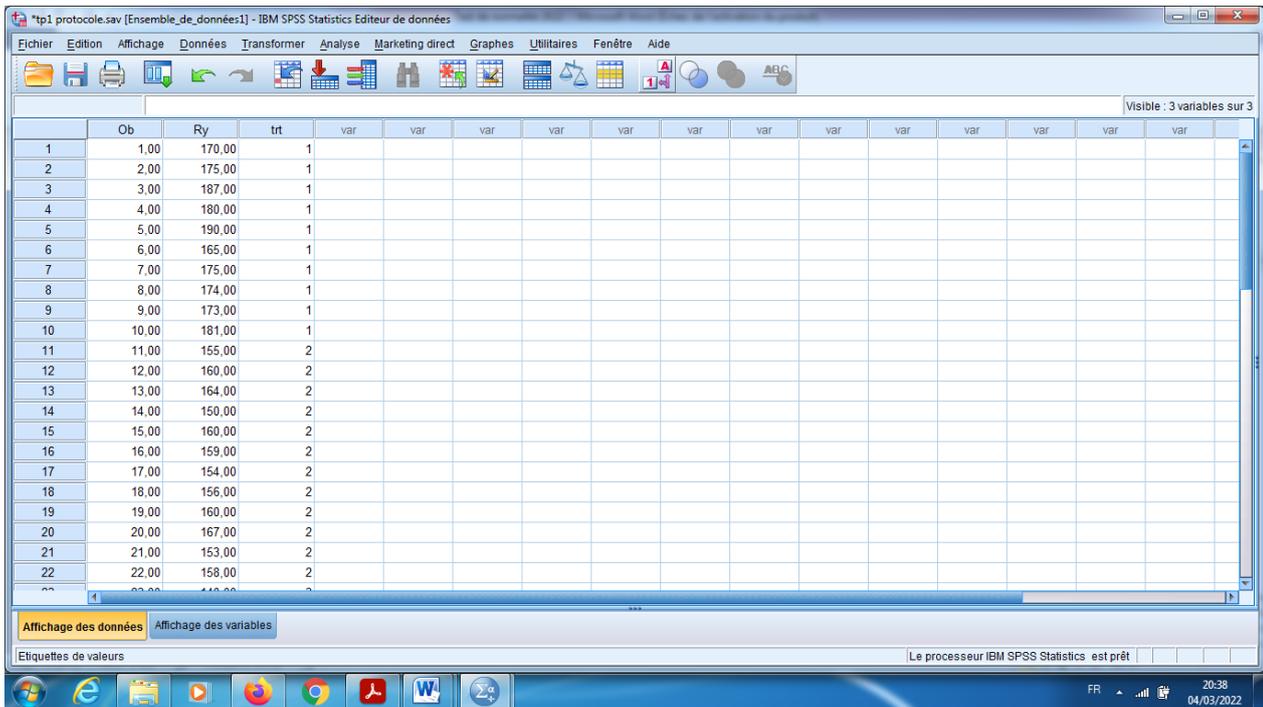
- a) Il faut définir dans la barre en bas « Affichage des variables » : les variables (qualitative et quantitative) suivantes : observations, Rythme cardiaque, et les Types de traitement (Groupes pour $k=3$).
- b) Il faut faire attention sur le 'Type' de la variable qualitative.



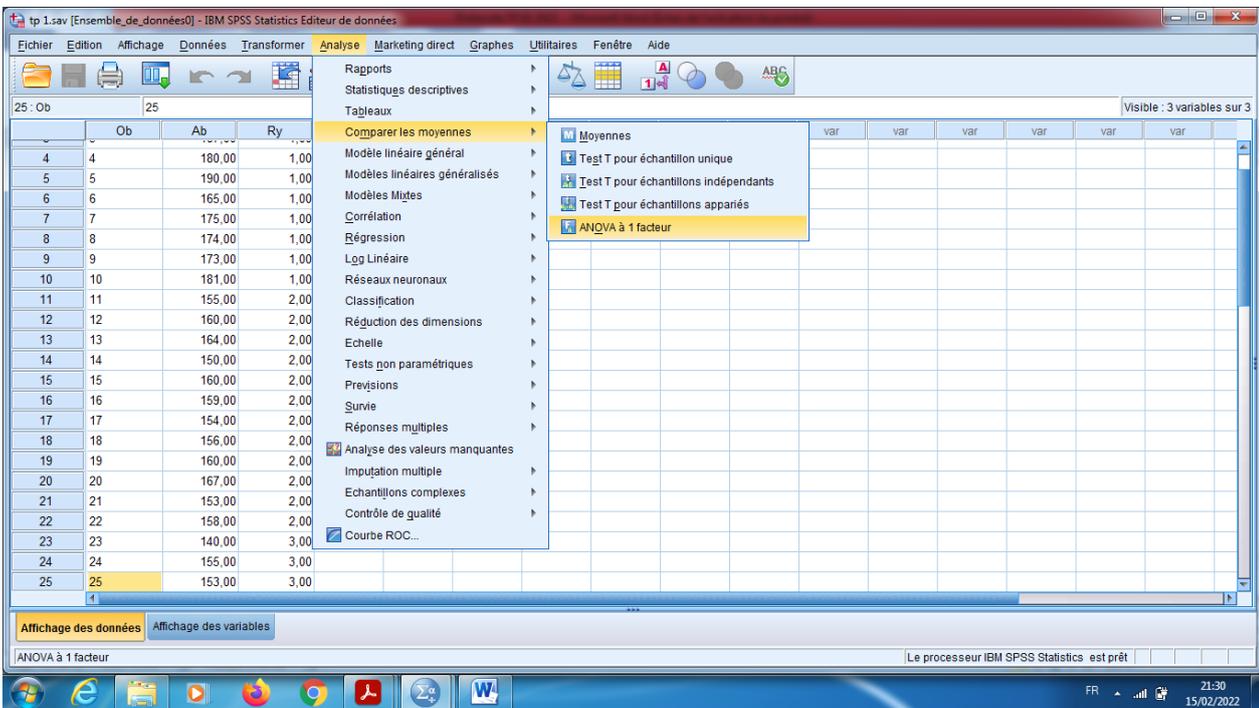
- c) On choisit les modalités pour la variable qualitative qui représente les échantillons (type de traitement) dans « valeurs », on peut prendre comme un exemple (la valeur 1 pour groupe témoin, et la valeur 2 pour la méthode de traitement 2, et 3 pour la méthode de traitement 3 (aussi pour la méthode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



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

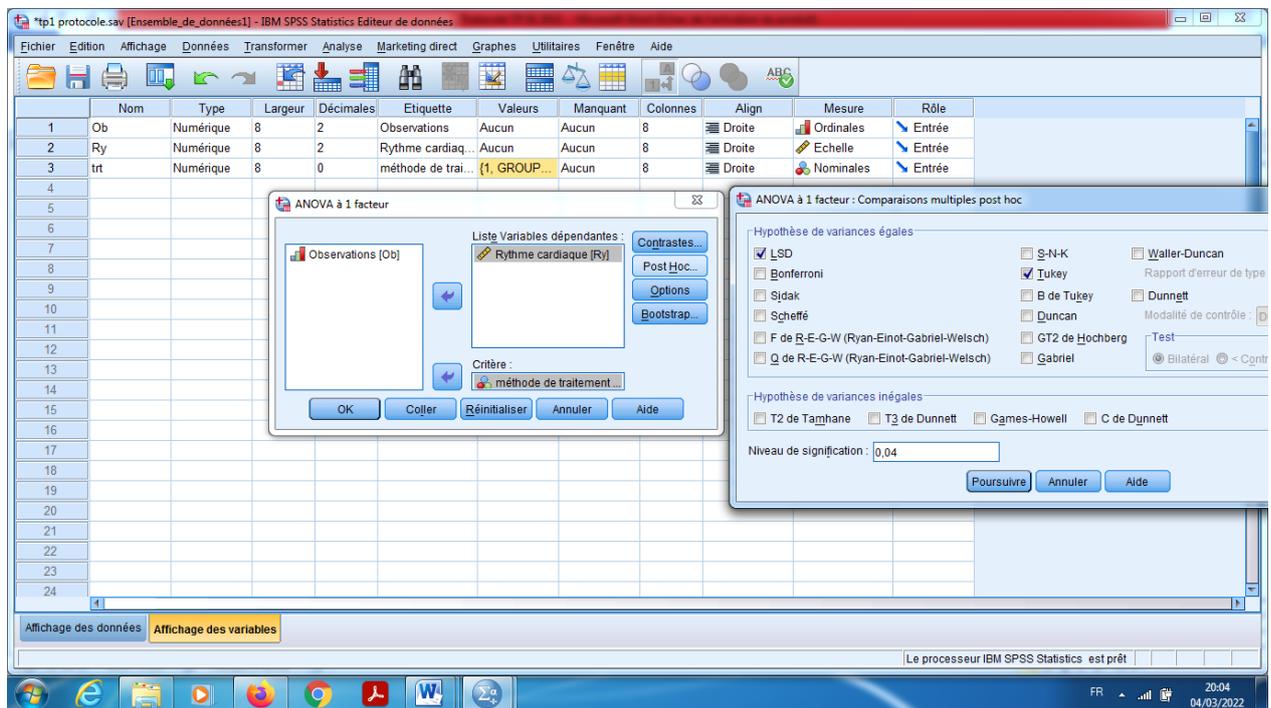


Analyse, puis comparer les moyennes, puis ANOVA à 1 facteur.

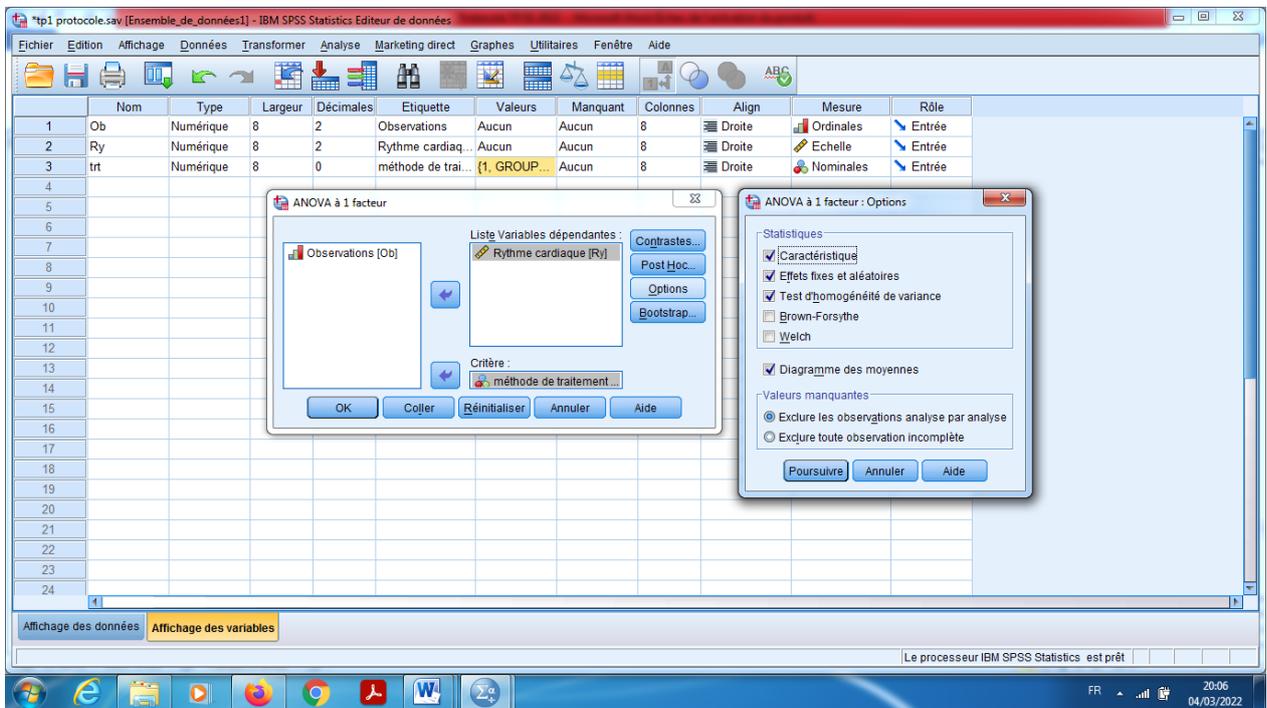


e) On pose la variable X dans le choix « liste des variables dépendantes », et dans le « critères » on pose la variable qui indique le facteur A.

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



f) 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 ya influence du facteur type de traitement sur la variable rythme cardiaque, alors les trois méthodes ne sont pas homogènes (la différence est significative).

Descriptives

Rythme cardiaque	N	Moyenne	Ecart-type	Erreur standard	Intervalle de confiance à 95% pour la moyenne		Minimum	Maximum	Variance inter-composantes
					Borne inférieure	Borne supérieure			
GRUPE TEMOIN	10	177,0000	7,60117	2,40370	171,5625	182,4375	165,00	190,00	
méthode de traitement 2	12	158,0000	4,74821	1,37069	154,9831	161,0169	150,00	167,00	
méthode de traitement 3	9	152,2222	9,64077	3,21359	144,8117	159,6328	140,00	167,00	
Total	31	162,4516	12,65922	2,27366	157,8082	167,0951	140,00	190,00	
Modèle									
Effets fixes			7,34739	1,31963	159,7485	165,1548			
Effets aléatoires				7,36853	130,7474	194,1558			155,39742

Test d'homogénéité des variances

Rythme cardiaque			
Statistique de Levene	ddl1	ddl2	Signification
3,253	2	28	,054

ANOVA à 1 facteur

Rythme cardiaque					
	Somme des carrés	ddl	Moyenne des carrés	F	Signification
Inter-groupes	3296,122	2	1648,061	30,529	,000
Intra-groupes	1511,556	28	53,984		
Total	4807,677	30			

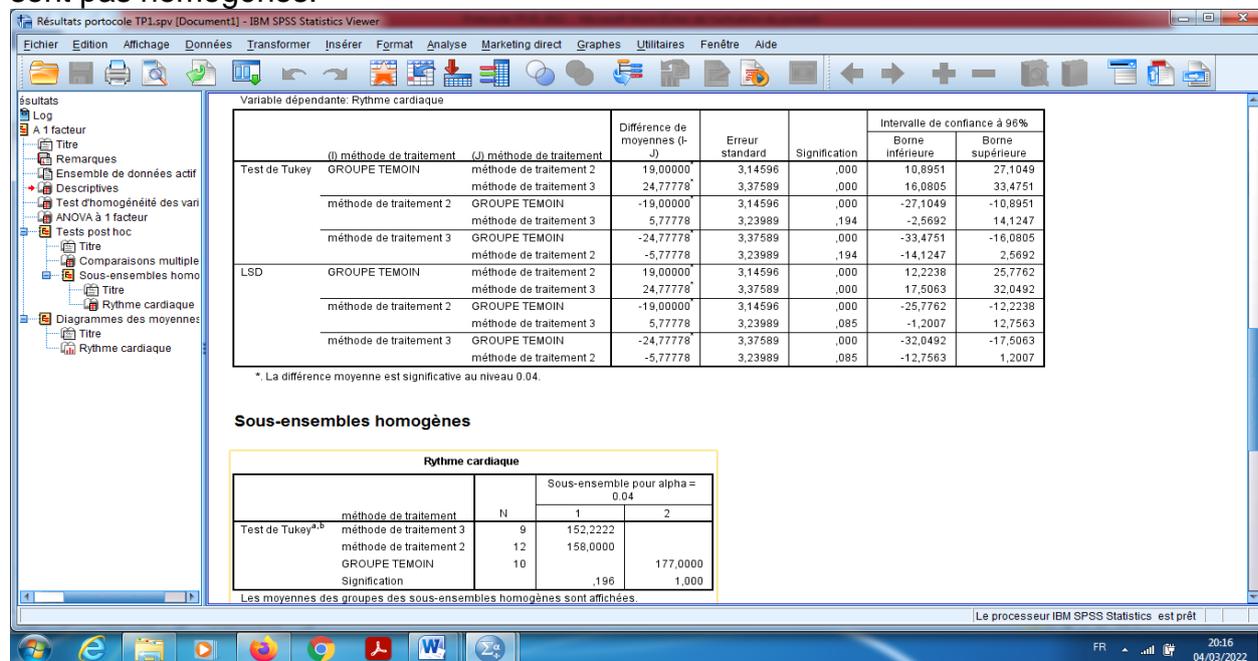
Alors on cherche les groupes qui sont homogènes deux à deux, on effectue le test de LSD ou bien test de Tukey.

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.

