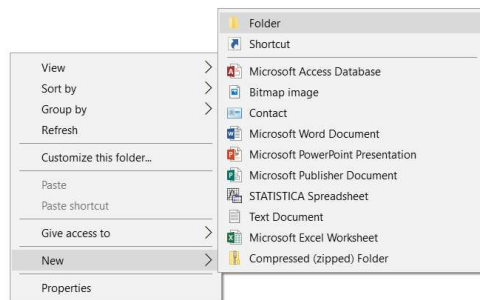


Creation des dossiers:

- Pour créer un nouveau dossier (ex. TP10), on utilise l'option New→Folder (voir l'image suivante):


CHOISIR LE BON TEST STATISTIQUE pour comparer les distributions d'une variable quantitatives sur deux groupes (voir la table ci-dessous) :

Type variable	Nb sujets	Nature des données	Statistique comparée	Test utilisé
Deux groupes indépendants				
Quantitative	$n_1, n_2 \geq$ ou < 30	Normalement distribuées	Différence des moyennes	Test de Student-t
	$n_1, n_2 \geq$ ou < 30	Normalement distribuées	Différence des moyennes	Test de Student-t
Deux groupes dépendants (appariées)				
Quantitative	$n_1 = n_2 \geq$ ou < 30	Normalement distribuées,	Moyenne des différences	Test de Student-t

FORMULATION DE L'HYPOTHESE NULLE DU TEST DE T DE STUDENT - BILATERAL :

L' hypothèse **nulle** (H_0) : **Il n'y a pas de différence statistiquement significative** entre **les MOYENNES de la variable d'intérêt dans les deux sous-populations** (depuis lesquelles les deux échantillons ont été extraites) (la différence entre deux moyennes est égale à zéro)

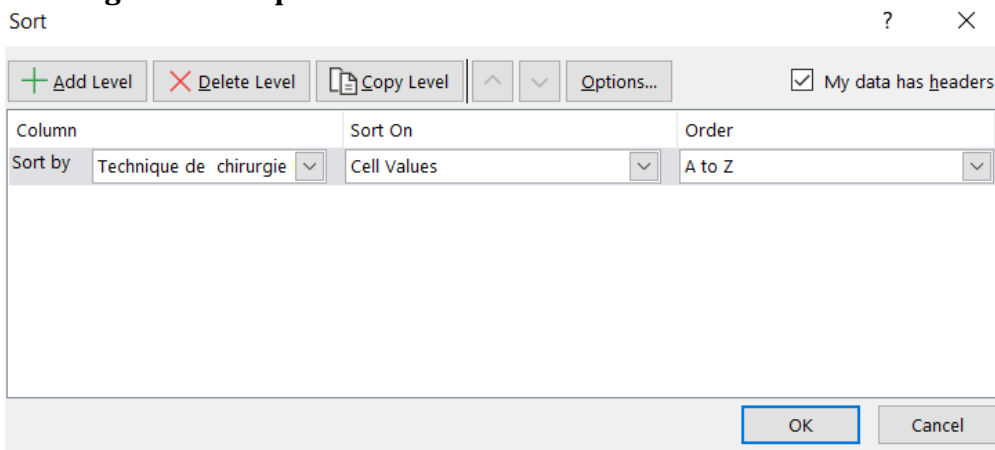
L' hypothèse **alternative** (H_1) : **Il y a une différence statistiquement significative** entre **les MOYENNES de la variable d'intérêt dans les deux sous-populations** (la différence entre deux moyennes n'est pas égale à zéro)

TRIER les données en ascendant/en descendant

- Sélectionnez toute le tableau de données-> **Home**, la section **Editing** (*Editage*), click sur le bouton **Sort & Filter** (*Trier et filtration*).

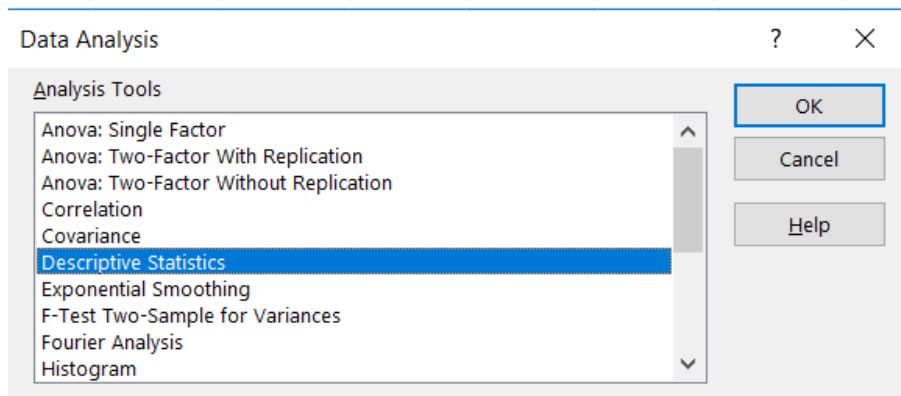


- Choisissez **Custom sort** et puis la variable qui se nomme **Technique de la chirurgie bariatrique**

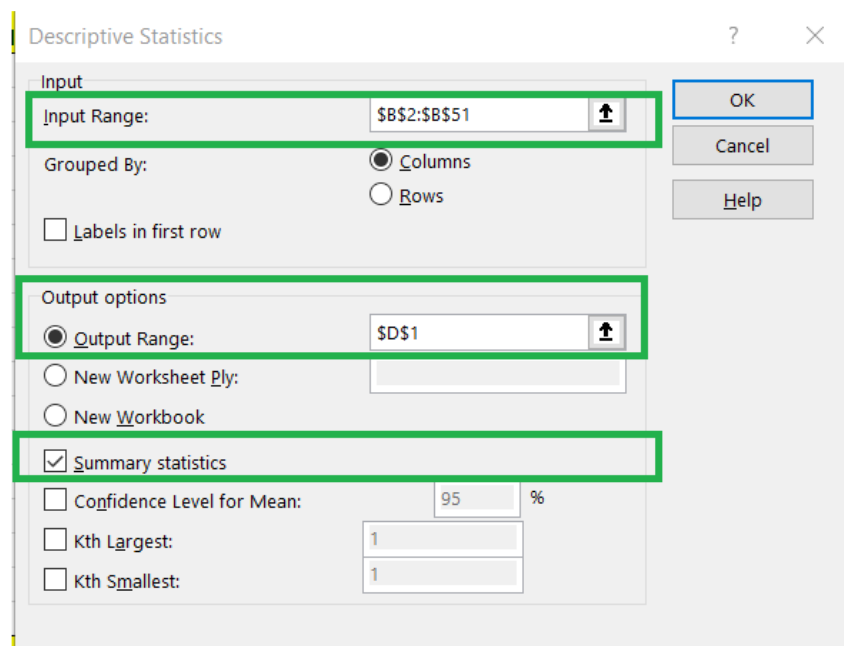


CALCUL DES STATISTIQUES DESCRIPTIVES en utilisant DATA ANALYSIS

- pour calculer les indicateurs descriptives on va utiliser l'option **DESCRIPTIVES STATISTICS** trouvé dans le Data Analysis



- On va remplir l'option *Input range* avec toutes les cellules correspondantes a la variable quantitative (dans notre cas : l'IMC) **chez les patients traités par BG**
- Output range : la cellule désiré pour le résultat (par exemple D1)
- On choisit les *summary statistics*=les statistiques (indicateurs) descriptives



- Puis on va calcules les statistiques descriptives de l'IMC **chez les patients traités par GM**

Descriptive Statistics

Input
 Input Range:

Grouped By: ☒ Columns ☐ Rows

☐ Labels in first row

Output options
☒ Output Range:
☐ New Worksheet Ply:
☐ New Workbook

☒ **Summary statistics**
☐ Confidence Level for Mean: %
☐ Kth Largest:
☐ Kth Smallest:

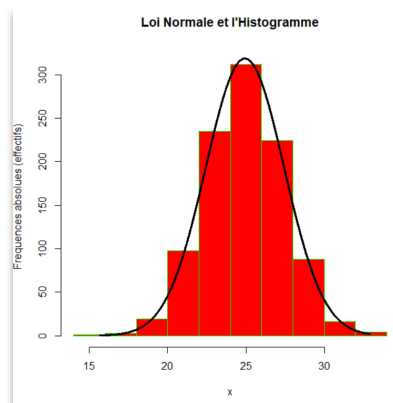
OK Cancel Help

VERIFIER LA NORMALITE DE DONNEES : Approches empiriques (statistiques descriptives) et graphiques

Normalité de données : **i)** la moyenne est approximative égale à la médiane (et le mode) ; **ii)** Le coefficient d'aplatissement est proche à 0 (dans l'intervalle [-1,1]) et **iii)** le coefficient d'asymétrie est proche à 0 (dans l'intervalle [-1,1]).

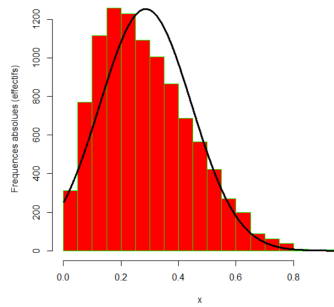
Si l'une des deux (coefficient d'aplatissement et le coefficient d'asymétrie) est en dehors de l'intervalle [-1, 1] \Rightarrow on considère que la distribution **NE suit pas** la **loi Normale**.

Voir en bas un **histogramme** d'une variable ayant une **distribution Normale**.

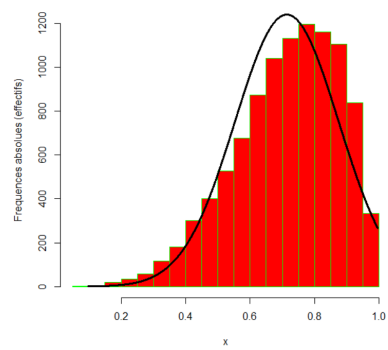


Déviations à la normalité de données : la moyenne est différente de la médiane (et/ou le mode). Le coefficient d'aplatissement **ou/et** le coefficient d'asymétrie sont différents de 0 (>1 ou <-1). Voir en bas un histogramme d'une variable avec distribution avec des **déviations de la loi Normale**.

Distribution avec l'asymetrie a la droite et l'Histogramme

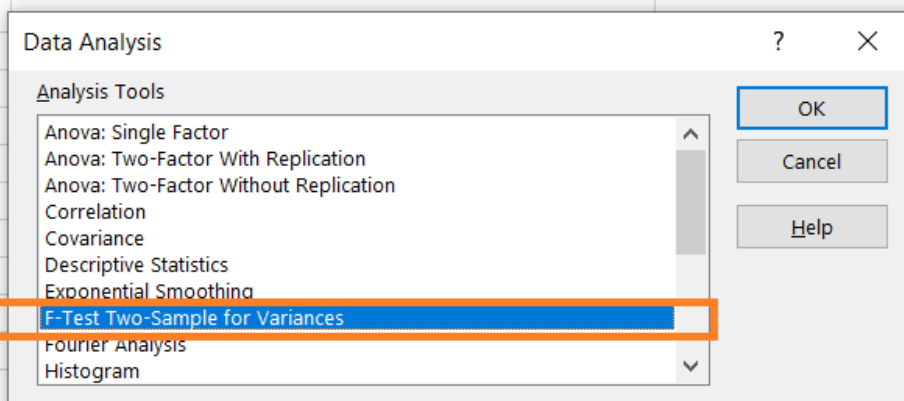


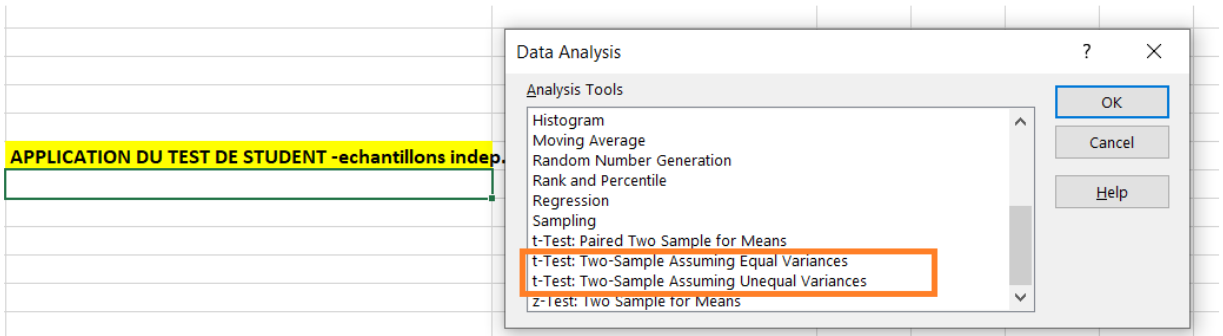
Distribution avec l'asymetrie a la gauche et l'Histogramme



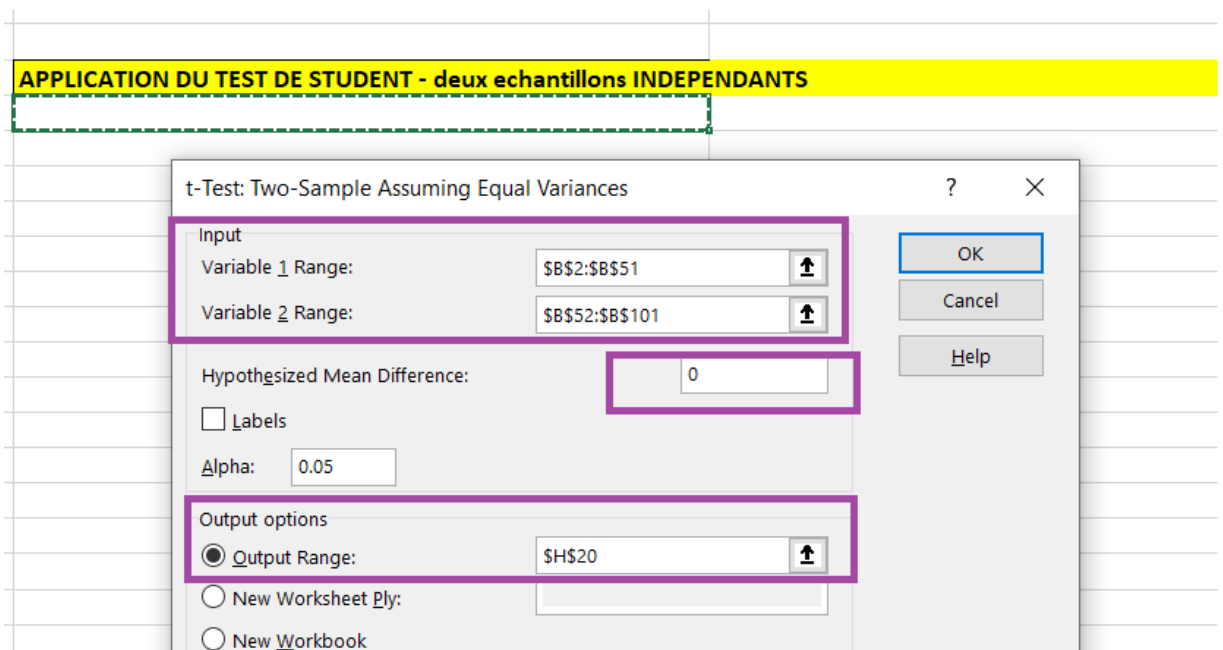
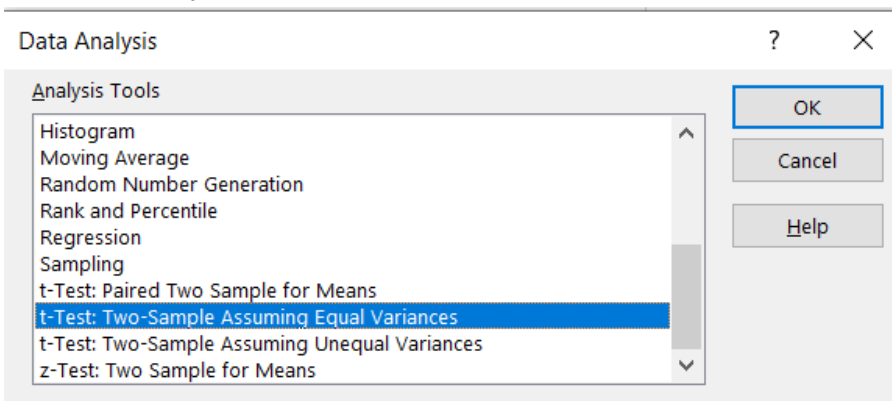
Comment faire un test de FISHER pour deux groupes **INDEPENDANTES** sous Excel??

APPLICATION DU TEST DE FISHER





- On choisit le test t de Student avec des variances EGALES (conformément au résultat du test de FISHER)



APPLICATION DU TEST DE STUDENT - deux échantillons INDEPENDANTS		
t-Test: Two-Sample Assuming Equal Variances		
	Variable 1	Variable 2
Mean	28.64676792	33.9612
Variance	8.35571476	9.110496
Observations	50	50
Pooled Variance	8.733105407	
Hypothesized Mean Difference	0	
df	98	
t Stat	-8.991716335	
P(T<=t) one-tail	9.32139E-15	
t Critical one-tail	1.660551217	
P(T<=t) two-tail	1.86428E-14	
t Critical two-tail	1.984467455	

Interprétation du résultat du test de STUDENT : parce que la p-valeur du test =1.884=0.00...01884<0.05

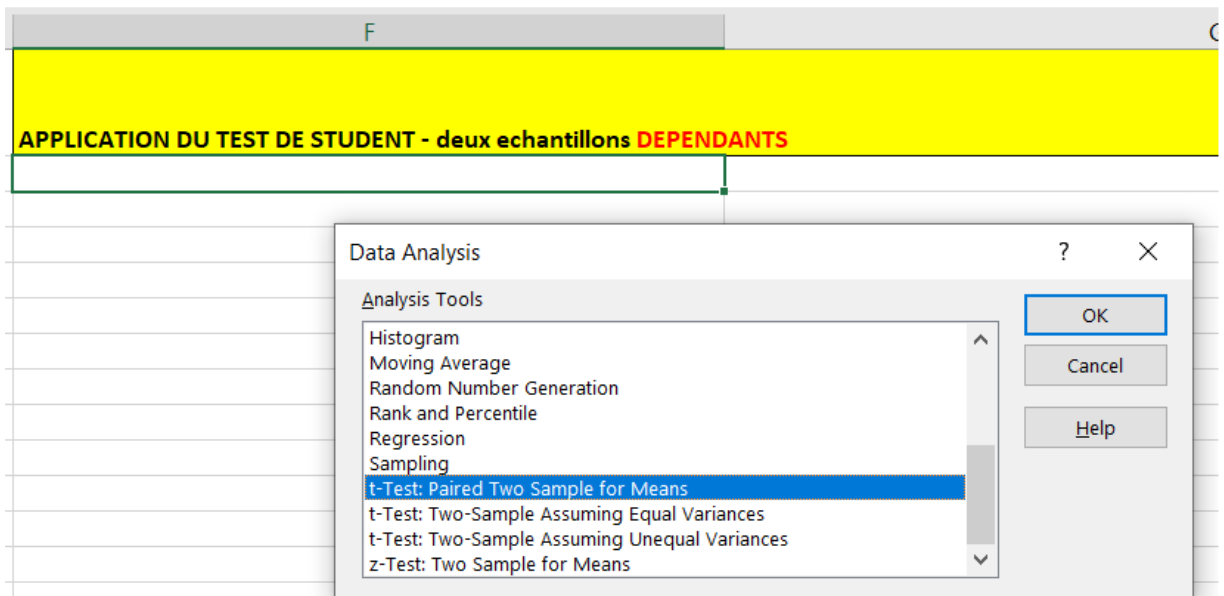
La 14eme decimale sera differente de 0 (et égale a 1)

⇒ on rejette H0 ⇒ nous sommes en faveur de H1 ⇒ Au risque de 5%, il y a une différence significative entre les moyennes de l'IMC une année après chirurgie chez les patients obèses DT2 traités par BG et ceux traités par GM.

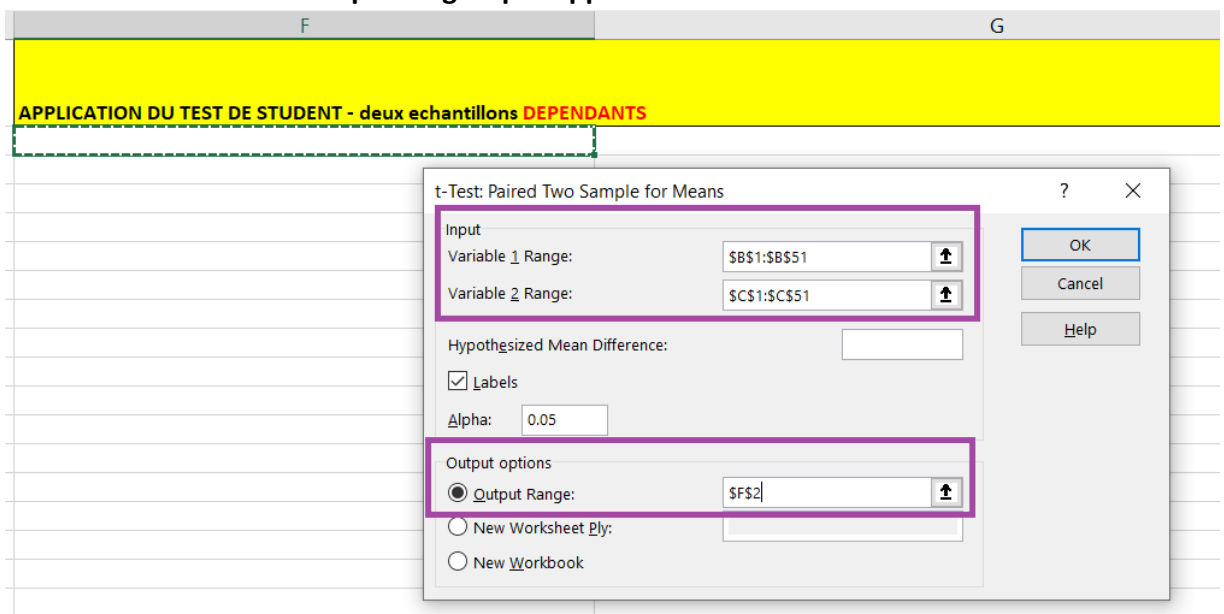
Etapes du test de STUDENT	TABLEAU 1
Etape 1 :	
Formulation de l'hypothèse nulle (H0)	il n'y a pas de différence statistiquement significative entre les moyennes de l'IMC une année après chirurgie chez les patients obèses DT2 traités par BG et ceux traités par GM
Formulation de l'hypothèse alternative (H1)	il y a une différence statistiquement significative entre les moyennes de l'IMC une année après chirurgie chez les patients obèses DT2 traités par BG et ceux traités par GM
Etape 2 :	
Statistique (paramètre) du test	il n'est pas nécessaire de l'écrire
Etape 3 :	
le niveau de signification (alpha)	0.05
Etape 4 :	
Région de rejet (RR)	$(-\infty; -1.981] \cup [1.98; +\infty)$
Etape 5 :	
Valeur calculée (observée) de la statistique du test	-8.99
Etape 6 :	
La décision statistique en fonction de la région de rejet (RR)	parce que -8.99 se trouve dans le RR ⇒ on rejette H0 ⇒ nous sommes en faveur du H1 ⇒ au niveau de signification 5%, il y a une différence statistiquement significative entre les moyennes de l'IMC une année après chirurgie chez les patients obèses DT2 traités par BG et ceux traités par GM
Etape 6' :	
La décision statistique en fonction de la p-valeur	parce que $p < 0.05$ ⇒ on rejette H0 ⇒ nous sommes en faveur du H1 ⇒ au niveau de signification 5%, il y a une différence statistiquement significative entre les moyennes de l'IMC une année après chirurgie chez les patients obèses DT2 traités par BG et ceux traités par GM

Comment faire un test t de Student pour deux groupes **DEPENDANTES** sous Excel??

- On utilise l'option Data Analysis



- choisir le test t de Student pour 2 groupes appariés :



Variable 1 Range: les valeurs de la variable quantitative mesurées au début du traitement

Variable 2 Range: les valeurs de la variable quantitative mesurées à la fin

Alpha: la probabilité d'erreur de type I pour le test statistique.

F	G	H
APPLICATION DU TEST DE STUDENT - deux échantillons DEPENDANTS		
t-Test: Paired Two Sample for Means		
	IMC_Avant (kg/m2)	_Après(kg/m2)
Mean	41.11552114	28.64677
Variance	16.12721528	8.355715
Observations	50	50
Pearson Correlation	-0.034153233	
Hypothesized Mean Difference	0	
df	49	
t Stat	17.53699123	
P(T<=t) one-tail	4.64861E-23	
t Critical one-tail	1.676550893	
P(T<=t) two-tail	9.29722E-23	
t Critical two-tail	2.009575237	

Interprétation du résultat du test de Student : parce que la p-valeur du test ($9.2972E-23 < 0.0001$) $< 0.05 \Rightarrow$ on rejette $H_0 \Rightarrow$ nous sommes en faveur de $H_1 \Rightarrow$ Au risque de 5%, il y a une différence significative entre les moyennes de l'IMC avant et une année après chirurgie chez les patients obèses DT2 traités par BG.

Etapes du test de STUDENT	TABLEAU 1
Etape 1 :	
Formulation de l'hypothèse nulle (H_0)	Il n'y a pas de différence statistiquement significative entre les moyens de l'IMC avant et après chez les patients obèses DT2 traités par BG
Formulation de l'hypothèse alternative (H_1)	Il y a une différence statistiquement significative entre les moyens de l'IMC avant et après chez les patients obèses DT2 traités par BG
Etape 2 :	
Statistique (paramètre) du test	il n'est pas nécessaire de l'écrire
Etape 3 :	
le niveau de signification (α)	0.05
Etape 4 :	
Région de rejet (RR)	$(-\infty; -2.01] \cup [2.01; +\infty)$
Etape 5 :	
Valeur calculée (observée) de la statistique du test	17.54
Etape 6 :	
La décision statistique en fonction de la région de rejet (RR)	parce que 17.54 se trouve dans le RR \Rightarrow on rejette $H_0 \Rightarrow$ nous sommes en faveur du $H_1 \Rightarrow$ au niveau de signification 5%, il y a une différence statistiquement significative entre les moyens de l'IMC avant et après chez les patients obèses DT2 traités par BG
Etape 6' :	
La décision statistique en fonction de la p-valeur	parce que $p < 0.05 \Rightarrow$ on rejette $H_0 \Rightarrow$ nous sommes en faveur du $H_1 \Rightarrow$ au niveau de signification 5%, il y a une différence statistiquement significative entre les moyens de l'IMC avant et après chez les patients obèses DT2 traités par BG

TESTER LES DIFFÉRENCES DE MOYENNES SUR DEUX ÉCHANTILLONS DÉPENDANTS À L'AIDE DE L'INTERVALLE DE CONFIANCE

- dans la feuille Test-t-dep, on va définir la nouvelle variable nommée: **Perte corporelle**

	A	B	C	D
1	Technique de chirurgie bariatrique	IMC_Avant (kg/m2)	IMC_Apres(kg/m2)	Perte Pondérale
2	BG	41.27	29.80	=B2-C2
3	BG	37.89	29.30	
4	BG	40.16	26.97	
5	BG	39.01	22.94	
6	BG	38.78	26.07	
7	BG	32.98	24.72	
8	BG	45.81	23.66	
9	BG	38.40	33.89	
10	BG	46.23	27.28	
11	BG	42.23	26.37	
12	BG	35.95	32.37	
13	BG	34.11	28.01	
14	BG	41.57	34.95	

	A	B	C	D
1	Technique de chirurgie bariatrique	IMC_Avant (kg/m2)	IMC_Apres(kg/m2)	Perte Pondérale
2	BG	41.27	29.80	11.47
3	BG	37.89	29.30	8.60
4	BG	40.16	26.97	13.18
5	BG	39.01	22.94	16.07
6	BG	38.78	26.07	12.70
7	BG	32.98	24.72	8.26
8	BG	45.81	23.66	22.15
9	BG	38.40	33.89	4.51
10	BG	46.23	27.28	18.96
11	BG	42.23	26.37	15.86
12	BG	35.95	32.37	3.58
13	BG	34.11	28.01	6.10
14	BG	41.57	34.95	6.63
15	BG	47.94	27.54	20.39
16	BG	35.18	26.09	9.09
17	BG	41.01	28.79	12.22

- Choisir l'option DATA ANALYSIS:

D	E	F
Perte Pondérale		
11.47		
8.60		
13.18		
16.07		
12.70		
8.26		
22.15		
4.51		
18.96		
15.86		
3.58		
6.10		
6.63		
20.39		
9.09		
12.22		

Descriptive Statistics

Grouped By: ☒ Columns ☐ Rows

☒ Labels in first row

Output options: ☒ Output Range: \$F\$39
 ☐ New Worksheet Ply:
 ☐ New Workbook

☒ Summary statistics

☒ Confidence Level for Mean: 95 %

☐ Kth Largest: 1

☐ Kth Smallest: 1

Perte Pondérale	
Mean	12.46875322
Standard Error	0.710997289
Median	12.34851178
Mode	#N/A
Standard Deviation	5.027510044
Sample Variance	25.27585724
Kurtosis	-0.76441875
Skewness	0.156398221
Range	20.0764116
Minimum	3.575280779
Maximum	23.65169238
Sum	623.437661
Count	50
Confidence Level(95.0%)	1.428802545

$$ME = t_{\alpha/2, n-1} * \frac{s}{\sqrt{n}};$$

ou $\frac{s}{\sqrt{n}} = ES =$
erreur standard

- Créez en dessous des résultats obtenus, un tableau similaire à celui ci-dessous dans lequel l'IC à 95% sera calculé :

Perte Pondérale	
Mean	12.46875322
Standard Error	0.710997289
Median	12.34851178
Mode	#N/A
Standard Deviation	5.027510044
Sample Variance	25.27585724
Kurtosis	-0.76441875
Skewness	0.156398221
Range	20.0764116
Minimum	3.575280779
Maximum	23.65169238
Sum	623.437661
Count	50
Confidence Level(95.0%)	1.428802545
Borne inferieure de l'IC	=G41-G54
Borne superieure de l'IC	
95% IC	

Perte Pondérale	
Mean	12.46875322
Standard Error	0.710997289
Median	12.34851178
Mode	#N/A
Standard Deviation	5.027510044
Sample Variance	25.27585724
Kurtosis	-0.76441875
Skewness	0.156398221
Range	20.0764116
Minimum	3.575280779
Maximum	23.65169238
Sum	623.437661
Count	50
Confidence Level(95.0%)	1.428802545
Borne inferieure de l'IC	11.03995067
Borne superieure de l'IC	=G41+G54
95% IC	

Perte Pondérale	
Mean	12.46875322
Standard Error	0.710997289
Median	12.34851178
Mode	#N/A
Standard Deviation	5.027510044
Sample Variance	25.27585724
Kurtosis	-0.76441875
Skewness	0.156398221
Range	20.0764116
Minimum	3.575280779
Maximum	23.65169238
Sum	623.437661
Count	50
Confidence Level(95.0%)	1.428802545
Borne inferieure de l'IC	11.03995067
Borne superieure de l'IC	13.89755577
95% IC	[11.04; 13.90]

Interpretation de l'IC pour la difference des moyennes: Nous sommes sûrs à 95 % que l'intervalle [12,41 ; 14,77] contiendra la différence des moyennes d'IMC avant et un an après l'intervention chez les patients obèses DT2 traités par BG. Puisque 0 n'est pas dans cette plage → **il existe une différence significative entre l'IMC moyen avant et un an après la chirurgie chez les patients obèses DT2 traités par BG (avec un risque d'erreur de 5 %).**