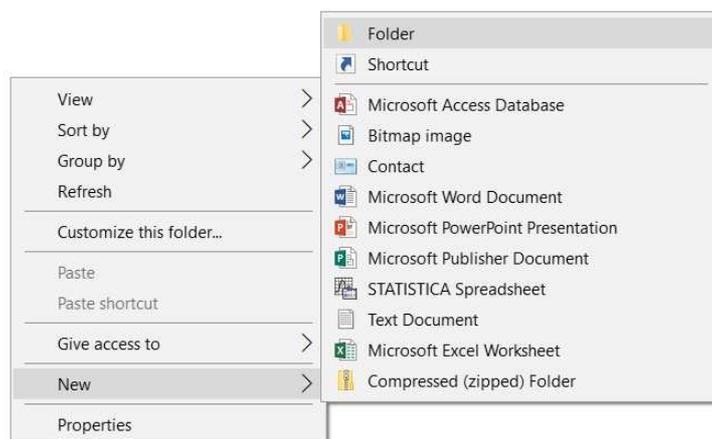


N'oubliez que:

Creation des dossiers:

- créer le dossier nommé **TP08NP sur le bureau (Desktop) de l'ordianteur.**
- Pour créer un nouveau dossier (ex. TP08NP), on utilise l'option New→Folder (voir l'image suivante)

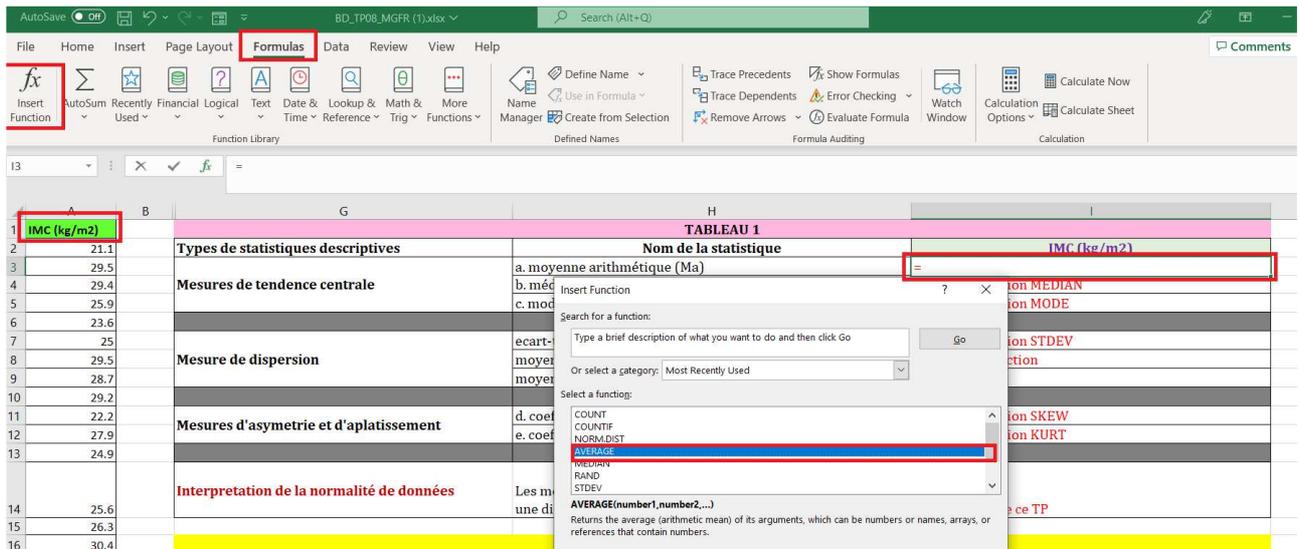


Mesures descriptives calculées en utilisant les fonctions predefinis de l'Excel

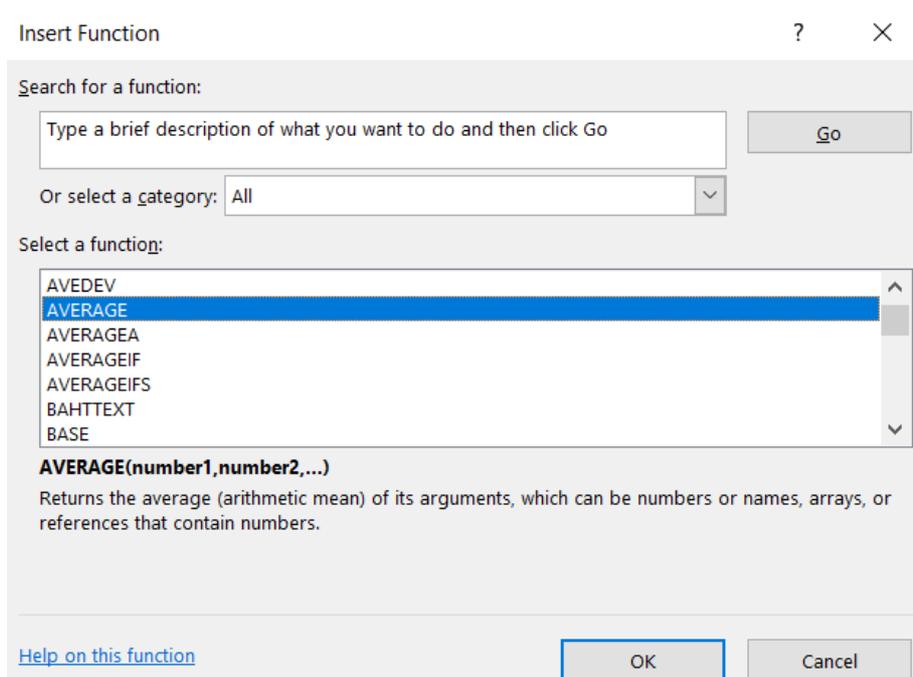
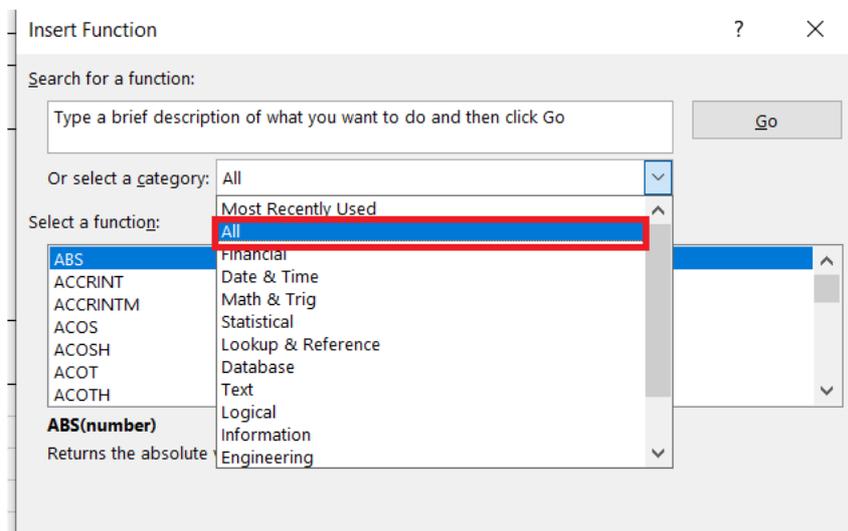
Mesure	Fonction dans Excel en Anglais	En Français
Moyenne	AVERAGE	MOYENNE
Médiane (= quartile 2)	MEDIAN	MEDIANE
Mode	MODE	MODE
Déviatiion standard/écart type	STDEV	ECARTYPE
Coefficient d'aplatissement	KURT	KURTOSIS
Coefficient d'asymétrie	SKEW	COEFFICIENT.ASYMETRIE

Exemple du calcul de la moyenne :

- Sélectionner la cellule du tableau contenant les mesures descriptives pour la variable IMC qui attends le résultat de la moyenne(dans la cellule I3)



○ Choisir FORMULAS-> INSERT FUNCTIONS-> ALL-> AVERAGE



○ **Sélectionner tous les valeurs de l'IMC :**

Function Arguments ? ×

AVERAGE

Number1 = {21.1;29.5;29.4;25.9;23.6;25;29.5;28.7;29.2;22.2;27...

Number2 = number

= 24.88870968

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.

Number1: number1,number2,... are 1 to 255 numeric arguments for which you want the average.

Formula result = 24.88870968

[Help on this function](#)

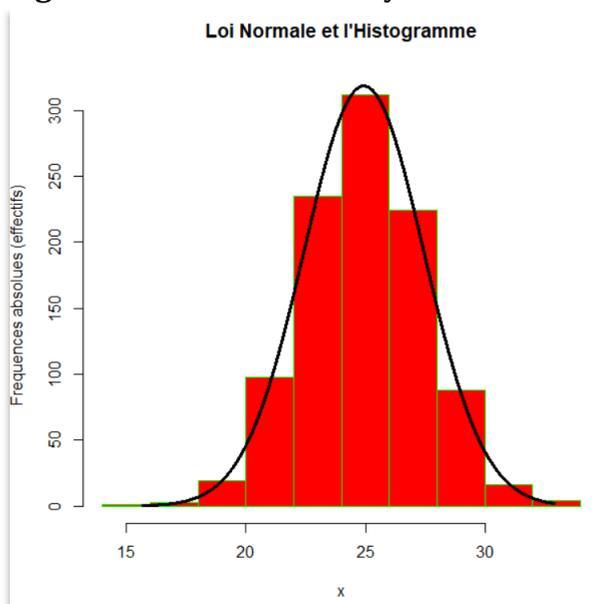
○ **L'IMC moyen sera égale a 24.89 kg/m2.**

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]**).

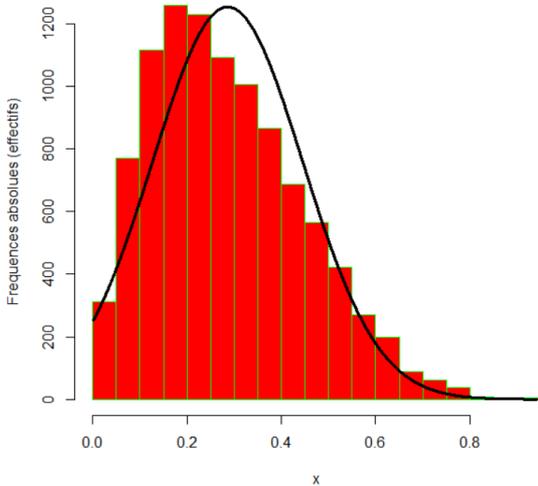
Si l'une des deux (coefficient d'aplatissement et le coefficient d'asymétrie) est en dehors de l'intervalle **[-1, 1]** ⇒ on considère que la distribution **NE suit pas la loi Normale**.

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

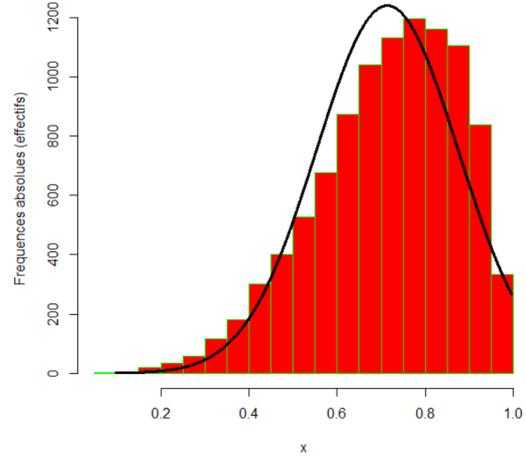


Déviations à la normalité de données : la moyenne est différente de la médiane (et/ou le mode). Le coefficient d'aplatissement **ou/et** le coefficient d'asymétrie sont différentes de 0 (>1 ou <-1). Voir en bas un histogramme d'une variable avec distribution avec des **déviations de la loi Normale**.

Distribution avec l'asymétrie à la droite et l'Histogramme



Distribution avec l'asymétrie à la gauche et l'Histogramme



CALCUL DES DIFFERENTS PROBABILITES EN UTILISANT LA LOI NORMALE

○ Pour calculer le pourcentage de sujets ayant le taux du IMC \leq moyenne+1.96DS choisir la cellule I8 du Tableau 2 de la page Statistique descriptive et puis chercher la fonction **NORM.DIST** parmi les fonctions de l'Excel :

TABLEAU 2		
Evènements	Probabilité	La valeur estimée de la probabilité est :
M = {avoir Calcium \leq 8 mg/dl }	Pr(M)= ?	=
N = {avoir 8 \leq Calcium \leq 10}	Pr(N)= ?	
T = {avoir moyenne-2·DS \leq Calcium \leq moyenne+2·DSCalcium }	Pr(T)= ?	

Insert Function ? X

Search for a function:

Type a brief description of what you want to do and then click Go

Or select a category: All

Select a function:

- NEGBINOM.DIST
- NEGBINOMDIST
- NETWORKDAYS
- NETWORKDAYS.INTL
- NOMINAL
- NORM.DIST**
- NORM.DIST

○ la syntaxe de la **fonction NORM.DIST** sera:

Function Arguments

NORM.DIST

X	19	=	32.03634877	← moyenne+ 1.96ecart-type
Mean	13	=	24.88870968	← moyenne
Standard_dev	17	=	3.646754637	← écart-type
Cumulative	TRUE	=	TRUE	

returns the normal distribution for the specified mean and standard deviation.

Cumulative is a logical value: for the cumulative distribution function, use TRUE; for the probability density function, use FALSE.

formula result = 0.975002105

[help on this function](#) OK Cancel

- Pour calculer le pourcentage de sujets ayant le taux du moyenne $-1.96DS \leq IMC \leq 1moyenne+1.96DS$, choisir la cellule F21 du Tableau 2 de la page Statistique descriptive et puis faites la différence entre les deux probabilités antérieurement calculés :

TABLEAU 2		
Evènements	Probabilité	La valeur estimée de la probabilité est :
M = {avoir IMC ≤ moyenne + 1.96·DS }	Pr(M)= ?	0.975002105
M = {avoir IMC ≤ moyenne - 1.96·DS }	Pr(N)= ?	0.29547351
Q = {avoir moyenne-1.96·DS ≤ IMC ≤ moyenne+1.96·DS }	Pr(Q)= ?	=I18-I19

- **Interpretation de l'intervale de fluctuation [moyenne-1.96DS, moyenne+1.96DS]:**
95% de jeunes adultes ont l'IMC compris dans l'intervalle: 17.74 – 32.04 kg/m².

Calcul de **L'INTERVALLE DE CONFIANCE DE LA MOYENE** par l'option **DATA ANALYSIS**

1. Dans la page nommée **ICMoyenne**, on va estimer la moyenne de l'IMC en utilisant l'option **DATA→ DATA ANALYSIS→ DESCRIPTIVE STATISTICS**

Descriptive Statistics

Input

Input Range: SAS1:SAS249

Grouped By: Columns Rows

Labels in first row

Output options:

Output Range: SDS1

New Worksheet Ply:

New Workbook

Summary statistics

Confidence Level for Mean: 95 %

Kth Largest: 1

Kth Smallest: 1

OK Cancel Help

2. Les résultats obtenus sont les suivantes :

D	E	F	G	H	I	J
IMC (kg/m2)						
Mean	24.88870968					
Standard Error	0.231569151					
Median	24.55					
Mode	25.9					
Standard Deviation	3.646754637					
Sample Variance	13.29881938					
Kurtosis	0.554638764					
Skewness	0.638469384					
Range	20.3					
Minimum	16.5					
Maximum	36.8					
Sum	6172.4					
Count	248					
Confidence Level(95.0%)	0.456102017					

Marge d'erreur (ME) = $t_{\alpha/2, n-1} * \frac{S}{\sqrt{n}}$

3. Pour déterminer les 2 bornes de l'IC (la limite inférieure et la limite supérieure), la formule pour 95% IC est la suivante :

$$\left[m - t_{\alpha/2, n-1} * \frac{S}{\sqrt{n}} ; m + t_{\alpha/2, n-1} * \frac{S}{\sqrt{n}} \right]$$

Borne inférieure de l'IC

Borne supérieure de

Ou **n**=taille de l'échantillon ; **m**=moyenne ; **S**=déviation standard d'échantillonnage

D	E
IMC (kg/m2)	
Mean	24.88870968
Standard Error	0.231569151
Median	24.55
Mode	25.9
Standard Deviation	3.646754637
Sample Variance	13.29881938
Kurtosis	0.554638764
Skewness	0.638469384
Range	20.3
Minimum	16.5
Maximum	36.8
Sum	6172.4
Count	248
Confidence Level(95.0%)	0.456102017
Borne inférieure de l'IC	=E3-E16
Borne supérieure de l'IC	
95% IC	

D	E
<i>IMC (kg/m2)</i>	
Mean	24.88870968
Standard Error	0.231569151
Median	24.55
Mode	25.9
Standard Deviation	3.646754637
Sample Variance	13.29881938
Kurtosis	0.554638764
Skewness	0.638469384
Range	20.3
Minimum	16.5
Maximum	36.8
Sum	6172.4
Count	248
Confidence Level(95.0%)	0.456102017
Borne inferieure de l'IC	24.43260766
Borne superieure de l'IC	=E3+E16

Borne inferieure de l'IC	24.43260766
Borne superieure de l'IC	25.34481169
95% IC	=

Insert Function ? X

Search for a function:

Type a brief description of what you want to do and then click Go Go

Or select a category: Most Recently Used

Select a function:

- CONFIDENCE T
- CONCATENATE**
- NORMDIST
- AVERAGE
- COUNT
- COUNTIF
- MEDIAN

CONCATENATE(text1;text2;...)
Joins several text strings into one text string

Borne inferieure de l'IC	24.43260766
Borne superieure de l'IC	25.34481169
95% IC	.34)

Function Arguments

CONCATENATE

Text1	24.43	= "24.43"
Text2	"-"	= "-"
Text3	25.34	= "25.34"
Text4		= text

ecrire les bornes de l'IC

= "24.43-25.34"

Joins several text strings into one text string.

Text3: text1,text2,... are 1 to 255 text strings to be joined into a single text string and can be text strings, numbers, or single-cell references.

Formula result = 24.43-25.34

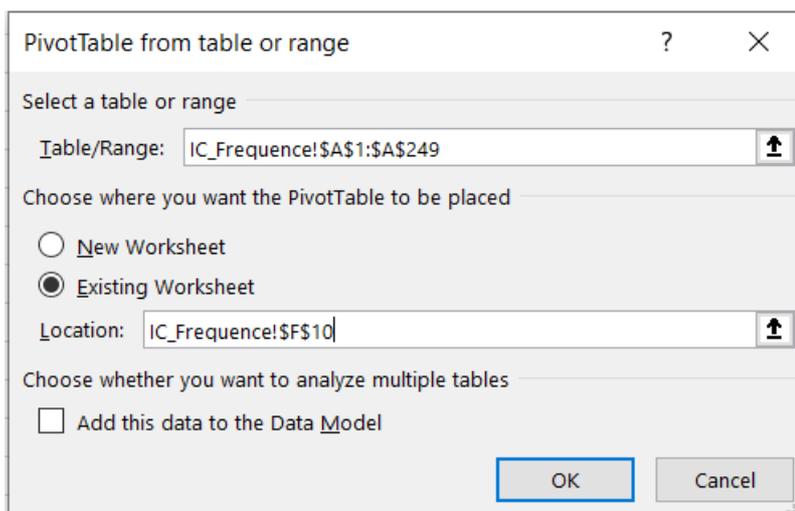
- 95% IC sera de la forme suivante:

Borne inferieure de l'IC	24.43260766
Borne superieure de l'IC	25.34481169
95% IC	24.43-25.34

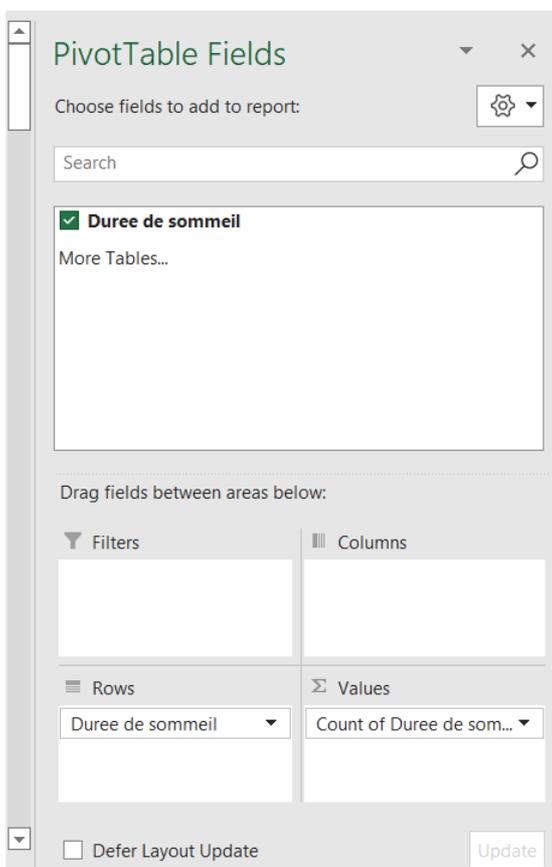
Borne inferieure de l'IC	24.43260766
Borne superieure de l'IC	25.34481169
95% IC	24.43-25.34
Interpretation de l'IC:	Nous sommes sûrs à 95% que l'intervalle de confiance [24.43, 25.34] contient l'IMC moyen de la population des jeunes adultes.

Travail sur l'Excel: **CALCUL DE L'INTERVALLE DE CONFIANCE D'UNE FREQUENCE**

1. Réalisez un tableau croise (engl.pivot table) pour la variable « **Durée de sommeil** »
2. On choisi une cellule de la base de donnees (par ex, A2) ->Insert-> Pivot Table



3. Choisir la variable d'intérêt suivant le model :



4. On obtient la table suivant :

Row Labels	Count of Duree de sommeil
optimale	161
reduite	87
Grand Total	248

5. Dans le TABLEAU 1 on calcule la fréquence relative de la gingivite (défini comme le rapport entre les sujets ayant une durée de sommeil réduite et l'effectif total de l'échantillon)

F	G
Tableau 1	
Fréquence relative de la durée REDUITE de sommeil	=87/248
Erreur standard de la fréquence (ES)	
Borne inférieure de l'IC	
Borne supérieure de l'IC	
95% IC	
Interpretation de l'IC	
Row Labels	Count of Duree de sommeil
optimale	161
reduite	87
Grand Total	248

- 6 L'intervalle de confiance de la fréquence de la gingivite a la formule suivante :

95% CI pour la fréquence :

$$\left[\underbrace{f - Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}}_{\text{Limite inférieure de l'IC}} ; \underbrace{f + Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}}_{\text{Limite supérieure de l'IC}} \right]$$

Ou ES=erreur standard = $\sqrt{\frac{f(1-f)}{n}}$ ou f=fréquence relative ; n=taille de l'échantillon

Alors, pour calculer l' erreur standard :

F	G
Tableau 1	
Fréquence relative de la durée REDUITE de sommeil	0.350806452
Erreur standard de la fréquence (ES)	=sqrt(G2*(1-G2)/248)
Borne inférieure de l'IC	
Borne supérieure de l'IC	
95% IC	
Interpretation de l'IC	
Row Labels	Count of Duree de sommeil
optimale	161
reduite	87
Grand Total	248

Les bornes de l'intervalle de confiance sont les suivantes :

F	G
Tableau 1	
Frequence relative de la duree REDUITE de sommeil	0.350806452
Erreur standard de la frequence (ES)	0.030303657
Borne inferieure de l'IC	=G2-1.96*G3
Borne superieure de l'IC	

F	G
Tableau 1	
Frequence relative de la duree REDUITE de sommeil	0.350806452
Erreur standard de la frequence (ES)	0.030303657
Borne inferieure de l'IC	0.291411284
Borne superieure de l'IC	=G2+1.96*G3

F	G
Tableau 1	
Frequence relative de la duree REDUITE de sommeil	0.350806452
Erreur standard de la frequence (ES)	0.030303657
Borne inferieure de l'IC	0.291411284
Borne superieure de l'IC	0.41020162
95% IC	=

The screenshot shows the 'Insert Function' dialog box. The search field is empty, and the category is set to 'All'. The function list includes: COMPLEX, CONCAT, CONCATENATE (highlighted), CONFIDENCE, CONFIDENCE.NORM, CONFIDENCE.T, and CONVERT.

Calcul de **L'INTERVALLE DE CONFIANCE DE LA MOYENE** par la fonction **CONFIDENCE.T** (fr. **INTERVALLE.CONFIANCE.STUDENT**)

- Choisir la **page Longueur de l'IC** et la cellule **E4** :

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	=	faire la soustraction entre la moyenn et le ME	faire l'addition entre la moyenn et le ME	utiliser la fonction CONCATENATE
248	24.89	0.05	3.65	ut			onction ONCATENATE
248	24.89	0.05	3.65	ut			onction ONCATENATE
248	24.89	0.01	3.65	ut			onction ONCATENATE
si l							
248	24.89	0.05	3.65	ut			onction ONCATENATE
248	24.89	0.05	1.5	ut			onction ONCATENATE

Insert Function

Search for a function:

Type a brief description of what you want to do and then click Go

Go

Or select a category: Most Recently Used

Select a function:

- CONCATENATE
- CONFIDENCE.T**
- NORM.DIST
- AVERAGE
- COUNT
- COUNTIF
- MEDIAN

CONFIDENCE.T(alpha,standard_dev,size)
Returns the confidence interval for a population mean, using a Student's T distribution.

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	=CONFIDENCE.T(C4,D4,A4)	faire la soustraction entre la moyenn et le ME	faire l'addition entre la moyenn et le ME	utiliser la fonction CONCATENATE
248	24.89	0.05	3				
248	24.89	0.05	3				
248	24.89	0.01	3				
248	24.89	0.05	3				
248	24.89	0.05	3				

Function Arguments

CONFIDENCE.T

Alpha C4 = 0.05

Standard_dev D4 = 3.65

Size A4 = 25

= 1.50664595

Returns the confidence interval for a population mean, using a Student's T distribution.

Size is the sample size.

Formula result = 1.50664595

Help on this function

OK Cancel

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	1.50664595	=B4-E4	faire l'addition entre la moyenn et le ME	utiliser la fonction CONCATENATE

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	1.50664595	23.3833541	=B4+E4	utiliser la fonction CONCATENATE

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	1.50664595	23.3833541	26.396646	=CONCATENATE(round(F4,2),"-",round(G4,2))
248	24.89	0.05	3.65				
si α varie							
248	24.89	0.05	3.65	utiliser fonction CONFID			
248	24.89	0.01	3.65	utiliser fonction CONFID			
si l'écart-type (S) varie							
248	24.89	0.05	3.65	utiliser fonction CONFID			

Function Arguments

CONCATENATE

Text1	round(F4,2)	= "23.38"
Text2	"-"	= "-"
Text3	round(G4,2)	= "26.4"
Text4		= text

= "23.38-26.4"

	A	B	C	D	E	F	G	H
1	TABLEAU 1							
2	Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
3	si la taille de l'échantillon varie							
4	25	24.89	0.05	3.65	1.50664595	23.3833541	26.396646	23.38-26.39

- Utiliser l'option HOME → FILL → DOWN pour copier les mêmes formules en bas :

TABLEAU 1							
Taille de l'échantillon (n)	Moyenne (m) de l'IMC	Alpha (α)	S	ME (marge d'erreur)	Borne inférieure de l'IC	Borne supérieure de l'IC	IC
si la taille de l'échantillon varie							
25	24.89	0.05	3.65	1.50664595	23.3833541	26.396646	23.38-26.39
248	24.89	0.05	3.65	0.456507916	24.4334921	25.346508	24.43-25.35

ESTIMATION PONCTUELLE DE LA DIFFERENCE DE MOYENNES (ECHANTILLONS APPARIES) ET PAR L'INTERVALLE DE CONFIANCE

- On va définir la nouvelle variable:

	A	B	C	D	E	F
1	Id_patient	Age(ans)	Sexe	VS_admission(mm)	VS_sortie (mm)	DifférencesVS
2	1	66	M	10.40	14.69	-4.29
3	2	46	F	12.48	17.95	
4	3	57	M	15.34	15.32	
5	4	61	M	34.31	10.81	
6	5	52	M	26.05	17.00	
7	6	28	F	39.35	10.58	
8	7	57	M	30.76	12.38	
9	8	63	M	24.87	8.74	
10	9	43	M	21.80	15.86	
11	10	55	M	35.94	11.87	
12	11	67	M	17.40	9.31	
13	12	51	M	50.02	4.71	
14	13	54	M	11.49	11.99	
15	14	69	F	26.58	9.26	
16	15	51	F	28.19	4.58	
17	16	43	F	17.14	20.19	
18	17	53	F	22.87	12.09	
19	18	59	F	28.38	7.20	

	A	B	C	D	E	F
1	Id_patient	Age(ans)	Sexe	VS_admission(mm)	VS_sortie (mm)	DifférencesVS
2	1	66	M	10.40	14.69	-4.29
3	2	46	F	12.48	17.95	-5.47
4	3	57	M	15.34	15.32	0.01
5	4	61	M	34.31	10.81	23.50
6	5	52	M	26.05	17.00	9.05
7	6	28	F	39.35	10.58	28.77
8	7	57	M	30.76	12.38	18.37
9	8	63	M	24.87	8.74	16.14
10	9	43	M	21.80	15.86	5.95
11	10	55	M	35.94	11.87	24.07
12	11	67	M	17.40	9.31	8.09
13	12	51	M	50.02	4.71	45.31
14	13	54	M	11.49	11.99	-0.50
15	14	69	F	26.58	9.26	17.32
16	15	51	F	28.19	4.58	23.61
17	16	43	F	17.14	20.19	-3.06
18	17	53	F	22.87	12.09	10.78
19	18	59	F	28.38	7.20	21.18

- On va utiliser l'option DATA ANALYSIS:

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:

J	K
<i>DifférencesVS</i>	
Mean	11.52359626
Standard Error	0.990358298
Median	11.01469528
Mode	#N/A
Standard Deviation	10.8488316
Sample Variance	117.6971471
Kurtosis	0.038631196
Skewness	0.314473857
Range	59.75078921
Minimum	-14.44145604
Maximum	45.30933316
Sum	1382.831551
Count	120
Confidence Level(95.0%)	1.961008344
Borne inferieure de l'IC	=K3-K16
Borne superieure de l'IC	
95% IC pour difference de moyennes	

J	K
<i>Différences VS</i>	
Mean	11.52359626
Standard Error	0.990358298
Median	11.01469528
Mode	#N/A
Standard Deviation	10.8488316
Sample Variance	117.6971471
Kurtosis	0.038631196
Skewness	0.314473857
Range	59.75078921
Minimum	-14.44145604
Maximum	45.30933316
Sum	1382.831551
Count	120
Confidence Level(95.0%)	1.961008344
Borne inferieure de l'IC	9.562587915
Borne superieure de l'IC	=K3+K16
95% IC pour difference de moyennes	

Borne inferieure de l'IC	9.562587915
Borne superieure de l'IC	13.4846046
95% IC pour difference de moyennes	9.56-13.48
Interpretation de l'IC	

Borne inferieure de l'IC	9.562587915
Borne superieure de l'IC	13.4846046
95% IC pour difference de moyennes	9.56-13.48
Interpretation de l'IC	Nous sommes sûrs à 95% que l'intervalle [9,56; 13,48] contient la différence de moyennes de VS. De plus, l'IC à 95% ne contient pas 0 → la différence significative de moyennes de VS a l'admission et a la sortie de l'hôpital