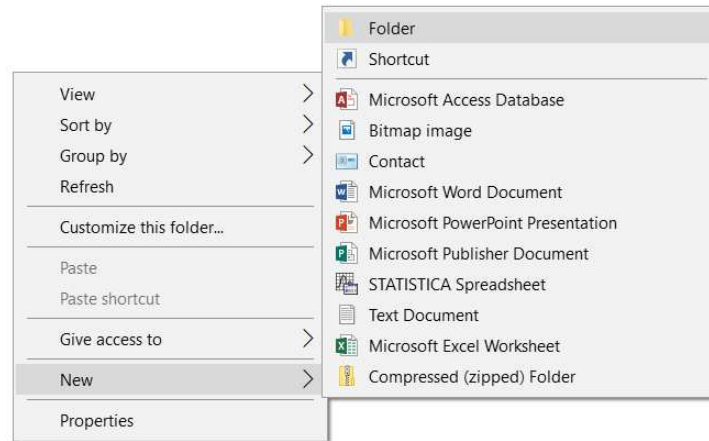


Creation des dossiers:

- Créer un nouveau dossier TP05_NP, ou N=Votre nom et P=votre prenom.
- vous créez un nouveau dossier nome TP05_NP par l'option New→Folder (voir l'image suivante)



L'enregistrement d'un fichier Excel

- File→Save As→chercher une place convenable sur votre ordinateur → dossier TP05NP

Inserez une nouvelle feuille de calcul

- Clic droit sur le nom implicite de la feuille de calcul (en bas de la feuille de calcul, vous aurez le nom Sheet 1=Feuille1) + Insert

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
Amplitude	Maximum - minimum	
Variance d'échantillonnage	VAR	VAR
Déviati on standard/écart type	STDEV	ECARTYPE
Erreur standard	$ES = \frac{s}{\sqrt{n}}$ ou s = la déviation standard, et n est la taille de l'échantillon	Pas de fonction predefinie, on applique la formule
Coefficient de variation	=déviati on standard/moyenne (STDEV/AVERAGE)	
Minimum	MIN	MIN
Quartile 1	= QUARTILE (plage de valeurs,1)	= QUARTILE (plage de valeurs 1)
Quartile 2 (= médiane)	= QUARTILE (plage de valeurs,2)	= QUARTILE (i plage de valeurs,2)
Quartile 3	= QUARTILE (intervalle,3)	= QUARTILE (intervalle,3)
Maximum	MAX	MAX

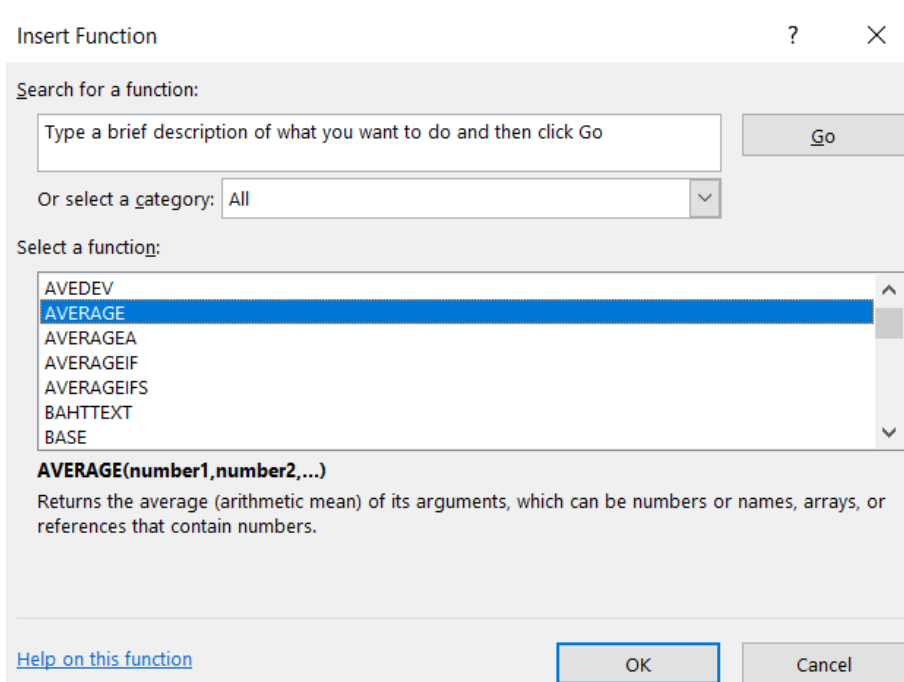
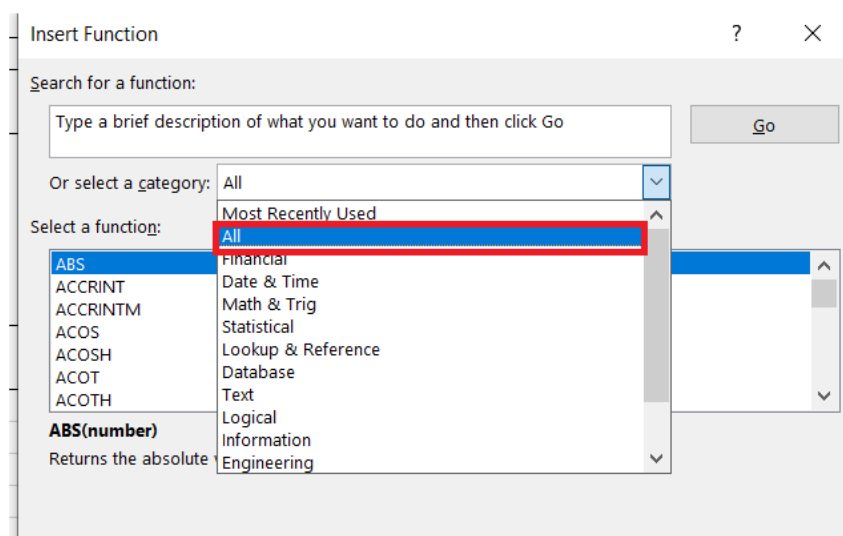
Percentile	= PERCENTILE (intervalle, percentile)	= CENTILE (intervalle, percentile)
Intervalle interquartile	La différence entre le quartile 3 et le quartile 1	
Coefficient d'aplatissement	KURT	KURTOSIS
Coefficient d'asymétrie	SKEW	COEFFICIENT.ASYMETRIE
Si	IF (condition, valeur si condition est vraie, valeur s'il est fausse)	SI
Ou	OR (condition 1, condition 2, ...)	OU
Et	AND (condition 1, condition 2, ...)	ET
Compter	COUNT – compter le nombre des cellules avec des valeurs numériques – pour trouver la taille de l'échantillon	COMPTE
Compter si	COUNTIF (intervalle des cellules, critère pour compter)	NB.SI
La somme	SUM	SOMME
La racine carrée	SQRT	RACINE

Exemple du calcul de la moyenne :

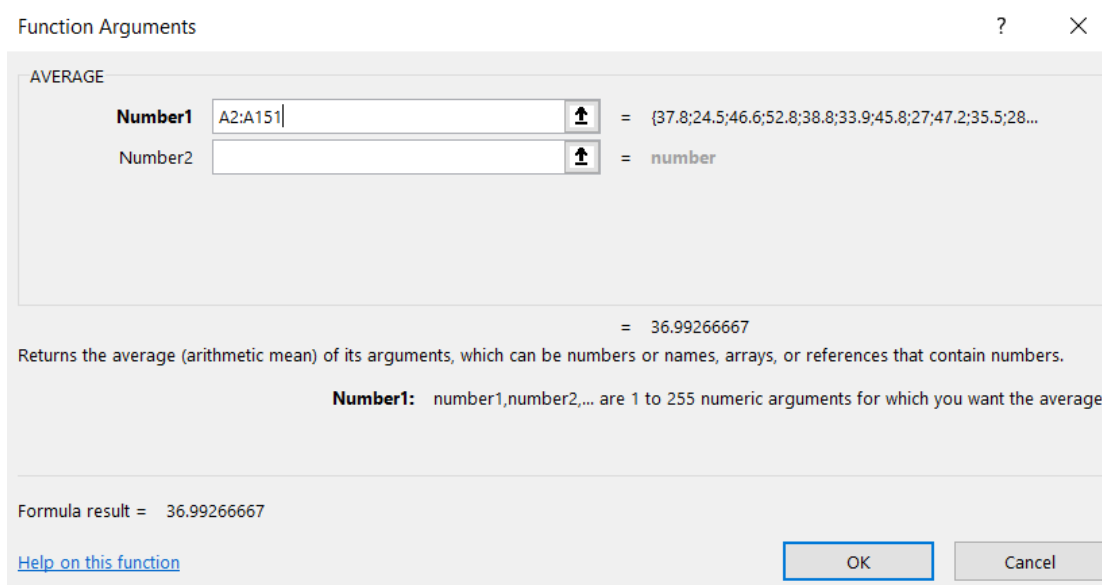
- Sélectionner la cellule du tableau contenant les mesures descriptives pour la variable Age qui attend le résultat de la moyenne (dans la cellule G4)

The screenshot shows the Excel interface with the 'Formulas' tab selected. The 'Insert Function' dialog box is open, showing a list of functions. The 'AVERAGE' function is highlighted. The background spreadsheet shows a table with 'Age(ans)' in column A and various statistical measures in columns E and F. The cell G4 is highlighted, indicating where the average will be calculated.

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

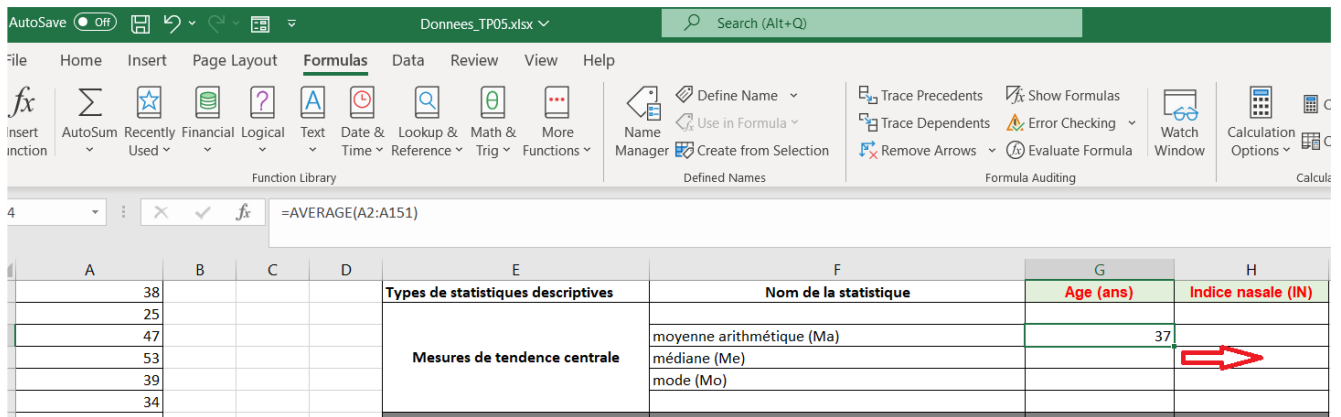


○ **Sélectionner tous les scores de la Variable AGE**

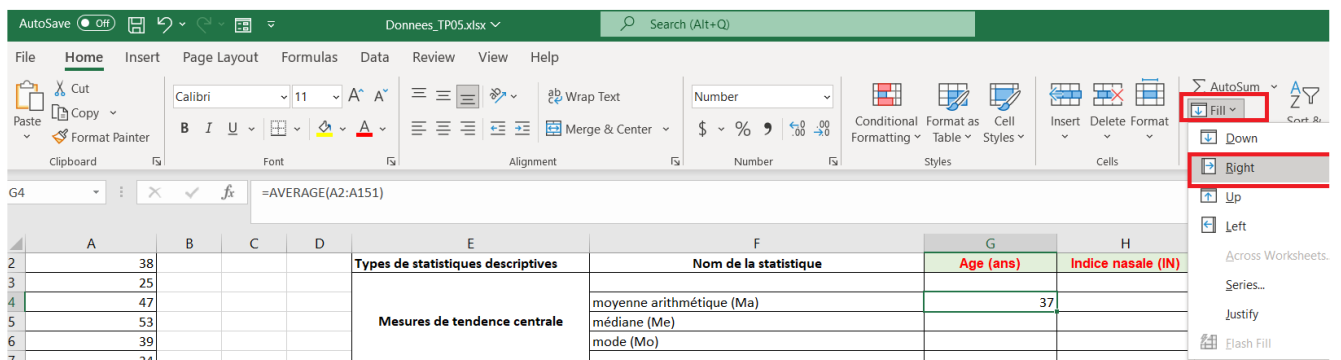


○ L'âge moyen sera égale a 36.9 ans.

○ !!! Pour calculer la moyenne de l'IN on va on va généraliser la formule précédente :



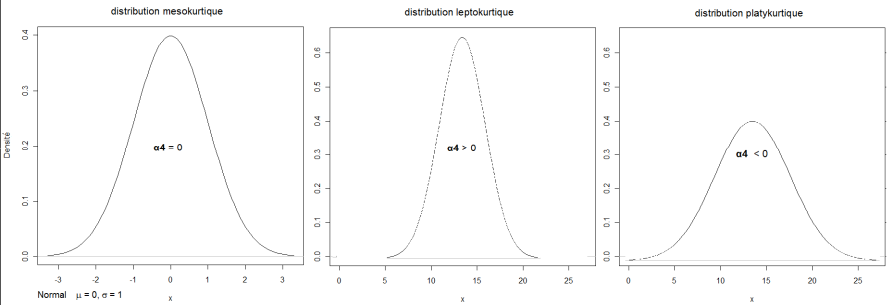
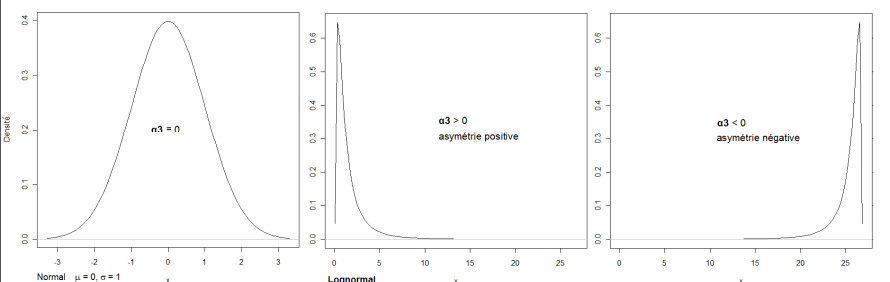
Ou a l'aide de l'option FILL→ RIGHT



Interpretation des medurtes de tendance centrale, emplacement, dispersion...

- **Les mesures de tendance centrale sont les suivants:** la moyenne, la médiane, le mode, la moyenne géométrique, la moyenne harmonique, la valeur la centrale.
- **Les mesures de l'emplacement sont:** les quartiles, les percentiles, les déciles
- **Les mesures de dispersion:** Amplitude, Intervalle interquartile, la variance, l'écart type, le coefficient de variation
- L'asymétrie – skewness (=coefficient d'asymetrie).
- L'aplatissement - kurtosis (=coefficient d'aplatissement).

Nom de la statistique descriptive	Interprétation
Moyenne arithmétique	L'observation au milieu des données si la distribution des données est normale
Médian	L'observation qui divise la distribution sur deux partis égaux ; La moyenne est approximative égale a la médiane si la distribution est normale
Mode	La valeur de l'observation associée à la fréquence la plus élevée. Si la fonctionne donne comme résultat NA ça signifie qu'il n y a pas une mode. Parfois on peut avoir des distribution avec plusieurs valeurs modales. La fonctionne va offrir comme résultat la première valeur modale.
Amplitude	La différence entre la plus grande et la plus petite valeur dans la série des données
Intervalle interquartile	La différence entre le quartile 3 et le quartile 1 dans la série des données.

Écart type (déviati�n standard)	<p>Est la racine carr�e de la variance. Montre la dispersion des donn�es autour de la moyenne.</p> <p>Pour une s�rie des donn�es avec une distribution normale (gaussienne):</p> <p>Dans l'intervalle $m \pm 1$ DS on trouve $\sim 68\%$ de la population.</p> <p>Dans l'intervalle $m \pm 2$ DS on trouve $\sim 95\%$ de la population.</p> <p>Dans l'intervalle $m \pm 3$ DS on trouve $99,7\%$ de la population.</p>
Erreur standard	<p>Est utilis�e dans le calcul de l'intervalle de confiance de 95% autour de la moyenne (pour une distribution normale), aussi est impliqu� dans l'inf�rence statistique.</p>
Quartiles	<p>Il y a trois quartiles: Q1, Q2, Q3</p> <p>le 1er quartile s�pare les 25% inf�rieurs des donn�es</p> <p>25% des valeurs sont $\leq Q1$, 75% sont $\geq Q1$</p> <p>le 2e quartile est la m�diane de la s�rie (50%)</p> <p>50% des valeurs sont $\leq Q2$/ m�diane, 50% sont $\geq Q2$/ m�diane</p> <p>le 3e quartile s�pare les 75% inf�rieurs des donn�es</p> <p>75% des valeurs sont $\leq Q3$, 25% sont $\geq Q3$</p>
Percentiles	<p>le 1er percentile s�pare le 1% inf�rieurs des donn�es;</p> <p>le 95e percentile s�pare les 95% inf�rieurs des donn�es;</p>
Coefficient d'aplatissement Kurtosis	<p>un coefficient d'aplatissement positif (>0) correspond � une distribution <i>pointue (leptokurtique)</i> et un coefficient d'aplatissement n�gatif (<0) correspond � une distribution <i>aplatie (platikurtique)</i>, proche � 0 [-1, 1] indique une distribution <i>mesokurtique</i>. Si la valeur appartient � l'intervalle [-1, 1] pourrait �tre consid�r� que les donn�es ont une distribution mesokurtique.</p> <div style="display: flex; justify-content: space-around;">  </div>
Coefficient d'asym�trie Skewness	<p>Un coefficient positif (>0) indique une distribution �tal�e vers la droite ;</p> <p>Un coefficient n�gatif (<0) indique une distribution �tal�e vers la gauche ;</p> <p>Si la valeur appartient � l'intervalle [-1, 1] pourrait �tre consid�r� que les donn�es ont une distribution sym�trique (du point de vue de l'asym�trie). Sinon la distribution est asym�trique</p> <p>$\alpha_3 \in [-0,5 ; 0,5]$ approximative sym�trique</p> <p>$\alpha_3 \in [-1 ; -0,5]$ ou $[0,5 ; +1]$ mod�r�e asym�trique</p> <p>$\alpha_3 < -1$ ou > 1 asym�trie importante</p> <div style="display: flex; justify-content: space-around;">  </div>

Coefficient of variation

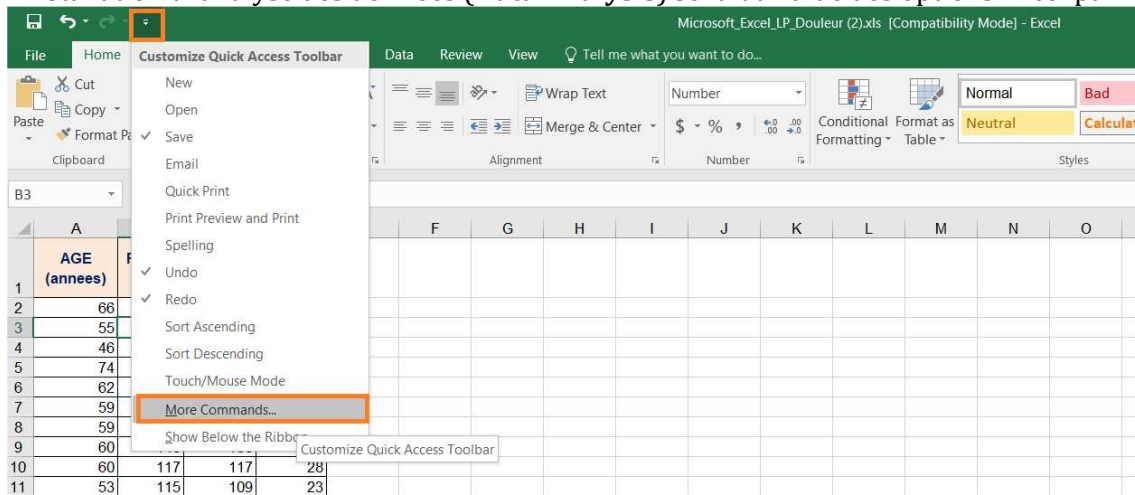
Pour l'évaluation de l'homogénéité d'une série de données
 Si **CV < 10%** la population/l'échantillon peut être considérée comme homogène;
 Si **CV est comprise entre $\geq 10\%$ - $< 20\%$** la population peut être considérée comme relativement homogène;
 Si **CV est entre $\geq 20\%$ à $< 30\%$** la population peut être considérée comme relativement hétérogène;
 Si **CV $\geq 30\%$** la population peut être considérée comme hétérogène.

L'évaluation de symétrie des données en utilisant les indicateurs de tendance centrale, et d'asymétrie.

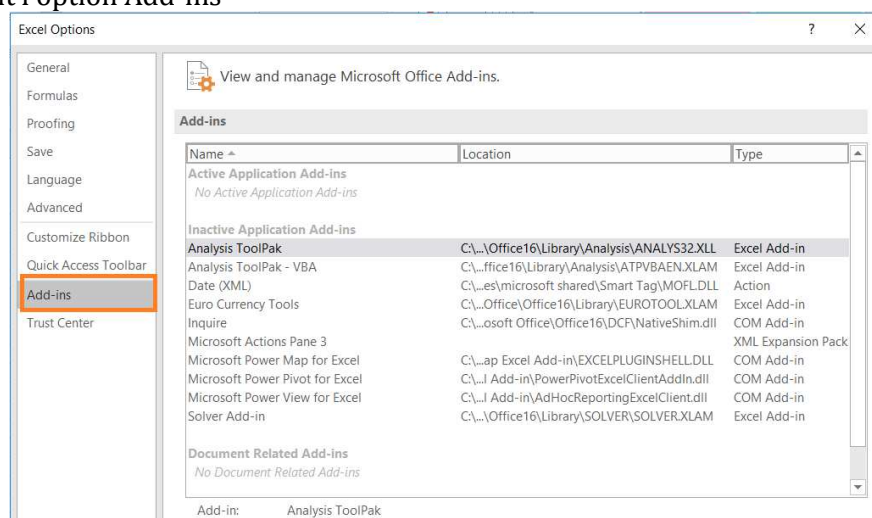
Indications de la symétrie : la moyenne est approximative égale à la médiane et le mode. Le coefficient d'aplatissement et le coefficient d'asymétrie (engl. skewness) sont proches à 0 (dans l'intervalle [-1 ; 1]).

Installation du module Data Analysis pour calculer les statistiques descriptives

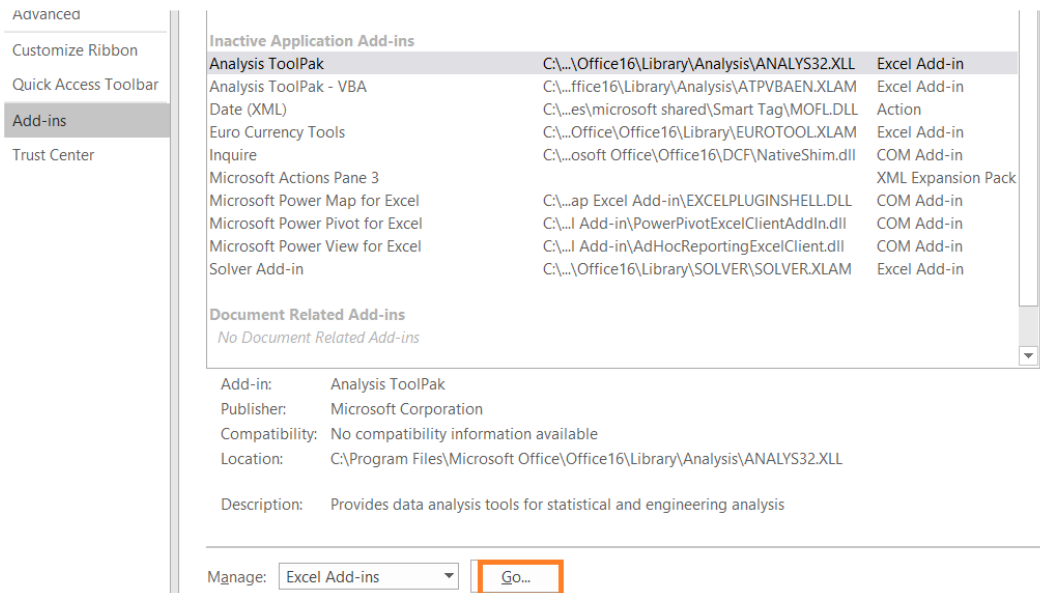
- l'installation d'analyse des données (**Data Analysis**) se fait à l'aide des options Excel par :



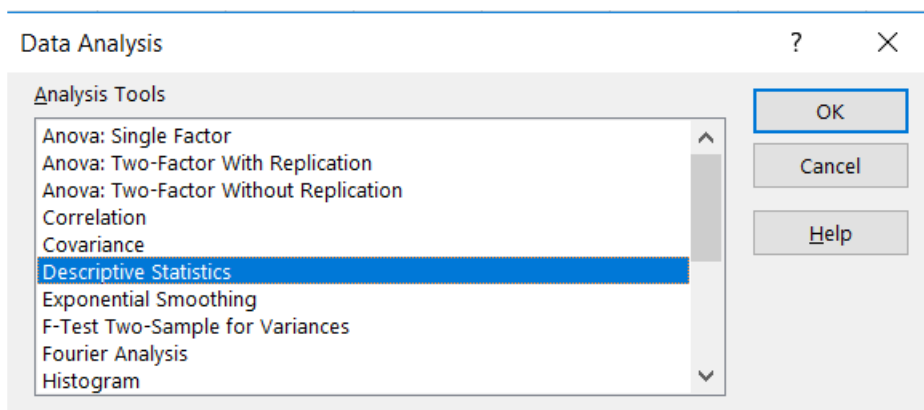
- on choisit l'option Add-ins



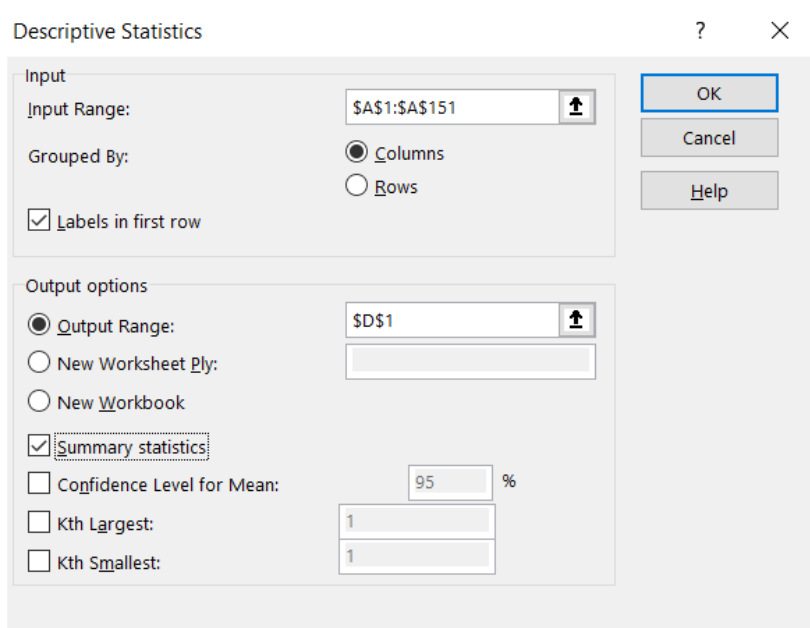
- on clic sur Go:



- Dans la fenetre, cochez l'option: Analysis ToolPak-OK;
- dans le tab **Data** se trouve le **Data Analysis** ;
- pour calculer les indicateurs descriptives on va utiliser l'option *Descriptives statistics* trouvé dans le *Data Analysis*

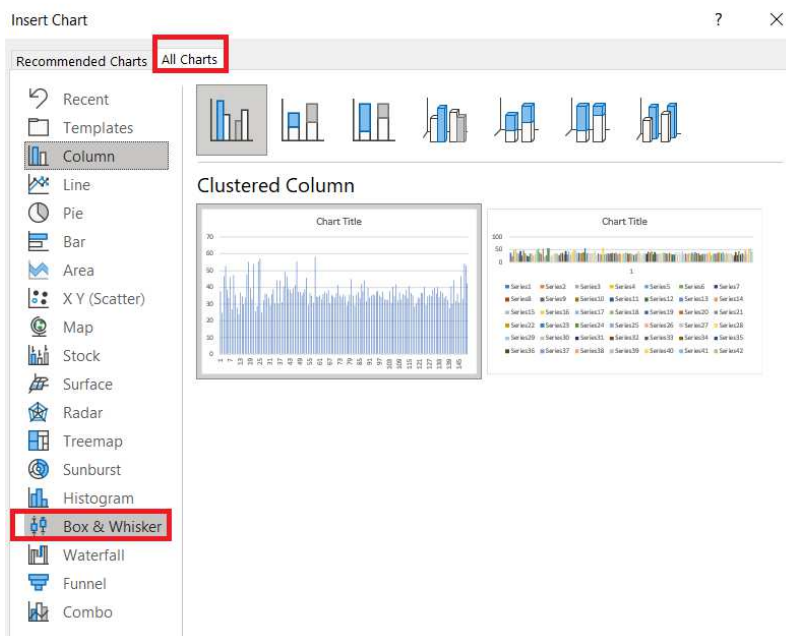


- On va remplir l'option *Input range* avec toutes les cellules correspondantes a la variable quantitative (dans notre cas : l'IN)
- Output range : la cellule désiré pour le résultat
- On choisit les *summary statistics*=les statistiques (indicateurs) descriptives

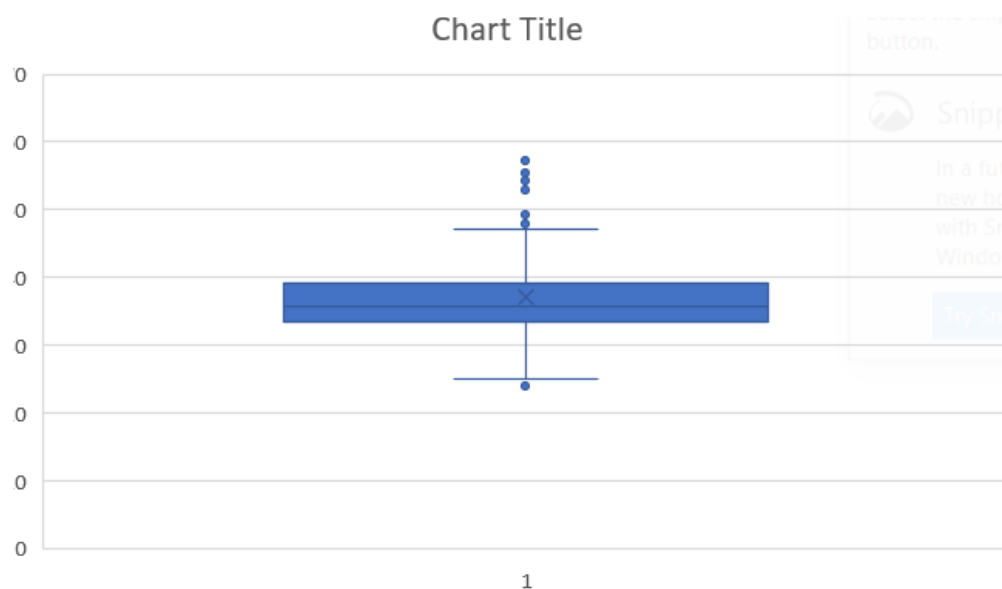


REALISATION D'UN GRAPHIQUE DE TYPE BOITE A DES MOUSTACHES (engl. box-whisker plot)

- On va selectionner les valeurs de la variable d'interet (les valeurs de l'age)
- On va utiliser l'option INSERT→ CHART→All charts



- Le graphique obtenue va avoir la forme suivante:



- A l'aide de l'option CHART DESIGN, realiser son formatage:

