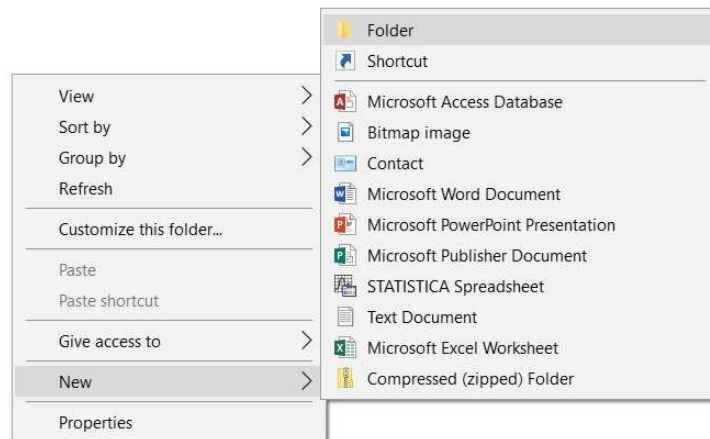


**N'oubliez que:****Creation des dossiers:**

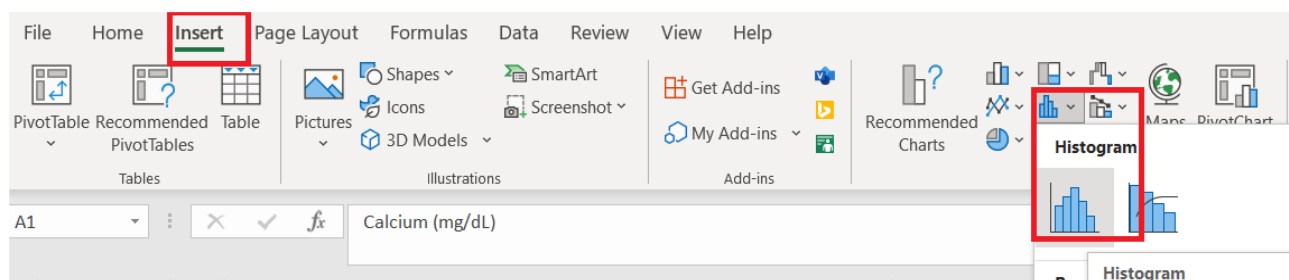
- créer le dossier nommé **TP07NP sur le bureau (Desktop) de l'ordinateur.**
- Pour créer un nouveau dossier (ex. TP07), on utilise l'option New→Folder (voir l'image suivante)

**L'enregistrement d'un fichier Excel et Renommer un fichier Excel**

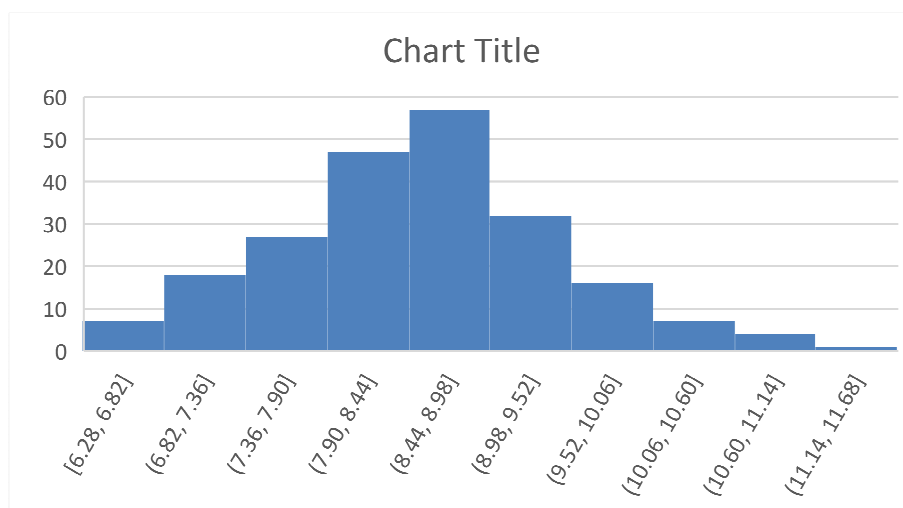
- File→Save As→ dossier TP07NP

**REALISER UN HISTOGRAMME**

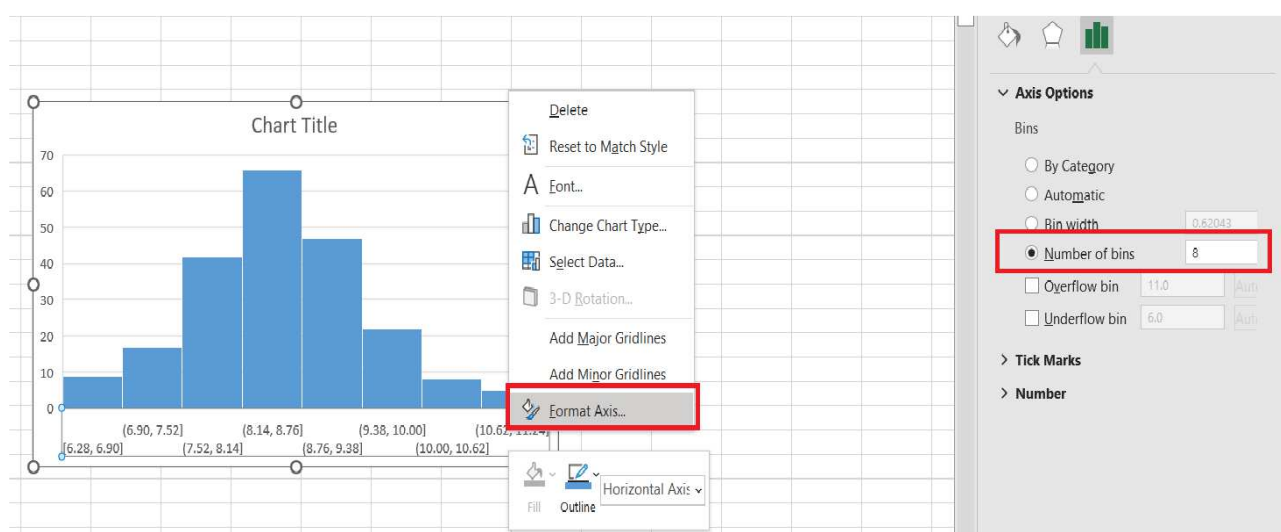
- Sélectionner les valeurs de la variable d'intérêt (dans notre cas, Calcium)
- Choisir l'option INSERT→ HISTOGRAMME (voir l'image)



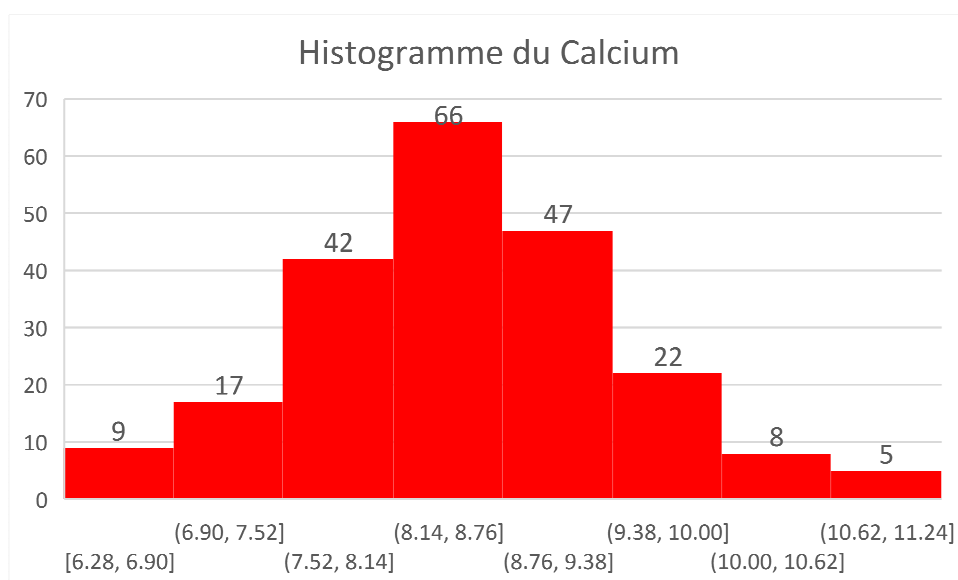
- Le graphique obtenu sera de la forme suivante:



○ Formater-le



○ pour obtenir le graphique suivant:

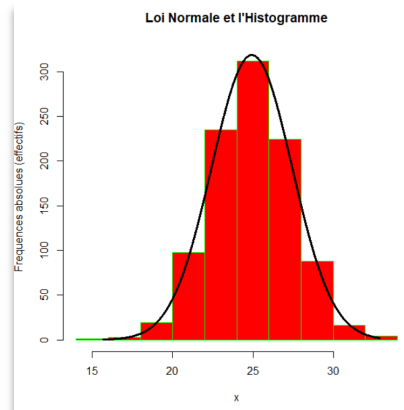


## VERIFIER LA NORMALITE DE DONNEES : Approches empiriques (statistiques descriptives) et graphiques

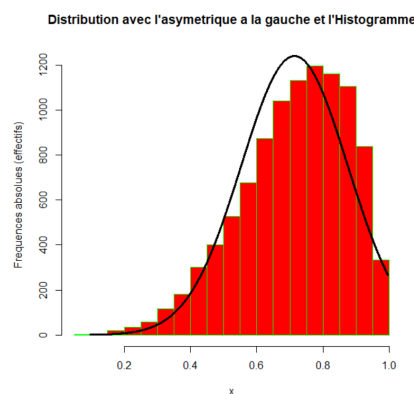
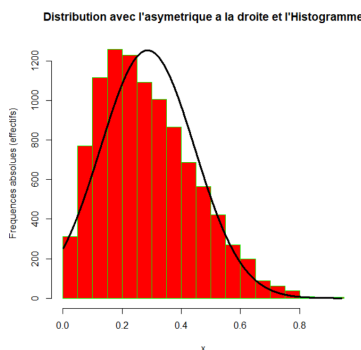
**Normalité de données** : **i)** la moyenne est approximative égale à la médiane (et le mode) ; **ii)** Le coefficient d'aplatissement est proche à **0** (dans l'intervalle **[-1,1]**) et **iii)** le coefficient d'asymétrie est proche à **0** (dans l'intervalle **[-1,1]**).

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**.



## CALCUL DES DIFFERENTS PROBABILITES EN UTILISANT LA LOI NORMALE

- Pour calculer le pourcentage de sujets ayant le taux du Calcium  $\leq 8$  mg/dl, choisir la cellule F19 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 $\text{moyenne} - 2 \cdot \text{DS} \leq$ Calcium $\leq$ $\text{moyenne} + 2 \cdot \text{DS}$ Calcium } | Pr(T) = ?   |   |
|  |             |   |
|  |             |   |
|  |             |   |
|  |             |   |

- la syntaxe de la fonction NORM.DIST sera:

Function Arguments

NORM.DIST

|              |       |                         |
|--------------|-------|-------------------------|
| X            | 8     | la valeur donnée        |
| Mean         | 8.52  | la moyenne du calcium   |
| Standard_dev | 0.929 | l'ecart-type du calcium |
| Cumulative   | TRUE  |                         |

= 0.287827832

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 = 0.287827832

[Help on this function](#)

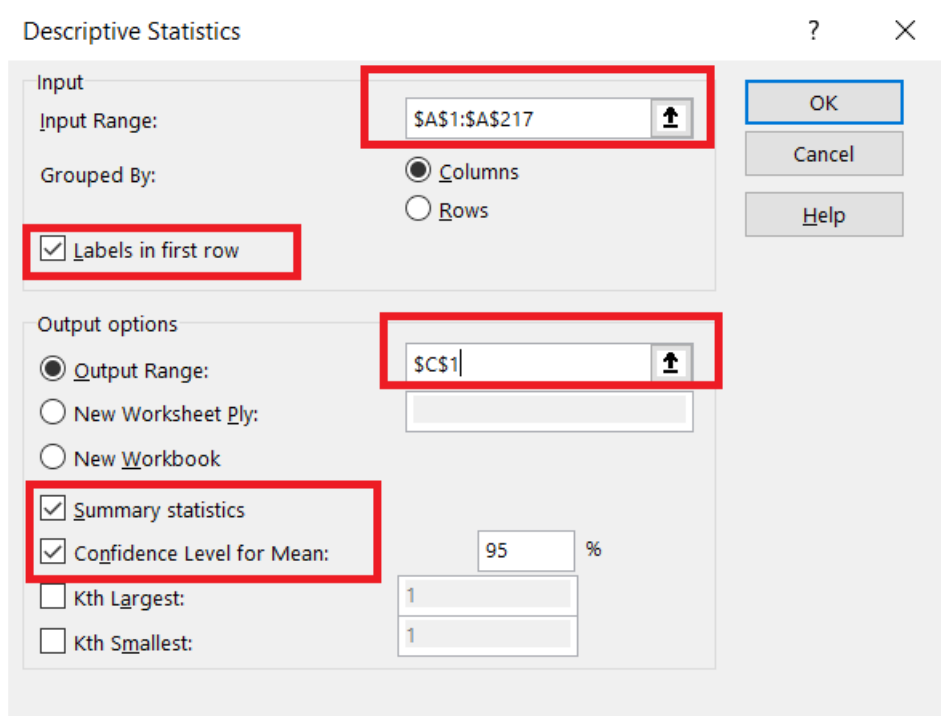
OK Cancel

- Pour calculer le pourcentage de sujets ayant le taux du  $8 \leq$  Calcium  $\leq$  10 mg/dl, choisir la cellule F21 du Tableau 2 de la page Statistique descriptive et puis faites la différence entre les deux probabilités antérieurement calculés :

|  |           |          |
|--|-----------|----------|
| P = {avoir $8 \leq$ Calcium $\leq$ 10} | Pr(P) = ? | =F20-F19 |
|--|-----------|----------|

## Travail sur l'Excel: calcul de L'INTERVALLE DE CONFIANCE DE LA MOYENNE

- Dans la page nommée **ICMoyenne**, on va estimer la moyenne du CALCIUM en utilisant l'option DATA → DATA ANALYSIS → DESCRIPTIVE STATISTICS



2. Les résultats obtenus sont les suivantes :

**Calcium (mg/dL)**

|                                |                    |
|--------------------------------|--------------------|
| <b>Mean</b>                    | <b>8.521718389</b> |
| Standard Error                 | 0.06326082         |
| Median                         | 8.536327549        |
| Mode                           | 9.044080876        |
| Standard Deviation             | 0.929740378        |
| Sample Variance                | 0.86441717         |
| Kurtosis                       | 0.176235695        |
| Skewness                       | 0.172549007        |
| Range                          | 4.963494575        |
| Minimum                        | 6.277911631        |
| Maximum                        | 11.24140621        |
| Sum                            | 1840.691172        |
| Count                          | 216                |
| <b>Confidence Level(95.0%)</b> | <b>0.124690816</b> |

$$t_{\alpha/2, n-1} * \frac{S}{\sqrt{n}}; \text{ ou } \frac{S}{\sqrt{n}} = ES = \text{erreur standard}$$

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

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

Limite inférieure de l'IC

Limite supérieure de l'IC

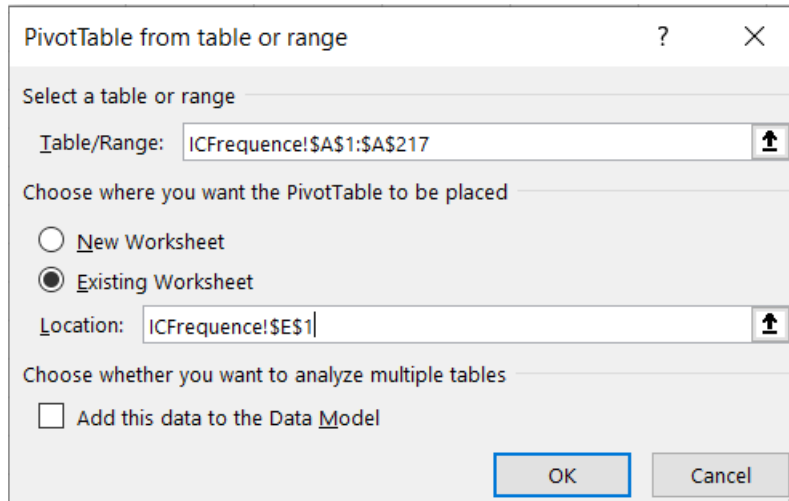
Ou n=taille de l'échantillon ; m=moyenne ; S=déviati n standard

| C                                 | D                  |
|-----------------------------------|--------------------|
| <b>Calcium (mg/dL)</b>            |                    |
| <b>Mean</b>                       | <b>8.521718389</b> |
| Standard Error                    | 0.06326082         |
| Median                            | 8.536327549        |
| Mode                              | 9.044080876        |
| Standard Deviation                | 0.929740378        |
| Sample Variance                   | 0.86441717         |
| Kurtosis                          | 0.176235695        |
| Skewness                          | 0.172549007        |
| Range                             | 4.963494575        |
| Minimum                           | 6.277911631        |
| Maximum                           | 11.24140621        |
| Sum                               | 1840.691172        |
| Count                             | 216                |
| <b>Confidence Level(95.0%)</b>    | <b>0.124690816</b> |
| <b>Limite inferieure de l' IC</b> | <b>=D3-D16</b>     |
| <b>Limite superieure de l'IC</b>  |                    |

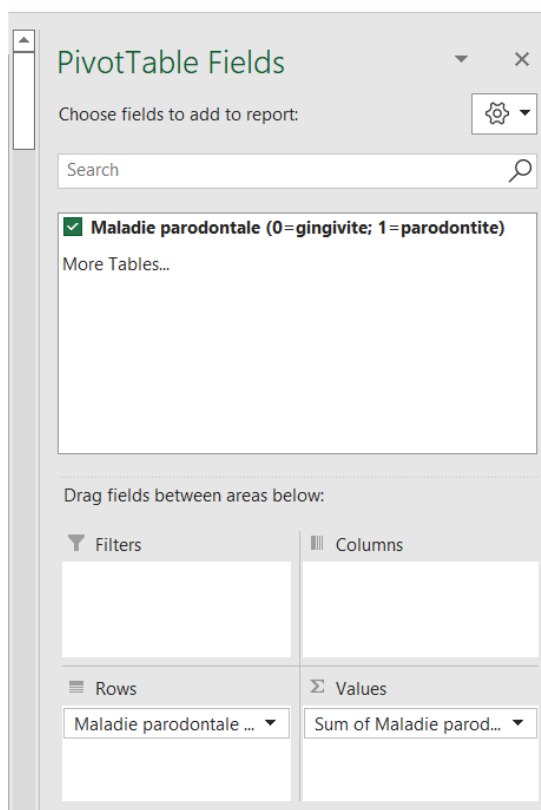
| C                                 | D                  |
|-----------------------------------|--------------------|
| <b>Calcium (mg/dL)</b>            |                    |
| <b>Mean</b>                       | <b>8.521718389</b> |
| Standard Error                    | 0.06326082         |
| Median                            | 8.536327549        |
| Mode                              | 9.044080876        |
| Standard Deviation                | 0.929740378        |
| Sample Variance                   | 0.86441717         |
| Kurtosis                          | 0.176235695        |
| Skewness                          | 0.172549007        |
| Range                             | 4.963494575        |
| Minimum                           | 6.277911631        |
| Maximum                           | 11.24140621        |
| Sum                               | 1840.691172        |
| Count                             | 216                |
| <b>Confidence Level(95.0%)</b>    | <b>0.124690816</b> |
| <b>Limite inferieure de l' IC</b> | <b>8.397027573</b> |
| <b>Limite superieure de l'IC</b>  | <b>=D3+D16</b>     |

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

1. Réalisez un tableau croise (engl.pivot table) pour la variable « Maladie parodontale »
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 :

| E   | F          |
|---|------------|
| Row Labels  |            |
| Count of Maladie parodontale (0=gingivite; 1=parodontite) |            |
| 0   | 137        |
| 1   | 79         |
| <b>Grand Total</b>  | <b>216</b> |

5. Dans le TABLEAU 1 on calcule la fréquence relative de la gingivite (défini comme le rapport entre les patients souffrant de la gingivite et l'effectif total de l'échantillon)

| Tableau 1                            |          |
|--------------------------------------|----------|
| frequence relative de la gingivite   | =137/216 |
| Erreur standard de la frequence (ES) |          |
| Limite inferieure de l'IC            |          |
| Limite superieure de l'IC            |          |

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

95% CI pour la fréquence :

$$\left[ f - Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}; f + Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}} \right]$$

Limite inferieure de l'IC
Limite supérieure de l'IC

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 :

| Tableau 1                            |                       |
|--------------------------------------|-----------------------|
| frequence relative de la gingivite   | 0.634259259           |
| Erreur standard de la frequence (ES) | =sqrt(F8*(1-F8))/216) |
| Limite inferieure de l'IC            |                       |
| Limite superieure de l'IC            |                       |

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

| Tableau 1                            |             |
|--------------------------------------|-------------|
| frequence relative de la gingivite   | 0.634259259 |
| Erreur standard de la frequence (ES) | 0.002229802 |
| Limite inferieure de l'IC            | =F8-1.96*F9 |
| Limite superieure de l'IC            |             |



| Tableau 1                            |             |
|--------------------------------------|-------------|
| frequence relative de la gingivite   | 0.634259259 |
| Erreur standard de la frequence (ES) | 0.002229802 |
| Limite inferieure de l'IC            | 0.629888847 |
| Limite superieure de l'IC            | =F8+1.96*F9 |