

Cours Modélisation et Programmation avec tableur

Eléments de base du tableur Excel

Avertissement : Ce document ne prétend pas être exhaustif et remplacer la documentation d'Excel ou l'aide en ligne; son but est d'aider à la prise en main d'un tableur, de donner quelques conseils qui permettent une utilisation propre et la construction de modèles transmissibles avec Excel, de rappeler les fonctions qui pourront être utiles en statistique et modélisation.

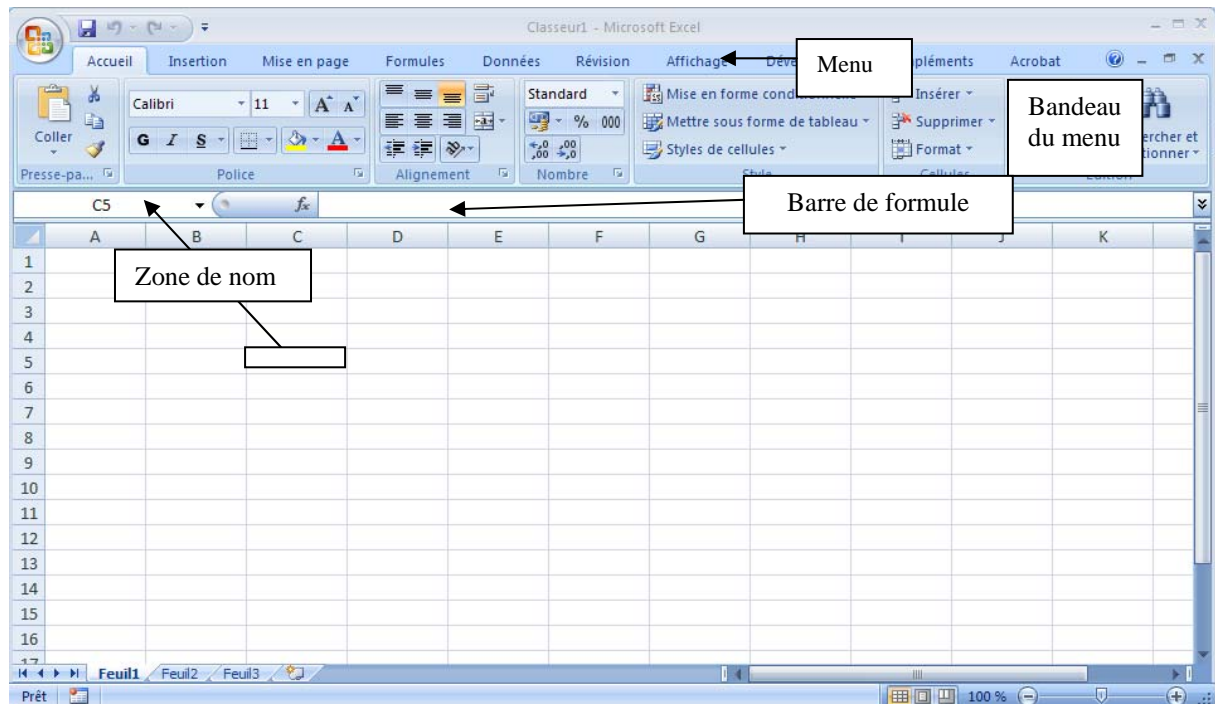
Eléments de base du tableur Excel

Sommaire

1	<u>RECOPIE DE CELLULES ET ADRESSAGE</u>	4
2	<u>QUELQUES REGLES D'UTILISATION D'UN TABLEUR</u>	5
3	<u>QUELQUES ELEMENTS DE BASE POUR L'UTILISATION D'EXCEL</u>	5
3.1	LES SELECTIONS DE CELLULES	5
3.2	LA FONCTION SOMME	6
3.3	LES FORMATS	6
3.4	LES TABLES	7
3.4.1	TABLE A UN PARAMETRE	7
3.4.2	TABLE A 2 PARAMETRES	8
3.5	LES GRAPHIQUES	8
4	<u>LES FONCTIONS</u>	9
4.1	FONCTIONS DE RECHERCHE DANS LES MATRICES : MAX, MIN, INDEX, EQUIV	9
4.2	LES AFFECTATIONS CONDITIONNELLES	10
4.2.1	FONCTION SI	10
4.2.2	FONCTION RECHERCHE	10
4.2.3	FONCTION RECHERCHEV	11
5	<u>LES FONCTIONS MATRICIELLES</u>	12
6	<u>BASES DE DONNEES (LISTES) SOUS EXCEL.</u>	13
6.1	STRUCTURE D'UNE LISTE DE DONNEES	13
6.2	DEFINITION D'UNE ZONE DE CRITERES	16
6.3	EXTRACTION D'UNE PARTIE DE LA BASE DE DONNEES	17
6.4	FONCTIONS BASE DE DONNEES	18
7	<u>EXEMPLE DE L'UTILISATION D'EXCEL EN STATISTIQUE DESCRIPTIVE</u>	19

Eléments de base du tableur Excel

Une fois lancée l'application Excel, apparaît à l'écran une fenêtre dédiée à l'application et contenant un classeur vierge :



L'application Excel travaille sur des objets, l'objet principal est le **classeur**, lui-même constitué de **feuilles**, les feuilles étant constituées de **lignes** et **colonnes**, elles aussi constituées de **cellules**.

L'objet élémentaire est donc la cellule repérée comme intersection d'une ligne et d'une colonne d'une feuille donnée. A un moment donné, seule une occurrence d'un objet est active, c'est à dire que l'utilisateur peut modifier cet objet : sur l'exemple la feuille active est 'Feuil1', la colonne active est la colonne 'C', la ligne active la ligne '5' et enfin la cellule active est notée 'C5'.

L'utilisateur ne peut modifier qu'un objet actif, il est donc nécessaire avant toute opération de sélectionner l'objet sur lequel on veut travailler.

Une cellule peut contenir du texte, des nombres ou des formules. Une formule est une expression qui fait dépendre la valeur d'une cellule des valeurs prises par d'autres cellules, et si l'utilisateur modifie la valeur d'une de ces cellules la formule est mise à jour et la valeur est donc modifiée.

Pour entrer le contenu d'une cellule :

- Sélectionner cette cellule en cliquant dessus
- Si le contenu est du texte ou un nombre entrer ce texte ou ce nombre, sinon entrer la formule en commençant par le signe =, et cliquer sur les cellules qui figurent dans l'expression, puis valider par la touche <Entrée> ou le bouton de validation.

L'expression du contenu de la cellule apparaît dans la barre de formule, la valeur prise par cette expression apparaît, elle, dans la cellule.

1 RECOPIE DE CELLULES ET ADRESSAGE

Un exemple : supposons que l'on veuille suivre mensuellement les ventes de différents produits. Chaque mois le chiffre d'affaires réalisé par produit sera obtenu en multipliant les quantités vendues par le prix unitaire, on peut présenter ce suivi de la façon suivante :

	B	C	D	E	F
3	Taux de TVA	20,6%			
4					
5	Produit	Quantité	Prix unitaire	C.A. H.T.	T.V.A.
6	Produit 1	156	200	=C6*D6	
7	Produit 2	240	125		
8	Produit 3	45	800		

Les formules que nous aurons à entrer dans les cellules du chiffre d'affaire auront toutes la même structure : ce sera le produit de la quantité (2 colonnes à gauche) par le prix unitaire (1 colonne à gauche). Plutôt que de rentrer individuellement chacune de ces formules, ce qui serait fastidieux s'il y avait une dizaine de produits, on peut n'entrer que la formule pour le premier produit et "recopier" cette formule vers le bas pour les autres produits. Ceci se fait en plaçant le curseur sur le petit carré noir en bas à gauche de la cellule active (CA du premier produit) puis quand le curseur a la forme d'un + en maintenant le bouton gauche de la souris appuyé et en tirant vers le bas suivant le nombre voulu de recopies.

On constate que la recopie "s'est bien passée" la formule en E7 est bien =C7*D7, **par défaut l'adressage est relatif**. C'est à dire que lorsque l'utilisateur rentre une formule, toutes les cellules intervenant dans la formule sont repérées par rapport à la cellule active (qui devient l'origine du tableur), par exemple, pour la formule entrée en E6, D6 n'est pas considérée comme la cellule se trouvant à l'intersection de la colonne D et de la ligne 6 mais comme la cellule se trouvant sur la même ligne et une colonne à gauche.

	B	C	D	E	F
3	Taux de TVA	20,6%			
4					
5	Produit	Quantité	Prix unitaire	C.A. H.T.	T.V.A.
6	Produit 1	156	200	31200	=E6*\$C\$3
7	Produit 2	240	125	30000	
8	Produit 3	45	800	36000	

En revanche si nous voulons procéder de la même façon pour calculer la TVA associée à chaque produit, nous ne pouvons pas entrer en F6 la formule = E6*C3 car en recopiant vers le bas la formule deviendrait en F7 : =E7*C4. Ici la cellule contenant la TVA doit être positionnée en absolue dans la feuille et non plus relativement à la cellule active, on parle alors **d'adressage absolu**. Pour obtenir un adressage absolu sous Windows, quand la cellule C3 est désignée dans la formule, appuyer alors sur *la touche de fonction F4*, il apparaît alors \$C\$3, ce qui est l'indication que la cellule est repérée par rapport à la feuille et non plus par rapport à la cellule active.

Remarques :

1. En appuyant plusieurs fois sur la touche F4 on obtient toutes les combinaisons possibles entre adressage absolu et relatif :
 - Colonne et ligne absolues (\$C\$3)
 - Colonne relative, ligne absolue (C\$3)
 - Colonne absolue, ligne relative (\$C3)
 - Colonne et ligne relatives (C3)

Voir l'exercice n°1 pour une application des différents types d'adressage.

2. Plutôt que d'utiliser l'adressage absolu, il est souvent plus clair d'utiliser un nom. Pour cela après avoir sélectionné la cellule C3 contenant le taux de TVA, taper dans la zone de nom, qui se trouve à gauche de la barre de formule et contient l'indication de la cellule active, un nom (sans blanc, par exemple taux_de_tva) et valider avec la touche <Entrée>. Chaque fois que vous désignerez cette cellule dans une formule son nom apparaîtra et elle sera considérée comme adressée en absolu.

2 QUELQUES REGLES D'UTILISATION D'UN TABLEUR

Un tableur est un outil d'aide à la décision, ce n'est pas une calculette qui permet de résoudre un problème pour un jeu de données et un seul. La bonne utilisation d'un tableur consiste à considérer que c'est un type de problème et non pas un problème particulier qui doit être représenté, modélisé. Pour cela on respectera les règles simples suivantes :

1. Ne jamais utiliser de **constantes** (tels que taux de tva, pourcentage de remise etc..) directement en valeur **dans une formule**, pour cela :
2. Créer dans une partie de la feuille ou sur une autre feuille une **zone de données** (zone de paramètres) du problème – il est bon de nommer chacune des cellules contenant les valeurs des données, ce qui rend le modèle plus lisible et plus facile à contrôler.
3. Dans une autre zone du classeur, regrouper les équations du modèle en détaillant au maximum les intermédiaires de calcul, de façon à éviter des formules trop compliquées difficilement contrôlables.
4. Ne jamais recopier (**copier/coller**) une cellule contenant une valeur numérique, en effet un tableur ne mémorise pas les opérations de copie, et si vous changez la valeur initiale les valeurs recopiées ne seront pas mises à jour. Utiliser une formule telle que '=cellule'

3 QUELQUES ELEMENTS DE BASE POUR L'UTILISATION D'EXCEL


3.1 Les sélections de cellules

Il est très souvent utile de pouvoir sélectionner un ensemble de cellules sur lesquels on effectuera les mêmes transformations. Cette sélection peut toujours être décomposée en une réunion de rectangles.

Pour sélectionner une zone rectangulaire de cellules (que nous appellerons plage), cliquer sur la cellule de l'angle supérieur gauche du rectangle, puis, en maintenant la touche <Maj> enfoncée, cliquer sur le coin inférieur droit du rectangle. La plage est inversée à l'écran sauf pour la cellule active qui est la cellule du coin supérieur gauche. Pour entrer une formule (après avoir bien fait attention à l'adressage relatif ou absolu) dans une sélection, il faut valider avec la combinaison de touche <Ctrl>+<Entrée>, c'est à dire maintenir la touche Ctrl enfoncée en appuyant sur la touche <Entrée>.

Pour ajouter un autre rectangle à une sélection, cliquer le coin supérieur gauche du rectangle à ajouter en maintenant la touche <Ctrl> enfoncée, puis procéder comme précédemment pour déterminer le rectangle.

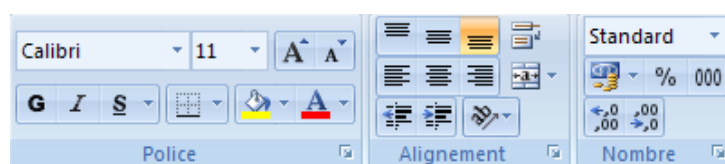
3.2 La fonction somme

Le bouton  du bandeau du menu accueil permet d'opérer des sommations automatiques, cela évite simplement de taper la fonction somme, la cellule active contient alors une formule telle que =**Somme(cellule1:cellule2)** (suivant l'environnement de la cellule active). Si cette plage ne vous convient pas, il vous suffit de sélectionner la plage dont vous voulez sommer le contenu. L'argument cellule1:cellule2 signifie que la somme se fait sur toute la plage rectangulaire ayant cellule1 comme coin supérieur gauche et cellule2 comme coin inférieur droit. En cliquant sur la liste de droite, on peut choisir entre différentes fonctions qui ont la même syntaxe que la fonction **Somme (Moyenne, Min, Max)**

3.3 Les formats

On appelle format toute présentation d'un résultat en informatique, il est important de noter qu'un format ne joue que sur **l'apparence** d'un objet et non sur son **contenu** qui reste inchangé.

Tous les choix peuvent être faits dans le menu format, les modifications portant sur l'objet actif. Pour une cellule l'utilisateur peut modifier la police de caractères, l'alignement du texte, l'affichage des nombres etc.

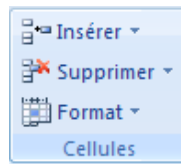


Pour les formats les plus courants, le bandeau d'accueil permet d'effectuer plus rapidement ces opérations :

- Choix de police, de taille, gras, italique, souligné.
- Alignement à gauche, centré, à droite, centrage sur plusieurs cellules.
- Format monétaire, pourcentage, milliers séparés par un blanc, ajout et suppression de décimales, retraits.
- Encadrement et fond de la cellule (bordure et trame) et enfin couleur du texte.

Pour les formats plus élaborés, les menus complets sont donnés par les boutons du bas du bandeau (Police, Alignement, Nombre).

Les autres formats tels que largeur de colonne, hauteur de ligne etc., s'obtiennent en cliquant sur le bouton format du bandeau d'accueil :



3.4 Les tables

Il est souvent utile de connaître un ou plusieurs résultats du modèle pour des valeurs différentes des paramètres (taux d'imposition, prix fixé etc.); il est bien sûr possible de modifier manuellement le contenu de la cellule et de noter les valeurs ainsi obtenues. En procédant ainsi on ne peut malheureusement pas conserver simultanément tous les résultats, chaque nouvelle valeur remplaçant la précédente.

Il est plus efficace d'utiliser alors les tables d'hypothèses, appelées simplement tables qui permettent alors de visualiser l'ensemble des résultats en regard des différentes valeurs des paramètres.

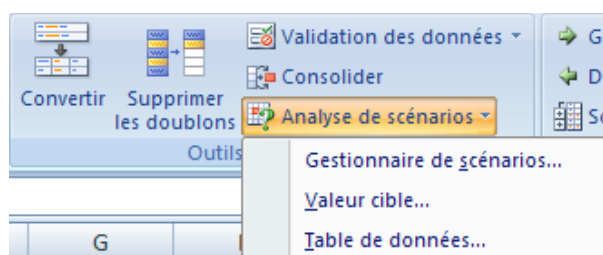
3.4.1 Table à un paramètre

Il s'agit ici de visualiser les valeurs de différentes cellules du modèle en fonction de différentes valeurs d'un paramètre.

La table va être une plage, dont vous alimenterez la première colonne avec les différentes valeurs testées pour le paramètre ; indiquez dans la première ligne quelles sont les cellules dont les résultats doivent être calculés pour ces valeurs :

	=cellule résultat1	=cellule résultat2
Valeur 1 testée		
Valeur 2		
Valeur 3		
Valeur n		

Une fois cette plage définie, il ne reste plus qu'à mettre en relation les valeurs avec la cellule de paramètre du modèle correspondant; ceci se fait par le menu **Table de Données** du bandeau associé au menu **Données / Analyse de scénarios** :



On obtient alors la boîte de dialogue suivante :



Les valeurs du paramètre sont ici en colonne, donc dans Cellule d'entrée en colonne on indiquera **l'adresse dans le modèle** de la cellule contenant le paramètre, puis on validera par OK. Excel recalcule alors le modèle en remplaçant la valeur du paramètre par les valeurs successives et stockera les résultats demandés en face de ces valeurs dans la colonne correspondante.

3.4.2 Table à 2 paramètres

Bien évidemment dans ce cas, étant donnée la nature bidimensionnelle d'une feuille, on ne peut calculer que les valeurs d'un seul résultat en fonction des différentes valeurs de deux paramètres. La table prend alors la forme suivante :

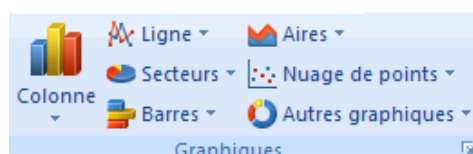
=cel resultat	Val1 Par en ligne	Val2 Par en ligne	Valp Par en ligne
Val1 Par en colonne				
Val2 Par en colonne				
Val3 Par en colonne				
Valn Par en colonne				

La première colonne et la première ligne contiennent les valeurs à tester pour les deux paramètres, la cellule du coin supérieur gauche contient une formule faisant référence à la cellule du modèle dont les résultats doivent être conservés.

En passant par le menu Données-Table, on remplit alors la boîte de dialogue en indiquant comme cellule d'entrée en colonne l'adresse dans le modèle du premier paramètre, et comme cellule d'entrée en ligne l'adresse dans le modèle du second paramètre.

3.5 Les graphiques

La construction d'un graphique se fait, après sélection des plages (en incluant les titres qui serviront de légendes) servant à la construction du graphique, soit en sélectionnant le type de graphique dans le bandeau du menu insertion, soit en utilisant l'assistant graphique correspondant au bouton « Graphiques » du bas (c'est ce que nous ferons ici) :



Une fois le type de graphique choisi ne pas oublier de vérifier qu'Excel affiche le graphique comme un nouvel objet dans la feuille de calcul, il est alors possible de le modifier et de l'enrichir grâce au bandeau « outils graphiques » qui apparaît.



Attention : dans tous les types de graphiques sauf le type nuage de points, les abscisses sont considérées comme des variables qualitatives et donc Excel ne prend pas en compte leur valeur comme nombre, mais la considère comme un intitulé même si elle est numérique. Par conséquent dans un graphique de type Courbes, les valeurs en abscisses seront équidistantes et dans l'ordre de la ligne ou colonne où elles se trouvent.

4 LES FONCTIONS

Quand l'utilisateur tape le signe = dans la barre de formule la zone de nom est remplacée par une liste déroulante de fonctions regroupant les dix dernières fonctions utilisées et "Autres fonctions ..." qui appelle une boîte de dialogue de l'assistant fonctionnel. Chaque fonction possède une aide en ligne appelée par le ?. Nous ne parlerons ici que des fonctions de recherche dans les matrices et des fonctions d'affectation de valeurs conditionnelles (la fonction **SI** et les fonctions **RECHERCHE***), laissant au lecteur le soin de découvrir en fonction de ses besoins les fonctions qui lui seront utiles.

De façon générale une fonction sous Excel a un nom suivi de la liste des arguments entre parenthèses, chacun des éléments de la liste étant séparé du suivant par un ";" en environnement français.

Le nom de la fonction est toujours traduit en majuscule par Excel, si vous entrez le nom sans passer par l'assistant fonctionnel il est de bonne pratique de l'écrire en minuscule pour déceler les fautes de frappe éventuelles.

4.1 Fonctions de recherche dans les matrices : Max, Min, Index, Equiv

Les fonctions **Max(zone1 ; zone2 ;...)** et **Min(zone1 ; zone2 ;...)** retournent respectivement le plus grand et le plus petit des éléments de la réunion des différentes zones.

La fonction **Index(matrice; ligne; colonne)** retourne l'élément qui se trouve à l'intersection de la ligne numéro ligne et de la colonne numéro colonne d'une plage nommée matrice ou de l'adresse de cette plage. Exemple :

	C	D	E	F
7	2	3	5	13
8	23	45	12	22
9	0	2	1	5

=Index(C7:F9;2;3) retourne la valeur 12

La fonction **Equiv(valeur;matrice;0)** trouve la position d'une valeur dans une matrice ligne ou colonne, c'est à dire ne contenant *qu'une ligne* ou *qu'une colonne*. Avec le tableau précédent :

=Equiv(12;C8:F8;0) retourne la valeur 3 car 12 occupe la 3^{ème} position dans la plage C8:F8

=Equiv(Max(E7:E9);E7:E9;0) retourne la valeur 2, car le maximum de la plage E7:E9 est 12 qui est en deuxième position dans la colonne E7:E9

Remarque : la fonction **Equiv(valeur;matrice;approximation)** a comme troisième paramètre la valeur approchée à trouver dans le tableau. Trois valeurs sont possibles :

- 0 correspond à une valeur exacte, et retourne #NA si la valeur n'est pas présente dans la zone.
- 1 correspond à une valeur par défaut, c'est-à-dire la valeur de la zone immédiatement inférieure à la valeur cherchée. *Dans ce cas la zone doit être ordonnée en ordre croissant.* Attention c'est la valeur par défaut de la fonction **Equiv**, c'est-à-dire celle qui est prise si le dernier argument est omis.
- -1 correspond à une valeur par excès, c'est-à-dire la valeur de la zone immédiatement supérieure à la valeur cherchée. *Dans ce cas la zone doit être ordonnée en ordre décroissant.*

4.2 Les affectations conditionnelles

4.2.1 Fonction SI

La fonction **Si(condition ;valeur si vraie ;valeur si faux)** permet de donner une valeur à une cellule suivant la valeur logique de la condition. La condition est exprimée à l'aide des d'opérateurs de comparaison et éventuellement d'opérateur booléen.

Les opérateurs de comparaison sont : =, <, >, <=, >=. Ils permettent tout aussi bien de comparer des nombres que des chaînes de caractères, dans ce dernier cas l'ordre est l'ordre lexicographique lié au code ASCII (les majuscules étant inférieures aux minuscules et les caractères accentués supérieurs à toutes les autres lettres).

Les opérateurs booléens sont présents sous forme fonctionnels dans Excel. Si C1, C2, C3 désignent des conditions (valeurs booléennes), C1 et C2 et C3 s'écrira **ET(C1 ;C2 ;C3)** ; de même C1 ou C2 ou ... s'écrira **OU(C1 ;C2 ;..)**. Enfin la négation s'écrit NON(C).

4.2.2 Fonction Recherche

La fonction **Recherche(valeur; bornes inférieures; valeur sur l'intervalle)** correspond au concept mathématique de fonction étagée ou constante par intervalle. Prenons un exemple : soit une fonction définie de la façon suivante (x représente le poids transporté et f(x) le coût de location d'un camion ou du train si x est très grand) :

Si	$0 \leq x < 20$	alors $f(x) = 24$
Si	$20 \leq x < 50$	alors $f(x) = 30$
Si	$50 \leq x < 200$	alors $f(x) = 40$
Si	$200 \leq x$	alors $f(x) = 100$

Pour définir cette fonction de coût, il nous suffit de connaître les bornes *inférieures* des intervalles (puisque la borne inférieure d'un intervalle est la borne supérieure de l'intervalle précédent), et les valeurs prises par la fonction sur chacun des intervalles. Nous allons donc associer une plage aux bornes (bien évidemment ordonnées par ordre croissant) et une autre plage associée aux valeurs prises par intervalle (dans l'ordre des intervalles).

Pour calculer la valeur pour un tonnage donné il suffira alors de donner ce tonnage. Mettons ceci en pratique :

	A	B	C	D	E
1	Intervalles	Coût			
2	0	30 €		Tonnage	76
3	20	45 €		Coût	150
4	50	150 €			
5	200	400 €			

La formule à entrer dans la cellule E3 est la suivante :

=RECHERCHE(E2;\$A\$2:\$A\$5;\$B\$2:\$B\$5)

où E2 contient le tonnage pour lequel on veut le coût.

Remarque : penser à toujours mettre en absolu les adresses des matrices des bornes et des valeurs, ou mieux nommer ces plages.

4.2.3 Fonction Recherchev

Fonction **Recherchev(valeur ;tableau de recherche ;index ;valeur proche)** {et sa symétrique **Rechercheh(valeur ;tableau de recherche ;index ;valeur proche)** }.

Le dernier argument (valeur proche) prend la valeur vrai ou faux (1 ou 0). Par défaut (donc si ce paramètre est absent) la valeur est *vraie*. C'est donc par la syntaxe simplifiée que nous commencerons.

a) La fonction **Recherchev(valeur ;tableau de recherche ;index)** correspond au même concept que la fonction recherche, en en y différenciant simplement par la présentation des intervalles et de la valeur prise sur chaque intervalle. Les zones correspondant aux bornes des intervalles et aux valeurs prises sur les intervalles sont remplacées par les colonnes d'une zone rectangulaire appelée **tableau de recherche** :

- Les bornes des intervalles correspondent à la première colonne de ce tableau, elles doivent être évidemment en ordre croissant.

- Les valeurs prises sont dans la colonne index du tableau de recherche (indépendamment des numéros d'Excel).

Sur l'exemple précédent la formule entrée en E3, serait :

=RECHERCHEV(E2;\$A\$2:\$B\$5;2)

Remarque : penser à toujours mettre en absolu l'adresse du tableau de recherche, ou, mieux, nommer cette plage.

- b) La fonction **Recherchev(valeur ;tableau de recherche ;index;0)** associe à chaque valeur de la première colonne du tableau de recherche, la valeur qui se trouve sur la même ligne dans la colonne index du tableau de recherche. Dans ce cas il est bien sûr inutile que la première colonne soit ordonnée).

La fonction **Rechercheh(valeur ;tableau de recherche ;index)** correspond à une présentation horizontale du tableau de recherche.

Il est possible de reconstruire ces fonctions à l'aide des deux fonctions **Index** et **Equiv**, c'est un exercice laissé au lecteur.

5 LES FONCTIONS MATRICIELLES

Théoriquement, l'utilisateur d'Excel ne peut entrer une formule que dans une cellule, il peut être toutefois nécessaire parfois d'entrer une formule sur une plage complète d'Excel, c'est le cas en particulier pour tout ce qui concerne le calcul matriciel. On appelle donc fonction matricielle toute fonction dont les valeurs sont prises par non pas une seule cellule, mais par un ensemble jointif de cellules ; d'un point de vue mathématique ce sont des fonctions à valeur dans un espace vectoriel et non plus à valeur dans R.

Pour entrer une fonction vectorielle, il faut procéder de la façon suivante :

- Sélectionner l'ensemble des cellules devant contenir les résultats. Ces cellules doivent constituer une plage dont la dimension (ligne colonne) dépend de la fonction utilisée et des éléments intervenant dans la fonction.
- Entrer la fonction voulue, soit en la tapant, soit à l'aide de l'assistant.
- Pour valider, il est nécessaire de maintenir appuyées les touche \hat{I} (Majuscule) et Ctrl tout en appuyant sur la touche entrée, la formule doit alors apparaître entourée d'accolades $\{\}$.

On retrouve parmi ces fonctions les fonctions classiques du calcul matriciel :

Inverse(M) pour calculer l'inverse de M

Produitmat(M1,M2) pour le produit matriciel

Transpose(M) pour la transposée,

par exemple, mais on trouve aussi des fonctions dont le résultat est un vecteur ou une matrice, pour le calcul de fréquences (voir le deuxième exemple ci dessous) ou la régression par exemple. Ces dernière fonctions ne sont pas des opérations matricielles, mais sont des fonctions (ou méthodes) donnant comme résultat un ensemble de valeurs.

Exemple (fichier Invmat.xls) : partant de la matrice : $M = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, nous calculons son inverse

M-1 et vérifions que $M \cdot M^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Voici les formules et les résultats :

	A	B
1	Matrice M	
2		
3	1	3
4	2	4
5		
6	Inverse M	
7		
8	=INVERSEMAT(A3:B4)	=INVERSEMAT(A3:B4)
9	=INVERSEMAT(A3:B4)	=INVERSEMAT(A3:B4)
10		
11	M*InverseM	
12		
13	=PRODUITMAT(A3:B4;A8:B9)	=PRODUITMAT(A3:B4;A8:B9)
14	=PRODUITMAT(A3:B4;A8:B9)	=PRODUITMAT(A3:B4;A8:B9)

Formules

	A	B
1	Matrice M	
2		
3	1	3
4	2	4
5		
6	Inverse M	
7		
8	-2	1,5
9	1	-0,5
10		
11	M*InverseM	
12		
13	1	0
14	0	1

Valeurs

Remarques :

1. Les accolades n'apparaissent pas dans l'affichage Formules, on voit cependant que la formule de l'inverse (ou du produit) n'a pas été obtenue par recopie d'après les adresses. Les accolades en revanche apparaissent dans la barre de formule.
2. La formule matricielle se reconnaît cependant dans la mesure où les adresses, bien que relatives, n'ont pas été changées comme dans le cas d'une recopie.

6 BASES DE DONNEES (LISTES) SOUS EXCEL.

Un fichier de données regroupe des informations structurées (nom, prénom, âge, etc.), pour un ensemble d'individus. Pour chaque individu sont recueillies les mêmes informations. Chaque information est associée à un champ du fichier de données qui est reconnu par son nom (*intitulé*), l'ensemble des informations recueillies pour un individu (ou, ce qui revient au même, l'ensemble des valeurs prises par chaque champ pour un individu) est appelé un enregistrement.

6.1 Structure d'une liste de données

A l'origine appelées bases de données, puis listes, les structures de données correspondant à un fichier se présentent sous une forme précise :

- Une base de données correspond à un ensemble rectangulaire de lignes et de colonnes contiguës, il est même recommandé pour qu'Excel reconnaisse implicitement une base de données, que la plage soit entourées de colonnes et lignes vides ou soit sur les bords de la feuille.
- Les champs correspondent aux colonnes, la première ligne de la plage contient les intitulés des champs.
- Les lignes correspondent aux enregistrements, c'est à dire aux valeurs prises par les champs pour les différents individus. Les enregistrements commencent donc à la deuxième ligne de la plage.

Pour illustrer notre propos nous utiliserons le fichier Excel BaseSal.xls, dont voici un extrait :

	A	B	C	D	E	F	G	H	I
1	NOM	PRENOM	SALAIRE	PRIMES	SEXE	Service	AGE	ANC.	Sal Tot
2	ridaba	vesa	928 €	455 €	M	Production	36	11	1 383 €
3	paroda	kola	1 347 €	801 €	M	Production	47	13	2 148 €
4	durezo	xike	1 402 €	345 €	M	Finance	38	0	1 747 €

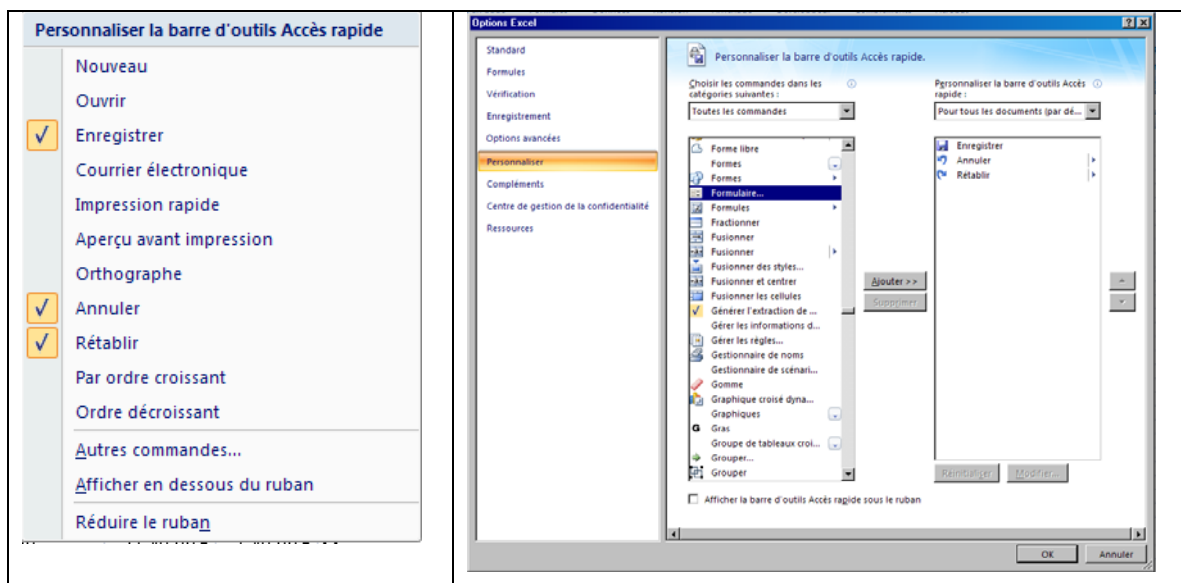
Les intitulés des champs sont dans la première ligne en gras, le champ "**Sal Tot**" en colonne I est un champ particulier dans la mesure où il n'est pas saisi mais calculé comme somme des valeurs des champs "**SALAIRE**" et "**PRIMES**".

Tous les menus ayant trait aux bases de données se trouvent en sous menu du menu *Données*. Pour utiliser ce menu, il est recommandé de mettre la cellule active à l'intérieur de la base de données, pour qu'Excel puisse reconnaître la plage correspondant à cette base. Les opérations élémentaires que l'on fait sur les bases de données sont de trois ordres :

- **Mise à jour : ajout, modification, suppression** ; ces opérations peuvent se faire directement sous Excel en modifiant ou supprimant les cellules correspondantes, ou alors après avoir ajouté la commande Formulaire dans la barre d'outil d'accès rapide



Pour ce faire, choisir la commande personnaliser la barre d'outils d'accès rapide, choisir autres commandes, puis toutes les commandes et enfin Formulaire :



La barre d'outil d'accès rapide devient alors :  . En cliquant sur l'icône formulaire on obtient alors :

Le champ calculé ne peut pas être saisi ou modifié avec cette grille, ce qui assure une meilleure cohérence des données que l'utilisation directe de la feuille.

- **Tri, sélection, extraction.** Le tri d'une base de données Excel se fait à l'aide du sous menu



Trier du bandeau du menu *Données*, on a la possibilité de sélectionner plusieurs clés de tri et l'ordre croissant ou décroissant sur chacune des clés. La sélection et l'extraction se font à l'aide de critères dont nous parlerons plus loin. Exemple du menu pour le tri :

Colonne	Trier sur	Ordre
Trier par	NOM	Valeurs
Puis par	SALAIRE	Valeurs
		De A à Z
		Du plus grand au plus petit

- **Statistiques descriptives.** Ceci est réalisé soit à l'aide de fonctions "base de données" que nous verrons ci dessous, soit à l'aide des sous menus Sous-totaux ou Rapport de tableau croisé dynamique. Pour utiliser le menu Sous-totaux, il faut que la base de données soit triée en conséquence, par exemple pour avoir les sous totaux par service, il faut auparavant avoir trié la base de données avec comme clé primaire le service. De plus le menu Sous-totaux crée des lignes supplémentaires, et donc modifie la Base. Nous verrons plus loin, avec les fonctions "base de données" comment obtenir les mêmes résultats dans une autre zone du classeur sans modifier la base.

6.2 Définition d'une zone de critères

L'extraction ou les fonctions statistiques base de données ne concernent pas tous les enregistrements, mais seulement les enregistrements dont certains champs doivent satisfaire des conditions ; ces conditions s'appellent critères de sélection ou plus simplement critères. Ces critères sont associés sous Excel à des plages qui doivent respecter une structure bien définie :

- La première ligne de la zone critère doit contenir le nom du champ sur le quel seront imposées des conditions, sauf dans le cas où le critère concerne plusieurs champs (par exemple prime>10% salaire), auquel cas on dit que l'on a affaire à un critère calculé.
- Les lignes suivantes contiennent les valeurs acceptables, si deux valeurs se trouvent sur la même ligne elles sont connecté par le connecteur logique "ET", sur deux lignes différentes par le connecteur logique "OU". Donnons quelques exemples pour illustrer cette notion, en commençant par des critères simples non calculés.

1. Pour opérer sur les salariés d'ancienneté supérieure à 10 ans **et** de salaire inférieur à 2000€, on créera la zone de critère suivante :

	L	M
1	ANC.	SALAIRE
2	>10	<2000

2. Mais pour opérer sur les salariés ayant une ancienneté supérieure à 10 ans **ou** de salaire inférieur à 2000€, on créera la zone de critère suivante (attention ici le ou est le ou mathématique, c'est à dire inclusif):

	L	M
1	ANC.	SALAIRE
2	>10	
3		<2000

3. Pour les hommes du service Comptabilité ayant un salaire >1800€ **ou** les femmes du service Marketing ayant un salaire >1800€, on utilisera la zone de critère :

	K	L	M
1	SEXE	Service	SALAIRE
2	M	Comptabilité	>1800
3	F	Marketing	>1800

Ne pas oublier de répéter >1800 sur la seconde ligne sinon les éléments sélectionnés par cette ligne n'auraient pas de condition sur le salaire.

4. Pour le critère 1000€<SALAIRE<2500€ on utilisera la présentation suivante :

	L	M
	SALAIRE	SALAIRE
	<2500	>1000

5. Création d'un critère calculé. Un tel critère portant généralement sur une relation entre plusieurs champs (bien qu'il soit possible d'utiliser un critère calculé sur un seul champ), la première ligne ne doit donc pas contenir le nom d'un champ et il est recommandé de la laisser vide. La valeur du critère est exprimée comme une formule portant sur les valeurs du premier enregistrement, c'est à dire de la deuxième ligne de la base de données. Les adresses correspondant aux valeurs des champs doivent être en relatif (au moins pour ce qui est de la ligne).

Exemple : prime supérieure à 10% du salaire :

	L
1	
2	=D2>10%*C2

Formule

L
FAUX

Valeur

Ne pas oublier le signe = pour le début de la formule, la valeur affichée (ici FAUX) représente la valeur logique du critère sur le premier enregistrement.

Pour la conjonction ou la disjonction de critère calculé, on peut soit utiliser la méthode présentée ci dessus en positionnant des lignes, soit utiliser les fonctions logiques d'Excel.

Par exemple pour le critère "**(prime>10%*salaire) ou (prime<5%*salaire)**", on pourra soit *ajouter au critère précédent la ligne L3* contenant =D2<5%*C2, soit *modifier la ligne L2* : =OU(D2>10%*C2;D2<5%*C2).

Remarque : pour vérifier que les critères "fonctionnent" bien, il est recommandé d'utiliser le sous menu "**Filtre -Filtre élaboré**", après avoir mis la cellule active dans la plage de base de données, en choisissant l'action "filtrer sur place", et en indiquant la plage zone de critère. Les lignes correspondant aux enregistrements qui ne satisfont pas le critère sont alors masquées, il suffit d'utiliser le sous menu "**Filtre - Afficher tout**" pour afficher toute la base.

6.3 Extraction d'une partie de la base de données

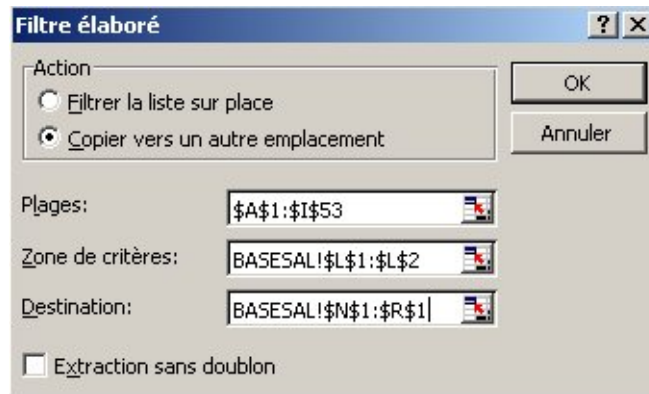
Il est souvent utile de pouvoir extraire certaines informations d'une base de données, pour y faire des traitements, sans modifier l'ensemble de la base. L'extraction consiste à sélectionner des enregistrements en fonction de la valeur de certains champs, mais aussi à n'afficher que les champs concernés par l'analyse. Par exemple, il sera généralement inutile de conserver les noms et prénoms de notre base, les traitements portant toujours sur les autres champs.

Pour faire cette extraction, il faut :

1. Créer une zone de critère, correspondant à la sélection des enregistrements. Par exemple les primes supérieures à 10% du salaire en L1:L2.
2. Créer une plage de réception des informations que l'on veut conserver, pour ce faire, il faut créer sur une ligne une zone de cellules adjacente contenant les intitulés des champs que l'on veut conserver. Exemple pour conserver les champs, SALAIRE, PRIME, SEXE, ANC. et SERVICE, on utilisera la plage N1:R1 :

	M	N	O	P	Q	R
1		SALAIRE	PRIMES	SEXE	ANC.	Service
2						

3. Puis dans le menu "**Données**", nous choisissons le menu "**Filtre – Filtre élaboré**" et remplissons la boîte de dialogue :



Remarques :

- La zone de destination et la plage de la base de données (et le critère, s'il contient un critère calculé) doivent être sur la même feuille. On pourra éventuellement déplacer les informations sur une autre feuille après leur extraction
- Attention : toute la partie située sous la ligne de titre de la zone de destination est détruite lors de l'extraction.

6.4 Fonctions Base de Données

Les fonctions base de données permettent d'obtenir un certain nombre de renseignements statistiques pour un sous ensemble de la base de données satisfaisant à un critère, par exemple nombre de valeurs, moyenne, écart type etc.

Le nom des fonctions base de données commence toujours par les deux lettres BD (en anglais DB), suivi du nom de la fonction statistique concernée. Ces fonctions comportent toujours trois arguments :

- L'adresse de la base de données
- Le nom champ sur le quel doit se faire le calcul (indiquer le nom du champ entre guillemets ou l'adresse de son intitulé)
- L'adresse de la zone de critère servant à sélectionner les enregistrements.

Par exemple, pour calculer le salaire moyen et le montant des primes des employés ayant des primes supérieures à 10% de leur salaire, on utilisera les deux formules suivantes :

=BDMOYENNE(\$A\$1:\$I\$53;"SALAIRE";L1:L2)

=BDSOMME(\$A\$1:\$I\$53;\$D\$1;L1:L2)

La base de données est dans la plage \$A\$1:\$I\$53, il serait préférable de lui donner un nom.

L1:L2 correspond à la zone de critère.

\$D\$1 est la cellule contenant l'intitulé du champ "Primes".

7 EXEMPLE DE L'UTILISATION D'EXCEL EN STATISTIQUE DESCRIPTIVE

Pour terminer nous allons ici donner un exemple de l'utilisation d'Excel en statistique descriptive, ce qui nous permettra d'approfondir un peu les graphiques et d'illustrer les fonctions matricielles et de base de données, ainsi que l'utilisation des tables.

Histogrammes des classes d'une variable numérique

Le classeur Frequences.xlsx est composé de deux feuilles, la première ("**Data**") contient les données dans une plage nommée "*donnees*", la seconde ("**Histo**") contient les calculs et le graphique. La plage *donnees* est composée de 100 cellules comportant ici des nombres compris entre 90 et 210.

Voici la structure de la feuille "**Histo**" :

	A	B	C	D	E	F	G	H	I	J
1	Milieux	Bornes	Effectif			Calculs intermédiaires			Choix	
2		90								
3	96	102	1			Minimum	90		nbre classes	10
4	108	114	5			Maximum	210			
5	120	126	5			pas	12			

Remarquez les accolades dans la zone de formule.

Les cellules de calculs intermédiaires et la cellule du nombre de classes sont nommées :

Cellule	Nom	Formule
G3	minval	=MIN(donnees)
G4	maxval	=MAX(donnees)
G5	pas	=(maxval-minval)/nbclas
J3	nbclas	

Dans la colonne B, on a créé une plage contenant les bornes des intervalles des classes (ici les 10 classes), en partant de la valeur minimale :

	B
1	Bornes
2	<i>=minval</i>
3	<i>=B2+pas</i>

La formule de B3 est recopiée vers le bas. La première et la dernière cellule de la plage sont mise en gras et italique car Excel n'en a pas besoin pour le calcul des fréquences, seule seront prises en compte les cellules B3:B11, c'est à dire 9 cellules excluant le minimum et le maximum ; la plage de réception des fréquences sera elle constituée de 10 cellules : la première contenant le nombre de données inférieures à B3, la dernière contenant le nombre de données supérieures ou égales à B11. Dans la plage C3:C12, on entre alors la formule matricielle :

	B	C
1	Bornes	Effectif
2	= <i>minval</i>	
3	=B2+pas	=FREQUENCE(donnees;B3:B11)
4	=B3+pas	
5	=B4+pas	
6	=B5+pas	
7	=B6+pas	
8	=B7+pas	
9	=B8+pas	
10	=B9+pas	
11	=B10+pas	
12	=B11+pas	

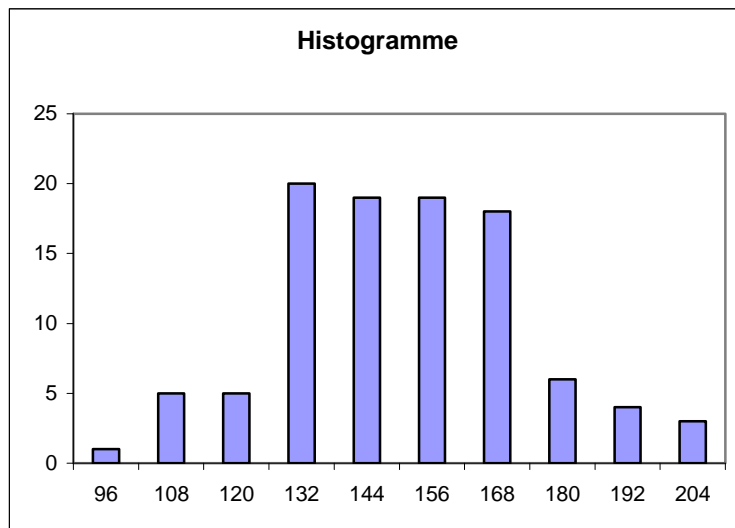
Avant ↑-Ctrl-Entrée

	B	C
1	Bornes	Effectif
2	= <i>minval</i>	
3	=B2+pas	=FREQUENCE(donnees;B3:B11)
4	=B3+pas	=FREQUENCE(donnees;B3:B11)
5	=B4+pas	=FREQUENCE(donnees;B3:B11)
6	=B5+pas	=FREQUENCE(donnees;B3:B11)
7	=B6+pas	=FREQUENCE(donnees;B3:B11)
8	=B7+pas	=FREQUENCE(donnees;B3:B11)
9	=B8+pas	=FREQUENCE(donnees;B3:B11)
10	=B9+pas	=FREQUENCE(donnees;B3:B11)
11	=B10+pas	=FREQUENCE(donnees;B3:B11)
12	=B11+pas	=FREQUENCE(donnees;B3:B11)

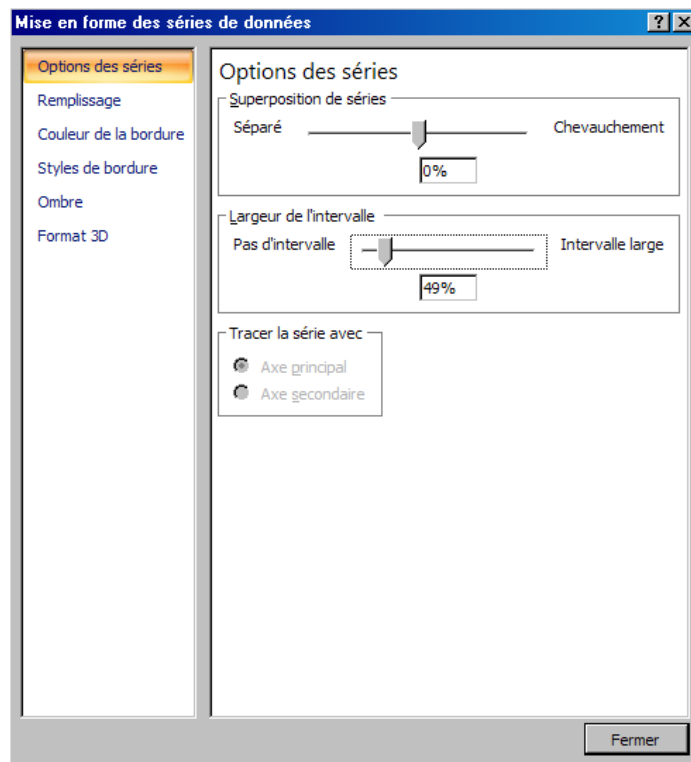
Après ↑-Ctrl-Entrée

Nous créons enfin dans la colonne A, dans la plage A3:A12, les valeurs médianes des classes qui serviront d'étiquettes aux blocs de l'histogramme. Il nous faudra donc introduire cette plage comme "Etiquettes des Abscisses" dans Le menu Sélection des données du bandeau Création de graphiques.

On obtient alors le graphique suivant :



Cette représentation ne convient pas à des données quantitatives, en effet il est d'usage dans ce cas que les blocs soient collés pour montrer la continuité de la variable. Après avoir cliqué sur un des blocs, dans le menu Mise en forme de la sélection du bandeau Mise en forme de graphique, on obtient la boîte de dialogue suivante :



Nous mettrons alors 0 dans la largeur de l'intervalle, nous pouvons aussi faire afficher les effectifs pour obtenir le graphique final suivant :

