

EXPÉRIENCE EN CLASSE SUR LE TABLEUR

Thème : STATISTIQUES en 4^{ème}

Alain JUILLAC (Collège Condorcet - Puy Guillaume)

Véronique JUILLAC (Lycée Montdory - Thiers)

Cet article est paru sur le Bulletin Vert de l'APMEP n°436.

Le fichier informatique (*) correspondant au format EXCEL est disponible sur le site de l'IREM de Clermont : <http://wwwmaths.univ-bpclermont.fr/irem/college/>

L'objectif de cette séquence est de mettre en application les nouvelles notions étudiées en cours de 4^e : effectifs cumulés, fréquences cumulées et moyenne pondérée d'une série à l'aide d'un tableur grapheur.

Cette séquence est composée de trois activités informatiques utilisant le logiciel EXCEL ou WORKS.

La première permet, d'abord, aux élèves de manipuler le tableur pour des calculs simples de sommes et de fréquences sur une série à caractère qualitatif. Ils découvrent ainsi l'intérêt du logiciel qui prend en charge des calculs répétitifs. Ensuite, les élèves peuvent, à l'aide du grapheur, produire rapidement les graphiques demandés et s'intéresser alors à leur pertinence.

La deuxième activité comporte toutes les étapes du traitement de données qui sont au programme : leur organisation (dépouillement, tableau d'effectifs), leur représentation (diagrammes, graphiques) et leur synthèse (effectifs cumulés croissants, effectifs cumulés décroissants, moyenne) afin d'étudier une série à caractère quantitatif discret. Les élèves doivent se servir de leurs connaissances et des fonctions du tableur grapheur.

La troisième activité enfin, porte sur l'étude d'une série à caractère quantitatif continu. Elle est composée de deux parties.

La première partie met en évidence l'importance du choix du nombre de classes et de leur amplitude. En effet, elle permet de mener expérimentalement, à l'aide du logiciel, la recherche d'une répartition en classes, adaptée à la situation, en visualisant rapidement les différents histogrammes associés.

Dans la deuxième partie la répartition en classes est donnée. Les élèves doivent alors utiliser leurs connaissances et les fonctions du tableur pour effectuer la synthèse des données (fréquences cumulées croissantes, fréquences cumulées décroissantes, moyenne).

I - Activités proposées

Présentation des pré-requis et objectifs de la séquence

En sixième et cinquième, les élèves ont travaillé sur la représentation de relevés statistiques sous forme de tableaux ou de graphiques. Ils ont également abordé les notions d'effectifs, de fréquences et de classes.

En quatrième, on prolonge ces notions avec les effectifs et fréquences cumulés et on aborde une nouvelle phase de la synthèse des données avec la moyenne pondérée d'une série statistique. Pour cette nouvelle notion on distinguera le cas où l'on dispose de données sur l'ensemble des éléments de la population étudiée, du cas où les données ne concernent qu'un regroupement en classes de la population. Dans ce dernier cas, on n'obtiendra donc qu'une valeur approchée de la moyenne de la série.

Cette séquence est composée de trois activités dans lesquelles vont être appliquées les nouvelles notions vues en cours. Elle se déroulera en salle informatique, durant deux séances, avec le logiciel EXCEL ou WORKS.

Les tableurs grapheurs, utilisés dès la cinquième en technologie introduisent une nouvelle manière de désigner une variable : par l'emplacement de la cellule dans un tableau.

Pour de telles activités informatiques, il est donc souhaitable que les élèves aient acquis quelques rudiments de fonctionnement du tableur grapheur dans le cadre de séances précédentes en cours de mathématiques (ou en technologie en cinquième) :

- savoir sélectionner des cellules, entrer un nombre ou du texte dans une cellule, entrer une formule dans une cellule, utiliser les références relatives et les références absolues, modifier le contenu d'une cellule, recopier des cellules ;
- savoir créer un graphique, changer de type de graphique, étiqueter un graphique.

Activité n°1

On étudie dans cette activité une série à caractère qualitatif à travers une étude de marché faite par un constructeur automobile.

Énoncé : « Couleurs de voitures »

On considère les résultats d'une enquête, réalisée par un constructeur automobile, portant sur la couleur des voitures vendues, à partir de bons de commandes de véhicules neufs.

- c) Interprétation des résultats :
- Que représente le nombre 34 dans la colonne des E.C.C. ?
 - Combien de familles ont au moins deux téléphones portables ?

4°) Moyenne de la série :

- a) Calculer la moyenne à partir des résultats de l'enquête, en utilisant la formule : $= \text{moyenne}(\dots)$. Interpréter ce résultat.
- b) Calculer la moyenne de la série à partir des données du tableau :
- remplir la 1^{ère} case de la colonne produit en utilisant la formule : $= \dots \times \dots$;
 - finir de compléter cette colonne (recopier vers le bas) et la case total produits ;
 - calculer alors la moyenne cherchée en utilisant la formule : $= \dots / \dots$.
- c) Que constate-t-on ?

Activité n°3

On étudie dans cette activité une série à caractère quantitatif continu à travers un exercice de sensibilisation aux problèmes d'excès de vitesse.

Énoncé : « Contrôle de vitesses »

Sur une route limitée à 90 km/h, les gendarmes ont relevé les vitesses.

Les élèves ont un document papier et ouvrent deux fichiers informatiques.

Premier fichier informatique(*) :

Il comprend la **liste brute de l'enquête** (sous forme d'une seule colonne) ainsi que le tableau suivant pour chaque exemple de classe choisi.

largeur des classes :

81	94	103	95
84	96	76	100
85	101	78	103
88	87	90	99
89	91	92	100
89	95	97	102
90	100	92	105

Vitesse comprise entre		Centres des classes	Effectifs
Borne inférieure incluse	Borne supérieure exclue		

Document papier :

- 1°) Recherche du regroupement en classes le mieux adapté à cette série :
- a) Trier ces données en suivant les mêmes démarches que dans l'activité n°2.
- b) Proposer, en remplissant les tableaux donnés, trois exemples de regroupement en classes d'amplitudes supérieures ou égales à 1.
- Pour cela, remplir les deux premières cases de la colonne «borne inférieure incluse» puis recopier vers le bas. Faire la même chose pour la colonne «borne supérieure exclue».
- Quelles largeurs de classes ont été choisies ?
- 1^{er} exemple : 2nd exemple : 3^{ème} exemple :
- c) Dans les trois exemples, construire l'histogramme.
- Pour cela, remplir la colonne «Centres des classes» :
- dans la 1^{ère} case de cette colonne en utiliser la formule : $= (\dots + \dots) / 2$;
 - finir de compléter la colonne ;
- Puis compléter la colonne «Effectifs» à partir des résultats de l'enquête.
- Ensuite, sélectionner les colonnes «Centres des classes» et «Effectifs» puis cliquer sur l'icône graphique. (Pour obtenir des rectangles accolés : cliquer sur les rectangles du graphique obtenu, aller dans «options» et mettre le chiffre 0 dans «largeur de l'intervalle»).
- Lequel paraît être le plus pertinent ? Justifier.

Deuxième fichier informatique(*) :

Il comprend la **liste brute de l'enquête** (sous forme d'une seule colonne), le tableau suivant ainsi que les différents types de calculs de moyenne.

Vitesse comprise entre		Centres des classes	Effectifs	Fréquences (%)	Fréquences Cumulées Croissantes (%)	Fréquences Cumulées Décroissantes (%)	Produits Centre des classes par Effectifs
Borne inférieure incluse	Borne supérieure exclue						
76	80						
80	84						
84	88						
88	92						
92	96						
96	100						
100	104						
104	108						
Totaux							

Moyenne de la série :

a) A l'aide des résultats de l'enquête :

b) A l'aide des données du tableau :

Document papier (suite) :

2°) Étude de la série à partir d'un regroupement en classe donné :

a) Trier de nouveau ces données.

b) Remplir la 1^{ère} case de la colonne « centres des classes » en utilisant la formule : $= (\dots + \dots) / 2$.

Finir de compléter la colonne.

c) Remplir la colonne «Effectifs» à partir des résultats de l'enquête puis la colonne «Fréquences».

d) Si ce regroupement en classes n'a pas été choisi dans la partie 1°) de cette activité, construire l'histogramme en sélectionnant les colonnes «Centres des classes» et «Effectifs».

e) Remplir les colonnes des Fréquences Cumulées Croissantes (F.C.C.) en pourcentages et des Fréquences Cumulées Décroissantes (F.C.D.) en pourcentages de la série :

- Pour les F. C. C. :
 - 1^{ère} case : = fréquence de la première valeur ;
 - 2nde case : = valeur de la 1^{ère} case + fréquence de la deuxième valeur ;
 - 3^{ème} case et les suivantes : recopier vers le bas.
- Pour les F.C.D. : utilisez vos connaissances et la question précédente.

Les deux affirmations suivantes sont-elles vraies ou fausses :

«Moins de 50 % des automobilistes contrôlés respectent la limitation de vitesse.».

«Plus de 25 % des automobilistes contrôlés roulent à 100 km/h et plus.».

f) Calculer la moyenne de la série :

- avec les données de départ en utilisant la formule : $= \text{moyenn}(\dots)$;

- avec les données du tableau :

· remplir la 1^{ère} case de la colonne produit en utilisant la formule : $= \dots \times \dots$;

· finir de compléter cette colonne (recopier vers le bas) et la case total produits ;

· calculer alors la moyenne cherchée en utilisant la formule : $= \dots / \dots$.

Que constate-t-on ? Justifier.

II - Commentaires sur les activités et sur le travail des élèves

Activité n°1

Commentaires sur l'activité :

Cette activité permet, dans un premier temps, aux élèves de manipuler le tableur sur des calculs simples de sommes et de fréquences. Ils voient l'intérêt du logiciel qui prend en charge des calculs répétitifs. Dans un deuxième temps, les élèves peuvent, à l'aide du grapheur, produire rapidement les deux graphiques demandés et s'intéresser alors à leur pertinence. Le fait que les fréquences des couleurs noires et grises soient proches permet de mettre en évidence que le diagramme en barres offre une lecture plus facile et précise que le diagramme circulaire.

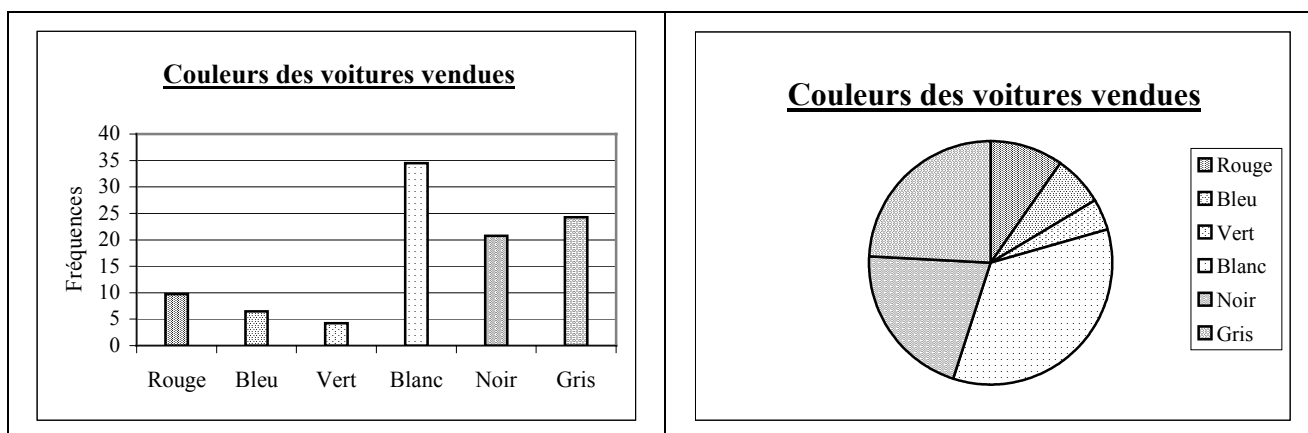
Résultats informatiques des élèves :

1°)

Couleurs	Effectifs	Fréquences (%)
Rouge	39	9,75
Bleu	26	6,5
Vert	17	4,25
Blanc	138	34,5
Noir	83	20,75
Gris	97	24,25
Totaux	400	100

2°)

Couleurs	Fréquences (%)
Rouge	9,75
Bleu	6,50
Vert	4,25
Blanc	34,50
Noir	20,75
Gris	24,25



Commentaires sur le travail des élèves :

1°) Les élèves retrouvent rapidement les fonctions du tableur nécessaires au remplissage des deux tableaux.

2°) Pour effectuer le diagramme en bâtons de la série, les élèves doivent choisir parmi les différentes sortes de graphiques proposées par l'ordinateur. Certains remarquent alors une divergence entre langage mathématique et langage informatique. En effet ce dernier utilise le terme histogramme pour désigner le diagramme en bâtons du langage mathématique. Ils effectuent ensuite sans problème le diagramme dans la mesure où ils ont fait la bonne sélection de données, facilité par la présentation d'un deuxième tableau (comportant uniquement couleurs et fréquences). Certains élèves donnent le titre «Fréquences des couleurs de voitures» à leur graphique, ce qui révèle une confusion entre le titre du graphique et la légende de l'axe des ordonnées.

3°) Peu d'élèves, dans le diagramme circulaire, pensent à utiliser les fonctions du grapheur pour faire correspondre les couleurs des secteurs angulaires du diagramme avec les couleurs de voitures. Pour ces diagrammes en bâtons et circulaires, certains élèves utilisent les représentations en trois dimensions.

4°) Les élèves se rendent compte assez facilement que le diagramme en bâtons est plus pertinent que le diagramme circulaire, mais ils ont beaucoup de difficultés à le justifier. En effet, ils ne discernent pas correctement la différence d'aspect des deux : plus quantitatif pour le premier et plus qualitatif pour le second.

Activité n°2

Commentaires sur l'activité :

Dans cette activité, il apparaît différentes étapes du traitement de données :

- organisation des données (dépouillement, tableau d'effectifs) ;
- représentation des données (diagrammes, graphiques) ;
- synthèse des données (E.C.C., E.C.D., moyenne).

En découvrant les résultats de l'enquête, on remarque la difficulté à lire une liste de données lorsque la population est nombreuse, d'où la nécessité de trier ces données.

La première partie de cette activité met évidence l'intérêt du logiciel pour effectuer un tri et donc dépouiller une enquête plus rapidement.

Dans la deuxième partie, il s'agit d'une approche dynamique d'une grande variété de graphiques (s'affichant «à la demande» sur l'écran). Cela conduit naturellement les élèves à examiner leur pertinence pour illustrer cette situation.

Dans la troisième partie, les élèves peuvent, à l'aide de leurs connaissances et des fonctions du tableur, calculer rapidement les E.C.C. et E.C.D. en évitant les calculs répétitifs. Ils doivent alors interpréter ces résultats. Le but de cette partie, plus difficile, est de faire comprendre aux élèves l'utilité de faire de tels calculs.

Dans la quatrième partie, les élèves calculent, à l'aide de leurs connaissances et des fonctions du tableur, la moyenne arithmétique simple puis pondérée de la série. On peut faire remarquer aux élèves que personne n'a 1,25 portable : la moyenne est un paramètre descriptif global.

Résultats informatiques des élèves :

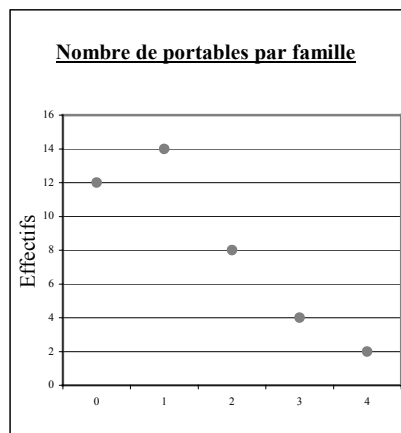
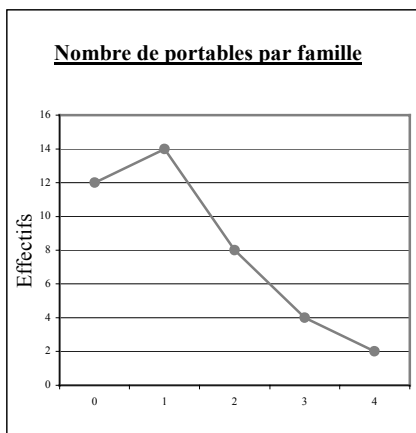
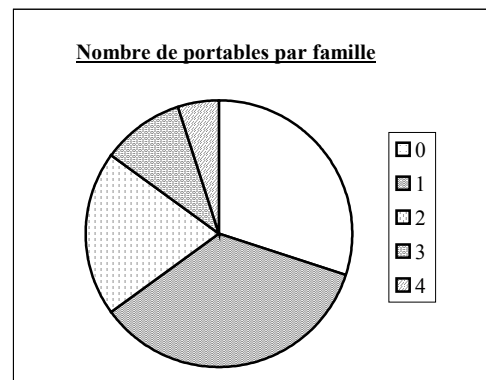
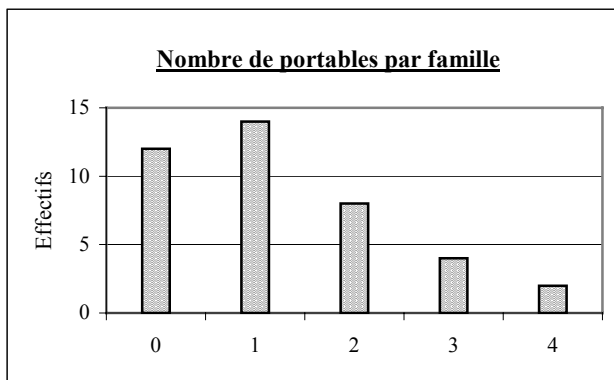
Nombres de portables	Effectifs	Effectifs Cumulés Croissants	Effectifs Cumulés Décroissants	Produits Nombres de Portables par Effectifs
0	12	12	40	0
1	14	26	28	14
2	8	34	14	16
3	4	38	6	12
4	2	40	2	8
Totaux	40			50

Moyenne de la série :

a) A l'aide des résultats de l'enquête :

1,25
1,25

b) A l'aide des données du tableau :



On fait remarquer aux élèves, que le premier des deux graphiques ci-contre n'a pas de signification mathématique : il n'y a pas lieu de rejoindre les points car il n'existe que des valeurs entières pour cette série (nombre de portables).

Commentaires sur le travail des élèves :

1°) Le dépouillement de la série est effectué aisément par les élèves car d'une part ils retrouvent rapidement la fonction tri du tableur, et d'autre part cette tâche est facilitée par le fait que les données sont présentées sous forme d'une seule colonne. En effet, des problèmes spécifiques à certains tableurs (Works) apparaissent si les données sont en plusieurs colonnes.

- 2°) Les élèves effectuent sans problème les trois types de graphiques demandés et discernent alors de façon plus claire la pertinence du diagramme en bâtons.
- 3°) Dans les calculs E.C.C. et E.C.D. les élèves voient l'utilité du tableur qui élimine les calculs répétitifs. Par contre des difficultés interviennent lors de l'interprétation de ces résultats.
- 4°) Les élèves calculent la moyenne de la série de deux façons différentes et constatent ainsi que la fonction moyenne du tableur (moyenne arithmétique) donne le même résultat que le calcul de la moyenne pondérée.

Activité n°3

Commentaires sur l'activité :

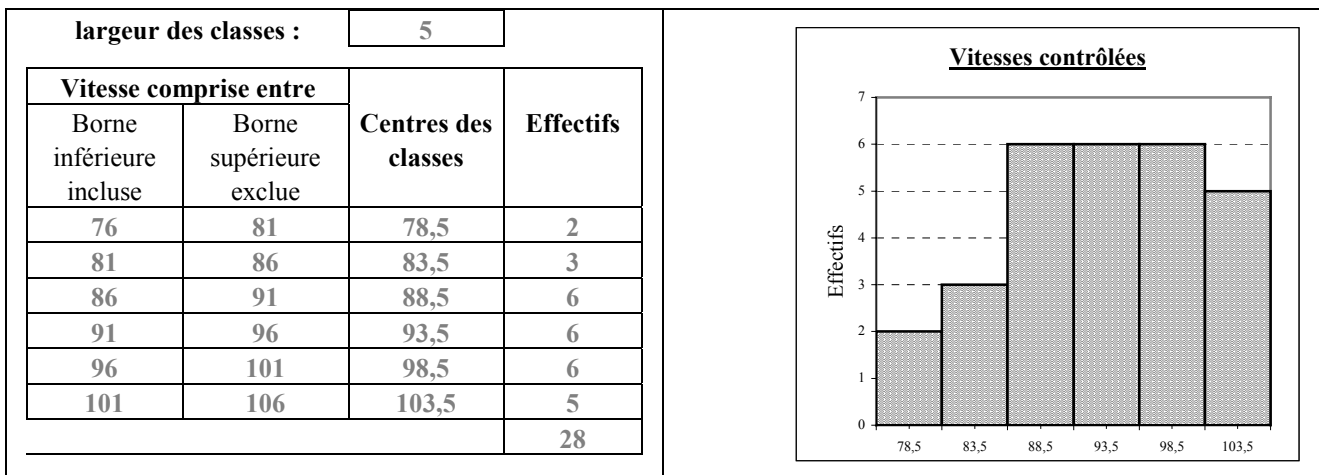
Si le caractère est continu, il est commode de regrouper ses valeurs en classes (une classe est définie par ses bornes : on convient d'inclure la borne inférieure et d'exclure la borne supérieure).

La première partie de l'activité met en évidence l'importance du choix du nombre de classes et de leur amplitude. En effet, elle permet de mener expérimentalement la recherche d'une répartition en classes, adaptée à la situation, en visualisant rapidement les différentes allures des histogrammes associés. Les élèves constatent, d'une part, qu'un trop grand nombre de classes apporte peu de simplification aux données de départ (triées) et d'autre part, qu'un nombre trop réduit cause une grande perte d'informations.

Dans la deuxième partie la répartition en classes est donnée. Les élèves doivent remplir les colonnes «Effectifs», «Fréquences», «Centres des classes», «F.C.C.», «F.C.D.» et «Produits Centres des classes par Effectifs» du tableau afin de répondre à deux questions d'interprétation et de calculer la moyenne arithmétique simple puis pondérée de la série. Ils constatent alors que cette dernière n'est qu'une valeur approchée de la première. De plus, il est intéressant de faire remarquer aux élèves que la moyenne ne divise pas la population étudiée en deux parties égales.

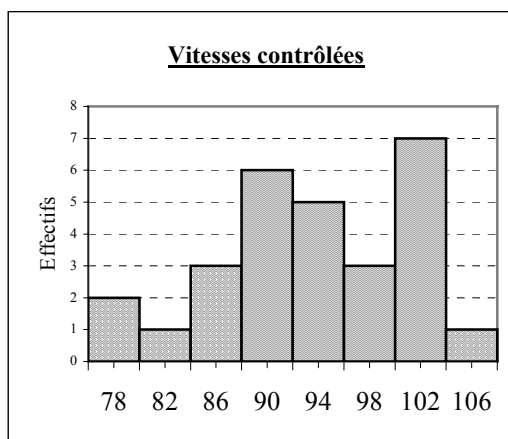
Pour répondre aux 2 questions d'interprétation, on utilise les F.C.C. et F.C.D.. On peut alors faire remarquer aux élèves que dans la mesure où les F.C.C. et les F.C.D. ont été calculées avec le centre des classes, il y a une perte d'information (+ ou - grande suivant le choix des classes). Les réponses faites à ces deux questions, dans ce cadre là, ne sont donc pas d'une rigueur absolue. Mais aux vues des résultats obtenus (la lecture du tableau donne : 42,9% roulent à moins de 92 km/h et 28,6% roulent à 100 km/h et plus), on peut donner des réponses satisfaisantes tout en gardant en tête une marge d'erreur dans cet exemple là. On peut alors demander aux élèves d'effectuer les mêmes calculs avec les données brutes (on trouve : 39,3% roulent à 90km/h au plus et 28,6% roulent à 100km/h et plus). On constate donc que les deux affirmations sont toujours vraies.

Résultats informatiques des élèves pour la partie 1 :



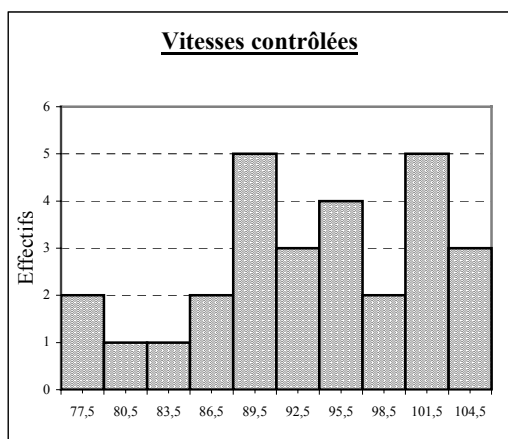
largeur des classes : 4

Vitesse comprise entre		Centres des classes	Effectifs
Borne inférieure incluse	Borne supérieure exclue		
76	80	78	2
80	84	82	1
84	88	86	3
88	92	90	6
92	96	94	5
96	100	98	3
100	104	102	7
104	108	106	1
Totaux			28



largeur des classes : 3

Vitesse comprise entre		Centres des classes	Effectifs
Borne inférieure incluse	Borne supérieure exclue		
76	79	77,5	2
79	82	80,5	1
82	85	83,5	1
85	88	86,5	2
88	91	89,5	5
91	94	92,5	3
94	97	95,5	4
97	100	98,5	2
100	103	101,5	5
103	106	104,5	3
Totaux			28



Résultats informatiques des élèves pour la partie 2° :

Vitesse comprise entre		Centres des classes	Effectifs	Fréquences (%)	Fréquences Cumulées Croissantes (%)	Fréquences Cumulées Décroissantes (%)	Produits Centres des Classes par Effectifs
Borne inférieure incluse	Borne supérieure exclue						
76	80	78	2	7,1	7,1	100,0	156
80	84	82	1	3,6	10,7	92,9	82
84	88	86	3	10,7	21,4	89,3	258
88	92	90	6	21,4	42,9	78,6	540
92	96	94	5	17,9	60,7	57,1	470
96	100	98	3	10,7	71,4	39,3	294
100	104	102	7	25,0	96,4	28,6	714
104	108	106	1	3,6	100,0	3,6	106
Totaux			28	100,0			2620

Moyenne de la série :

a) A l'aide des résultats de l'enquête :

92,92857

b) A l'aide des données du tableau :

93,57143

Commentaires sur le travail des élèves :

1°) Après le premier choix de largeur de classe, la plupart des élèves orientent de façon judicieuse leurs choix suivants. Cependant quelques élèves font des choix de largeur de classes qui les écartent complètement de l'objectif attendu. D'où l'intérêt de la deuxième partie de cette activité qui porte sur la même série et qui propose un regroupement en classes pertinent. Ainsi l'ensemble des élèves se rend compte de l'influence du choix des largeurs de classe sur la lecture d'un histogramme.

2°) Pour construire l'histogramme de la série, les élèves se heurtent aux limites des logiciels Works et Excel. En effet ceux qui travaillent sur Works constatent qu'ils ne peuvent obtenir un histogramme comme ils l'ont vu en cours car contrairement à Excel, on ne peut pas réduire à zéro l'espace entre les différents rectangles. D'autre part, quelque soit le logiciel sur lequel ils travaillent, les élèves ne peuvent faire apparaître que le centre de la classe et non pas la borne inférieure et supérieure de celle-ci.

Les élèves ont beaucoup de mal à interpréter les résultats obtenus dans les colonnes F.C.C. et F.C.D.. En effet, ils ont des difficultés à faire le lien entre les résultats du tableur et les deux phrases :

«Moins de 50 % des automobilistes contrôlés respectent la limitation de vitesse.»

«Plus de 25 % des automobilistes contrôlés roulent à 100 km/h et plus»

Pour un regroupement en classes données de la série, les élèves constatent facilement que la moyenne arithmétique obtenue avec la fonction moyenne du tableur et les données de départ d'une part, et la moyenne pondérée obtenue avec les centres des classes d'autre part ne donnent pas le même résultat. Par contre ils ont beaucoup de difficultés à justifier cette différence par une perte d'informations.

Sources des données des différentes activités

Dans les deux premières activités, les données ont été inventées ; dans la troisième, elles ont été inspirées du manuel Hachette 5/5, classe de 4^e, édition 1998, chapitre 7 exercice n°29 page 125.

Bibliographie

Les publications sur le tableur de l'IREM de Clermont-Ferrand :

« Mathématiques et tableur au collège » sous forme de 4 brochures :

- Classe de 4^e (1998)
- Classe de 3^e (1999)
- Classes de 4^e et 3^e (2000)
- An 2001

