

TD n°1 de Statistiques : histogrammes

Préliminaire sur les densités : Monaco : 32 543 h, Japon : 127 417 244 h (source INED 2005). Pour comparer la démographie de ces deux pays, ces nombres sont insuffisants. Il faut déterminer les **densités** de population, et pour cela tenir compte des superficies de ces pays. Monaco : 2,02 km², Japon : 378 000 km².

Calculer les densités de ces deux pays pour comparer leur démographie.

Monaco : $32\,543 \div 2,02 \approx 16110$ h/km², Japon : $127\,417\,244 \div 378\,000 \approx 337$ h/km².

Le calcul de la densité de ces pays montre que, malgré l'évidence d'une population totale bien supérieure au Japon qu'à Monaco, il y a moins de personnes par unité d'aire, dans ce pays dont les villes – très densément peuplées – ne représentent qu'une partie du territoire, l'autre partie de campagne et de montagnes, très étendue étant relativement très peu peuplée. Le territoire de Monaco est presque exclusivement urbain, donc très dense. Pour conclure, l'idée de densité permet de se faire une idée du rapport entre l'effectif d'une population et l'espace qu'elle occupe.

1] Série ayant des classes d'amplitudes égales

a) Le directeur des ressources humaines (DRH) d'une entreprise a relevé la distribution statistique de l'ancienneté des cadres de son entreprise, exprimée en années :

Classes	[6,5;8[[8;9,5[[9,5;11[[11;12,5[[12,5;14[[14;15,5[[15,5;17[Total
Effectifs	3	8	12	19	9	5	4	60

► Représenter sur le graphique ci-contre l'*histogramme des effectifs* (constitué de rectangles contigus). Les classes ayant la même *amplitude* on peut graduer l'axe vertical.

Nous avons tracé les rectangles en prenant pour hauteur les effectifs. Pour le polygone des effectifs, nous avons joint les milieux des côtés supérieurs de ces rectangles. Le cadre supérieur était destiné à un titre expliquant la signification de cette représentation.

b) Modifions légèrement cet exemple en regroupant les deux dernières en une seule : la classe [14;17[. De même pour les deux premières classes. Ces classes n'ont plus la même amplitude que les autres.

Pour représenter l'effectif d'une classe, il faut donner au rectangle de l'histogramme une *hauteur calculée* pour que son aire (et non sa hauteur) soit proportionnelle à l'effectif.

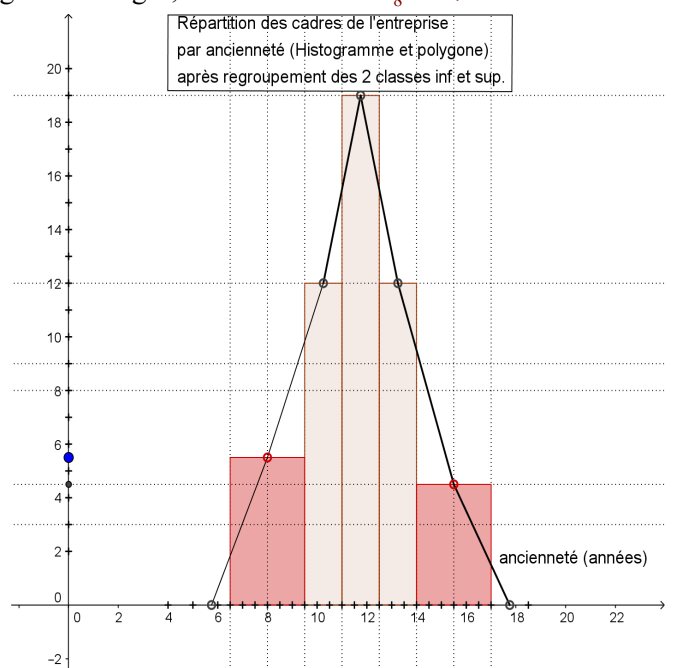
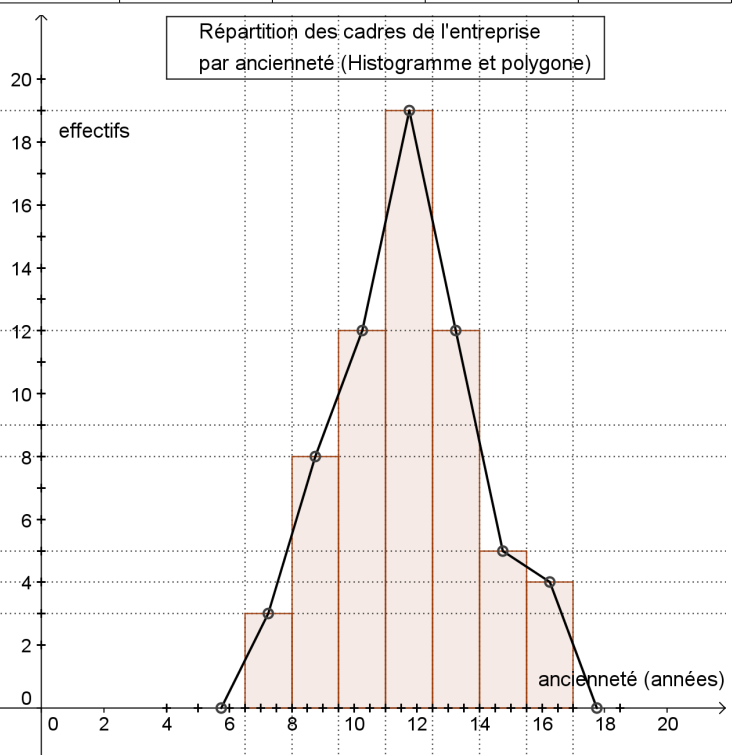
La bonne méthode est donc le calcul des *densités* (rapports effectif/amplitude), puis l'application d'un coefficient de proportionnalité (hauteur/densité) pour le calcul des hauteurs.

► Compléter le tableau pour déterminer la hauteur *h* des rectangles. Le coefficient de proportionnalité *k* pour le calcul des hauteurs est ici déterminé par la hauteur des rectangles inchangés, il vaut donc $k = \frac{12}{8} = 1,5$.

En déduire la hauteur *h* des rectangles.

Les densités sont calculées en divisant par 1,5 les effectifs de toutes les classes, sauf la première et la dernière qui ont une amplitude doublée : on divise donc ici par 3 (car $9,5 - 6,5 = 17 - 14 = 3$). Les hauteurs sont calculées à partir des densités, en appliquant un coefficient de proportionnalité choisi pour que l'histogramme ressemble au précédent (on aurait pu choisir n'importe quel coefficient).

Classes	[6,5;9,5[[9,5;11[[11;12,5[[12,5;14[[14;17[
Effectifs	11	12	19	9	9
Amplitude	3	1,5	1,5	1,5	3
Densités= <i>d</i>	$11 \div 3 \approx 6,7$	8	$19 \div 1,5 \approx 12,7$	$9 \div 1,5 = 6$	$9 \div 3 = 3$
Hauteur= <i>h</i>	$6,7 \times 1,5 = 5,5$	12	$19 \div 1,5 \times 6 \times 1,5 = 19$	$3 \times 1,5 = 4,5$	



Pour le premier et le dernier rectangle, aucun problème.

Leur aire est égale à la somme des deux rectangles qu'il remplace. Nous avons tracé le polygone des effectifs, l'équivalent de celui que nous avons tracé plus haut. Ce polygone joint les points d'abscisse : le centre de classe et d'ordonnée : l'effectif (dans le cas des classes de même amplitude) ou la densité (cas des classes d'amplitudes variables).

II] Série ayant des classes d'amplitudes inégales

a) La sécurité routière étudie l'accidentologie des passagers des véhicules de tourisme, âgés de 18 à 65 ans. Le tableau suivant indique le nombre de tués par tranche d'âge en 2005.

► Calculer les densités et les hauteurs de manière à ce que la classe de densité maximale soit représentée par un rectangle de hauteur 30.

Tracer ensuite l'histogramme de la série.

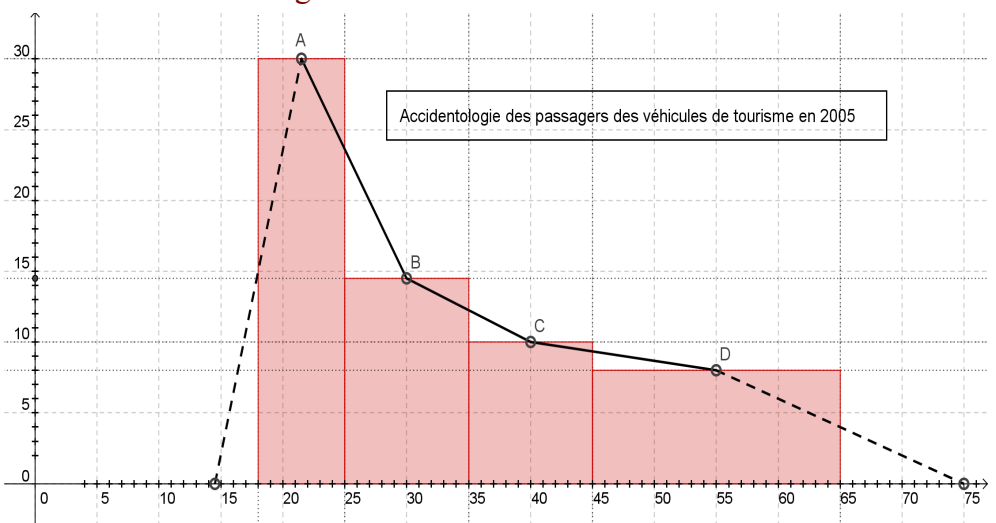
Âge	[18;25[[25;35[[35;45[[45;65[
Amplitude	7	10	10	20
Effectif	790	545	377	606
Densités	112.9	54.5	37.7	30.3
Hauteur	30	14.5	10	8.1

► Tracer alors l'histogramme de la série et le polygone des effectifs (au dos de la feuille). (déterminer un coefficient de proportionnalité tel que la classe de plus grande densité soit représentée par une hauteur de 30)

Pour la réalisation de l'histogramme, il n'y a pas de difficulté particulière. On applique ici un coefficient de proportionnalité égal à $30 \div 112,9$ ou plutôt, pour être exact, à $30 \times \frac{7}{790} = \frac{21}{79} \approx 0,2658$. De cette manière, la classe de plus grande densité (la première classe) aura une hauteur de 30, et on peut s'aider des graduations de l'axe des ordonnées pour trouver les autres hauteurs des rectangles.

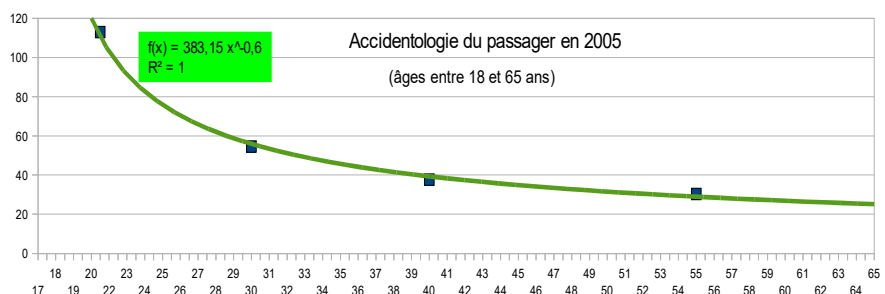
Le polygone des effectifs est plus difficile à déterminer, si on veut qu'il conserve la propriété de l'égalité des aires en dessous de la courbe avec celle des rectangles.

Le principe de la classe d'amplitude nulle à droite et à gauche peut être conservé, sauf qu'il n'a pas de réalité intrinsèque : comme il s'agit ici des passagers ici (et non des conducteurs), rien ne s'oppose à ce que l'âge des tués (je sais ce n'est pas gai..) aille jusqu'à 14,5 ans sauf que les statistiques indiquent les valeurs pour des passagers de plus de 18 ans. Que signifierai donc réellement 14,5 ans dans ce contexte? Pour pallier à cette difficulté, il conviendrait de ne pas prolonger le polygone au-delà des points A et D obtenus au sommet des vrais rectangles.



La réalité du phénomène étudié est une variation continue (pas des segments artificiels, mais une courbe plus ou moins régulière). Pour l'obtention de cette courbe, il faudra attendre un peu pour la théorie, mais un tableau peut en donner l'allure. On représente uniquement les quatre points A, B, C et D et l'on demande au tableur de réaliser une courbe qui passe au mieux par ces points (on a le choix entre affine : une droite, exponentiel, puissance et logarithmique).

La réalité du phénomène étudié est une variation continue (pas des segments artificiels, mais une courbe plus ou moins régulière). Pour l'obtention de cette courbe, il faudra attendre un peu pour la théorie, mais un tableau peut en donner l'allure. On représente uniquement les quatre points A, B, C et D et l'on demande au tableur de réaliser une courbe qui passe au mieux par ces points (on a le choix entre affine : une droite, exponentiel, puissance et logarithmique).

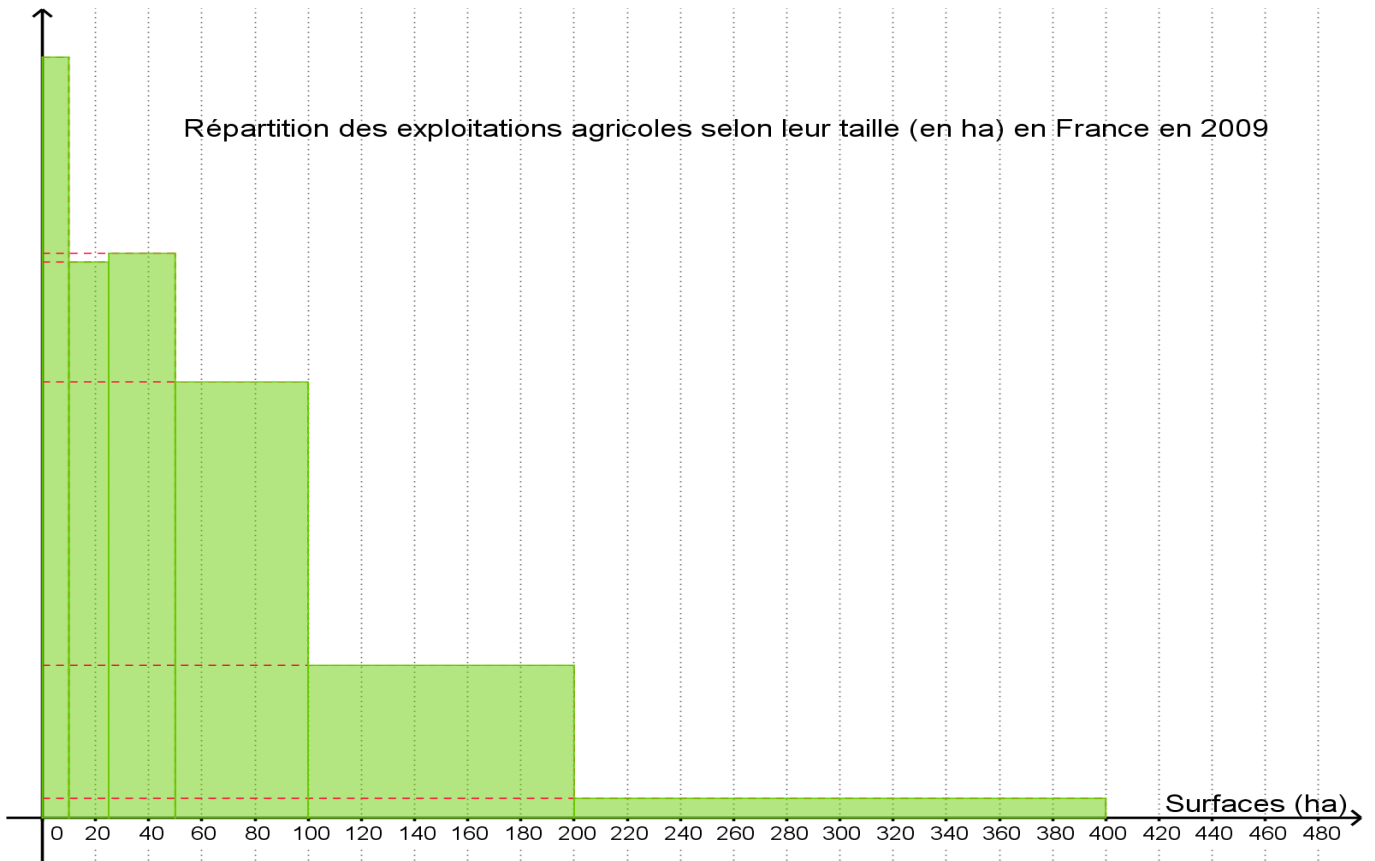


Taille exploitation	Effectif (milliers)	Amplitude	Densité	Hauteur
moins de 10 ha	34,9	10	3,49	20
de 10 à 25 ha	38,2	15	2,55	14,59
de 25 à 50 ha	64,8	25	2,59	14,85
de 50 à 100 ha	100,2	50	2	11,48
de 100 à 200 ha	69,5	100	0,7	3,98
plus de 200 ha	18,7	200	0,09	0,54

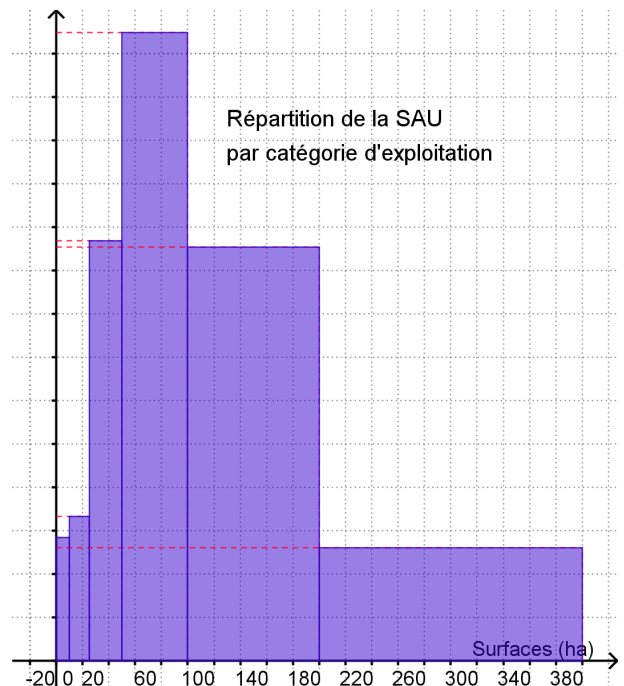
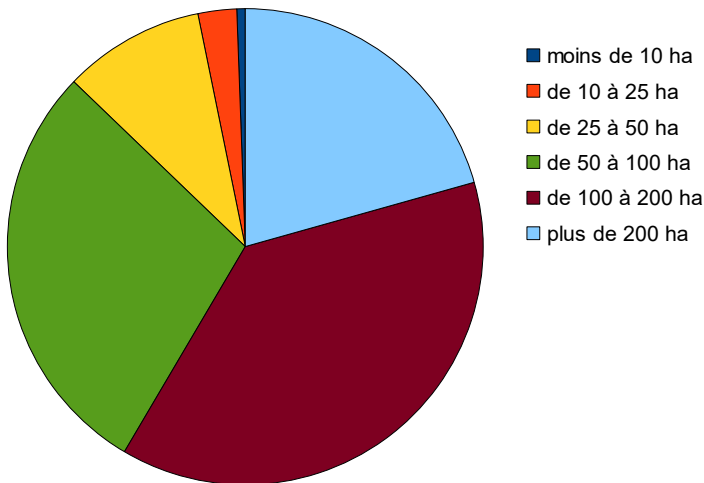
b) En 2009, la répartition des exploitations agricoles selon leur taille (en ha) en France était celle du tableau ci-contre (exercice 51 page 242).

Compléter ce tableau et tracer un histogramme pour représenter ces données (au dos de la feuille). (prendre 400 ha maximum pour la dernière classe)

Répartition des exploitations agricoles selon leur taille (en ha) en France en 2009



Cette représentation ne tient compte que des effectifs. Si on veut représenter la répartition de la Surface Agricole Utile, on aurait pu tracer un diagramme circulaire ou un autre diagramme en rectangles (en bleu), avec des aires de rectangles proportionnelles aux SAU. La part de SAU (Surface totale d'une classe divisée par la surface totale) attachée à chaque catégorie d'exploitation a été, pour cela, calculée préalablement. Il ne reste plus qu'à déterminer la hauteur du rectangle ou l'angle du secteur, selon le type de diagramme que l'on souhaite tracer.



L'avantage de ces représentations est de donner à chaque catégorie l'importance qui lui revient, en terme de surfaces cette fois, et non d'effectifs comme dans l'histogramme vert.