

**Exercice 1 — Vos cartables sont-ils trop lourds ?**

“ Enfin, il convient de rappeler que le chef d’établissement et les représentants des partenaires intéressés à la vie des élèves peuvent, lors du conseil d’administration, examiner les solutions pratiques à mettre en oeuvre, le cas échéant, pour aboutir à une réduction du poids des cartables (qui ne devrait pas dépasser 10% du poids moyen des élèves) ; on peut citer comme exemple la mise à disposition des élèves de casiers individuels fermant à clé.

BO n°39 du 26 octobre 1995, <http://c5eassoc.chez-alice.fr/diver2.html> ”

Proposez une étude pour répondre à la question en titre, puis réalisez-la afin d’y répondre.

**Exercice 2** (page 6–9 du livre *Exploratory data analysis* (1977) de John W. Tukey)

John W. Tukey a donné son nom au diagramme de Tukey (qu’on appelle plus communément diagramme en boîte à moustaches, qu’on verra à la fin de ce chapitre, exercices 5 à 8). Dans son livre, il propose une manière de représenter les données : le diagramme en tiges et feuilles.

Démarrons avec des publicités pour voitures d’occasion tirées du *Sunday Standard-Times* de New Bedford, Massachusetts, daté du 18 août 1968. Dans ces 18 publicités, on peut trouver entre autres des prix pour des Chevrolets d’occasion. Voici les prix demandés dans ces publicités (en dollars) :

250; 150; 795; 895; 695; 1699; 1499; 1099; 1693; 1166; 688; 1333; 895; 1775; 895; 1895; 795

Voici à quoi ressemblerait un diagramme en tiges et feuilles pour cet ensemble de données, si on choisit d’organiser nos tiges par centaines de dollars (dans cet exemple, le fait de marquer deux étoiles montre qu’on a coupé par centaines, et qu’il reste donc deux chiffres à lire dans la colonne des feuilles, pour chaque prix ; on n’est pas obligés de mettre des étoiles partout, Tukey a choisi d’en mettre à plusieurs endroits pour que ça soit plus clair, sans surcharger l’écriture) :

1**		50
2		50
3		
4		
5**		
6		95, 88
7		95, 95
8		95, 95, 95
9**		
10		99
11		66
12		
13**		33
14		99
15		
16		99, 93
17**		75
18		95

Expliquez avec des mots ce que représente la ligne du tableau qui correspond à la tige 16.

**Exercice 3** (page 14 du livre de Tukey)

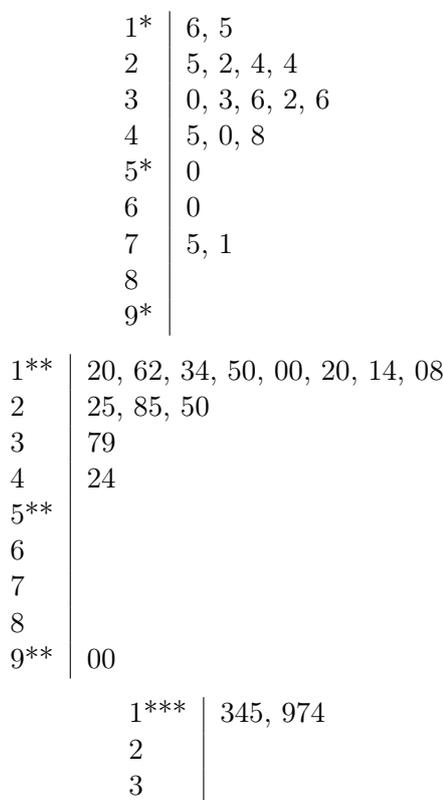
Parfois, les données ne sont pas adaptées pour que toutes les tiges soient similaires. Dans l'exemple à la page 9 du diaporama ([http://www.barsamian.am/2023-2024/S4P6/Chap5\\_diaporama.pdf](http://www.barsamian.am/2023-2024/S4P6/Chap5_diaporama.pdf)), une tige par heure était très adapté. Dans l'exemple de l'exercice 2, une tige par centaine de dollars était adapté. Cela se voit sur le diagramme qui est régulièrement rempli.

Parfois, et même souvent, les données ne sont pas régulièrement espacées. On a alors parfois recours à un diagramme en tiges et feuilles où les tiges n'ont pas toutes la même signification. En exemple, voici une liste de puissances de centrales hydroélectriques des États-Unis (en megawatts) :

100; 1345; 50; 162; 150; 36; 250; 114; 24; 1974; 40; 45; 285; 33; 15; 75; 225;

30; 16; 379; 120; 32; 108; 22; 25; 48; 24; 36; 900; 134; 120; 71; 424; 60

Pour ces données, Tukey choisit de montrer d'abord les petits nombres (entre 0 et 99) avec une première partie où les tiges représentent des dizaines, puis les nombres moyens (entre 100 et 999) avec une seconde partie où les tiges représentent des centaines, puis le reste avec une dernière partie où les tiges représentent des milliers :



Représenter ces mêmes données avec un diagramme en tiges et feuilles où les tiges sont uniquement des centaines. Pourquoi a-t-on une moins bonne idée des valeurs contenues dans la série avec cette seconde représentation ?

#### Exercice 4 — Altitudes aux États-Unis (<https://geology.com/state-high-points.shtml>)

Réaliser un diagramme en tige et feuilles pour les données suivantes.

Indication : commencer par chercher une bonne manière d'organiser les tiges.

State	Highest Point	Altitude (feet)
Alabama	Cheaha Mt., AL	2 405
Alaska	Mt. McKinley (Denali), AK	20 237
Arizona	Humphreys Peak, AZ	12 633
Arkansas	Magazine Mt., AR	2 753
California	Mt. Whitney, CA	14 494
Colorado	Mt. Elbert, CO	14 433
Connecticut	Mt. Frissell-S slope, CT	2 380
Delaware	Ebright Azimuth, DE	448
Florida	Britton Hill, FL	345
Georgia	Brasstown Bald, GA	4 784
Hawaii	Mauna Kea, HI	13 796
Idaho	Borah Peak, ID	12 662
Illinois	Charles Mound, IL	1 235
Indiana	Hoosier Hill Point, IN	1 257
Iowa	Hawkeye Point, IA	1 670
Kansas	Mt. Sunflower, KS	4 039
Kentucky	Black Mt., KY	4 139
Louisiana	Driskill Mt., LA	535
Maine	Mt. Katahdin, ME	5 267
Maryland	Backbone Mt., MD	3 360
Massachusetts	Mt. Greylock, MA	3 487
Michigan	Mt. Arvon, MI	1 979
Minnesota	Eagle Mt., MN	2 301
Mississippi	Woodall Mt., MS	806
Missouri	Taum Sauk Mt., MO	1 772
Montana	Granite Peak, MT	12 799
Nebraska	Panorama Point, NE	5 424
Nevada	Boundary Peak, NV	13 140
New Hampshire	Mt. Washington, NH	6 288
New Jersey	High Point, NJ	1 803
New Mexico	Wheeler Peak, NM	13 161
New York	Mt. Marcy, NY	5 344
North Carolina	Mt. Mitchell, NC	6 684
North Dakota	White Butte, ND	3 506
Ohio	Campbell Hill, OH	1 549
Oklahoma	Black Mesa, OK	4 973
Oregon	Mt. Hood, OR	11 239
Pennsylvania	Mt. Davis, PA	3 213
Rhode Island	Jerimoth Hill, RI	812
South Carolina	Sassafras Mt., SC	3 560
South Dakota	Harney Peak, SD	7 242
Tennessee	Clingmans Dome, TN	6 643
Texas	Guadalupe Peak, TX	8 749
Utah	Kings Peak, UT	13 528
Vermont	Mt. Mansfield, VT	4 393
Virginia	Mt. Rogers, VA	5 729
Washington	Mt. Rainier, WA	14 410
West Virginia	Spruce Knob, WV	4 861
Wisconsin	Timms Hill, WI	1 951
Wyoming	Gannett Peak, WY	13 804

Remarque : si vous vérifiez les données via le lien fourni, attention : les données proviennent des États-Unis, donc la virgule représente le séparateur de 3 chiffres, et pas le séparateur décimal (j'ai ici remplacé ces virgules par des espaces).

### Exercice 5

Une maternité a étudié les tailles de bébés nés à terme au cours d'une année. Elle constate que le plus petit bébé mesure 44 cm, que le plus grand mesure 54 cm. Un quart des bébés mesurent plus de 52 cm et les trois-quarts plus de 47 cm. De plus, la moitié des bébés mesurent plus de 49 cm. Représenter ces données par un diagramme en boîte à moustaches.

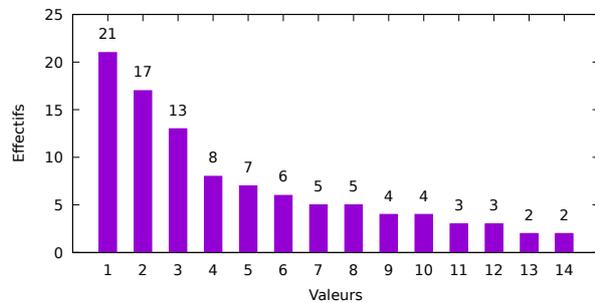
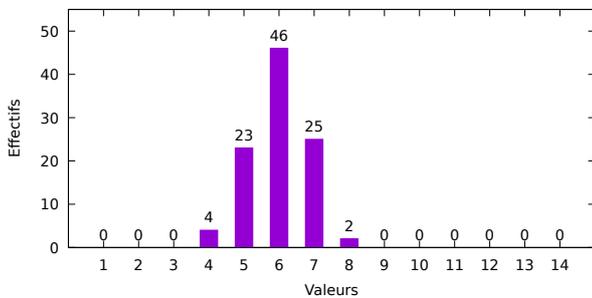
### Exercice 6

Représenter la série suivante par un diagramme en boîte à moustaches :

1; 2,5; 3; 4,2; 5,3; 7; 9; 10,2; 12; 15; 17; 20; 21,7; 25; 27; 50; 54; 60; 63

### Exercice 7

On considère les deux séries de données données par les histogrammes ci-dessous. Compléter le tableau puis réaliser un diagramme en boîte à moustaches de ces deux séries en utilisant la même échelle pour les deux diagrammes.



	Moyenne	Étendue	Médiane	Q1	Q3
Série 1					
Série 2					

### Exercice 8

Associer ces trois diagrammes en boîte à moustaches aux histogrammes.

