

## TIGE et FEUILLE ou encore STEM and LEAF

En 1977, dans son ouvrage *Exploratory Data Analysis*, John W. TUKEY a indiqué une méthode pour organiser ou représenter un ensemble de données numériques qu'il a appelée STEM and LEAF et que l'on traduit en Français par TIGE et FEUILLE.

Il s'agit d'une disposition originale des données numériques qui peut être utilisée pour

- dépouiller ces données,
- archiver ces données,
- donner une représentation semi-graphique de ces données.

Expliquons la méthode sur un exemple :

Soit l'ensemble de données suivant qui représente l'ensemble des tailles, en cm, des élèves d'une classe de BTSA :

145	172	161	170	155	161	169	154	168	172
161	175	155	158	156	156	170	173	183	170
197	178	175	165	160	178	167	161	165	181

Nous allons ranger ces nombres en 2 temps :

1°) chaque nombre est décomposé en 2 parties :

145 ----->	14	et	5
	partie principale		feuille

On range sur une même ligne (appelée TIGE) tous les nombres ayant la même partie principale (*starting part*). La partie principale est inscrite une seule fois en début de ligne, les feuilles sont inscrites au fur et à mesure du dépouillement. On obtient ainsi :

14*	5
15*	5 5 8 6 6 4
16*	1 1 1 5 0 8 1 5 7 9
17*	2 0 5 8 5 8 3 0 2 0
18*	3 1
19*	7

On sépare les parties principales des feuilles par une ligne verticale.

2°) On réécrit le tableau en rangeant, sur chaque tige, les feuilles dans l'ordre croissant, d'où la disposition finale :

4*	5
15*	4 5 5 6 6 8
16*	0 1 1 1 1 5 5 7 8 9
17*	0 0 0 2 2 3 5 5 8 8
18*	1 3
19*	7

On remarque que l'on obtient ainsi une sorte de diagramme en bâtons horizontal de la série statistique.

## Avantages du TIGE et FEUILLE

C'est une méthode pratique, simple et efficace pour réaliser un dépouillement en ordonnant les données.

Cette méthode donne rapidement, et sans outil de dessin, un aperçu graphique d'une série statistique sans perte d'information sur les valeurs numériques.

Cette méthode permet d'obtenir facilement les quartiles d'une série statistique :

Médiane : la série précédente comporte 30 valeurs, la médiane est donc la demi somme des valeurs des 15ème et 16ème individus, ces valeurs étant rangées dans l'ordre croissant. Il suffit donc de compter, à partir du haut du tige et feuille, les 15ème et 16ème feuilles sont 7 et 8 sur la tige 16 d'où la médiane **me = 167,5 cm**

Quartiles d'ordre 1 et 3 : il suffit de compter 8 feuilles à partir du haut du tige et feuille **Q1 = 160 cm**  
et 8 feuille à partir du bas du tige et feuille **Q2 = 173 cm**

Cette méthode peut être utilisée pour comparer visuellement deux séries statistiques : dans la série précédente, les 12 premières valeurs correspondent aux tailles des filles de la classe, les autres valeurs correspondent, bien sûr, aux tailles des garçons. On peut construire le double TIGE et FEUILLE suivant :

5	14*	
8 6 6 5 5	15*	4
1 1 1	16*	0 1 5 5 7 8 9
5 2 0	17*	0 0 2 3 5 8 9
	18*	1 3
	19*	7

Remarques : il existe plusieurs variantes du TIGE et FEUILLE selon l'ordre de grandeur des données à traiter ainsi que la plus ou moins grande dispersion des données : par exemple, si l'on a la série suivante :

1453 1128 937 472 1175 1017 1215 1183  
1036 1260 1151 1238 1189 1310 1247

on prendra des feuilles à 2 chiffres :

4**	72,
5**	
6**	
7**	
8**	
9**	37,
10**	17, 36,
11**	28, 51, 75, 83, 89,
12**	15, 38, 47, 60,
13**	10,
14**	53,

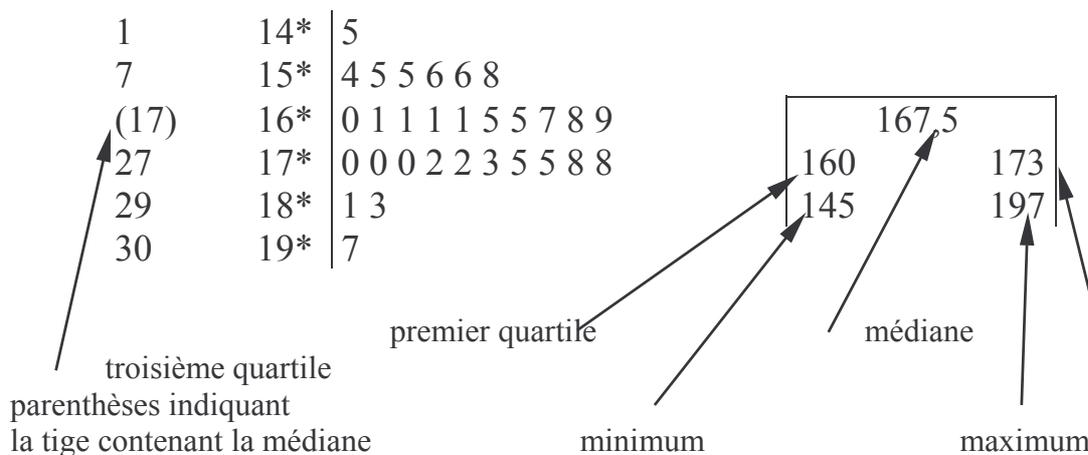
On constate que la valeur 472 se « détache » du reste de la série.

On peut aussi arrondir à la dizaine tous les nombres et représenter alors le nombre de dizaines (ce qui revient à changer d'unité). Il y a dans ce cas une légère perte d'information :

4*	7	unité 10
5*		
6*		
7*		
8*		
9*	4	
10*	2 4	
11*	3 5 7 8 9	
12*	1 4 5 6	
13*	1	
14*	5	

On peut rajouter sur le diagramme certains renseignements statistiques :

- les effectifs cumulés croissants,
- une marque indiquant la classe qui contient la médiane,
- un résumé en 5 nombres de la série statistique.



Ce résumé en 5 nombres permet de faire une représentation graphique de la série appelée Boîte à moustaches, imaginée toujours par le même TUKEY dans son ouvrage cité en début d'article.

#### REFERENCES :

- Exploratory Data Analysis      John W. TUKEY  
ADDISON-WESLEY Publishing Company
- Les certitudes du hasard      Arthur. ENGEL  
ALEAS Editeur
- Statistiques concepts et méthodes      Sabin Lessard Monga  
Masson Editeur