

Définition 3.4.3.b : étendue inter quartile

L'**étendue inter quartile** est la distance séparant le troisième quartile du premier. $EtendueIQ = Q3 - Q1$

Pour [POINTS], l'étendue inter quartile est $30,25 - 14 = 16,25$. Pour la variable [B₅], l'étendue inter quartile est $10 - 7 = 3$.

Variable	Nature	Q0	Q1	Q2	Q3	Q4	Étendue	Inter quartile
[RESULTAT]	Ordonnée	Perdu	Perdu	Gagné	Gagné (bonus)	Gagné (bonus)	<NA>	<NA>
[POINTS]	Continue	0	14	21,5	30,25	42	0	16,25
[CARTON]	Discrète	0	0	1	1	3	3	1
[METEO]	Nominale	<ND>	<ND>	<ND>	<ND>	<ND>	<NA>	<NA>
[COMMENTAIRE]	Ordonnée	Ultra nul	Mauvais	Moyen	Moyen	Excellent	<NA>	<NA>

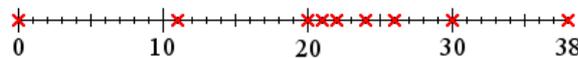
TAB. 3.4.3.a – RCV, les quartiles et étendue

3.4.4 Représentation graphique : la boîte à moustache

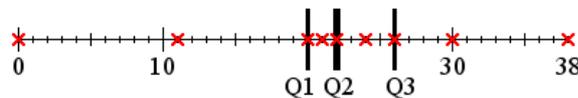
Dans le cas d'une variable quantitative, la médiane et les quartiles se prêtent à une représentation graphique particulièrement astucieuse, la **boîte à moustache** (ou **box plot en anglais**). Cette boîte représente d'un seul coup la centralité, la dispersion et met en valeur d'éventuelles points extrêmes à surveiller. Pour la tracer, nous avons besoin des quartiles Q1, Q2, Q3 et de nouveaux concepts propres aux boîtes à moustache : les barrières et les adhérences. Nous les introduirons au fur et à mesure de nos besoins.

Pour commencer et pour plus de simplicité, abandonnons momentanément nos Villeneuvois pour nous intéresser à leur grand rivaux de toujours : Le State de Labarthe. Leur variable [POINT] vaut {0,11,20,21,22,24,26,30,38} (9 match, déjà ordonnées).

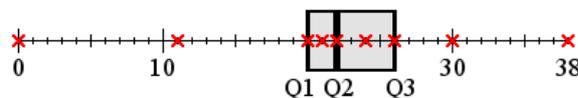
Pour commencer, plaçons sur une règle graduée allant du minimum 0 au maximum 38 (avec les observations)



un trait vertical large à l'emplacement de la médiane, des traits verticaux à l'emplacement des deux quartiles Q1=20 et Q3=26



et joignons les quartiles Q1 à Q3 par des segments :



La boîte ainsi obtenue a la propriété de contenir 50% des observations puisqu'elle délimite l'espace inter quartile.

Il nous reste à tracer les moustaches. Pour cela, nous avons besoin des barrières :

Définition 3.4.4.a : barrière inférieure

La **barrière inférieure** est une valeur imaginaire en dessous de laquelle une observation est considérée comme extrême. Cette valeur est classiquement fixée à une distance de $1,5 \times EtenduIQ$ au dessous du premier quartile.

$$BarInf = Q1 - 1,5 \times EtenduIQ.$$

Définition 3.4.4.b : barrière supérieure

La **barrière supérieure** est une valeur imaginaire au dessus de laquelle une observation est considérée comme extrême. Cette valeur est classiquement fixée à une distance de $1,5 \times EtenduIQ$ au dessus du troisième quartile.

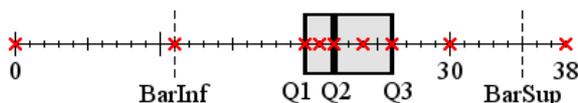
$$BarSup = Q3 + 1,5 \times EtenduIQ.$$

Définition 3.4.4.c : points extrêmes

Les **points extrêmes** sont les observations dont les valeurs sont inférieure à la barrière inférieure et supérieure à la barrière supérieure.

Les barrières sont donc :

- Inférieure : $BarInf = Q1 - 1,5 \times EtenduIQ = 20 - 1,5 \times 6 = 11$
- Supérieure : $BarSup = Q3 + 1,5 \times EtenduIQ = 26 + 1,5 \times 6 = 35$.



Mais ces barrières ne sont là que pour nous permettre de placer les valeurs adhérentes :

Définition 3.4.4.d : valeur adhérente inférieure

La **valeur adhérente inférieure** est la plus petite des observations non extrêmes.

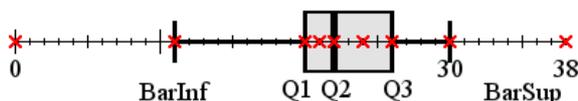
$$ValAdhInf = \min\{Obs \geq BarInf\}$$

Définition 3.4.4.e : valeur adhérente supérieure

La **valeur adhérente supérieure** est la plus grande des observations non extrêmes.

$$ValAdhSup = \max\{Obs \leq BarSup\}$$

Graphiquement, la valeur adhérente supérieure est donc l'observation la plus proche de la barrière supérieure tout en étant situé sur sa gauche (soit 30 sur notre exemple), la valeur adhérente inférieure est l'observation la plus proche de la barrière inférieure et situé sur sa droite (soit 11). Les "moustaches" sont des traits verticaux situés sur les valeurs adhérentes et reliés aux quartiles par une ligne :



La boîte à moustache est maintenant terminée :



La même chose faite par le logiciel "R" :