Mesures de position

Définition: une mesure de position nous renseigne sur la position d'une donnée par rapport aux autres données d'une distribution **ordonnée**.

Voici les mesures de position les plus utilisées.

Les quartiles

Ils séparent la distribution ordonnée en 4 groupes égaux (contient le même nombre de données).

Ils sont symbolisés par Q_1, Q_2, Q_3. Autrement dit, vous allez trouver 3 médianes.

Chaque groupe contiendra 25% des données.

Représentation:

Exemple: Voici une distribution ordonnée

1 3 4 7 9 11 11 13 15

$\boxed{Q_1 = 3{,}5}$ $\boxed{Q_2 = 9}$ $\boxed{Q_3 = 12}$

Procédure:

1- Commencer par trouver Q_2. n=9 Impair. Donc, (n+1)/2 ==> 10/2=5

Q_2 se retrouvera à la 5ième position.

Q_2 =9

2- Par la suite, trouver Q_1 en vous concentrant sur les 4 premières positions car la cinquième position est prise par Q_2. n=4 Pair. Donc, n/2 ==> 4/2 = 2.

Nous allons faire la moyenne entre la 2ième et la 3ième donnée.

$(3+4)/2 = 3{,}5$

Q_1 =3,5

3- Par la suite, trouver Q_3 en vous concentrant sur les 4 dernières positions car la cinquième position est prise par Q_2. n=4 Pair. Donc, n/2 ==> 4/2 = 2.

Nous allons faire la moyenne entre la 2ième et la 3ième données à partir de Q_2.

$(11+13)/2 = 12$

Q_2 = 12

<u>Données éloignées (aberrantes)</u>

Ces données fausses la représentation réelle de la distribution.

Formule :

$x < Q_1 - 1,5(Q_3 - Q_1)$

$x > Q3 + 1,5(Q_3 - Q_1)$

Il suffit de rejeter ces données lors de la construction du diagramme de quartile.

Diagramme de quartiles:

Le diagramme de quartiles se compose de deux rectangles et de deux tiges et met en évidence le minimum, le maximum, Q_1, Q_2, Q_3 et les données éloignées.

Il nous renseigne sur la dispersion et la concentration des données.

Exemple 1 :

Voici une distribution de données :

20 40 60 65 65 75 80 90 95 100

$X_{min} = 20$

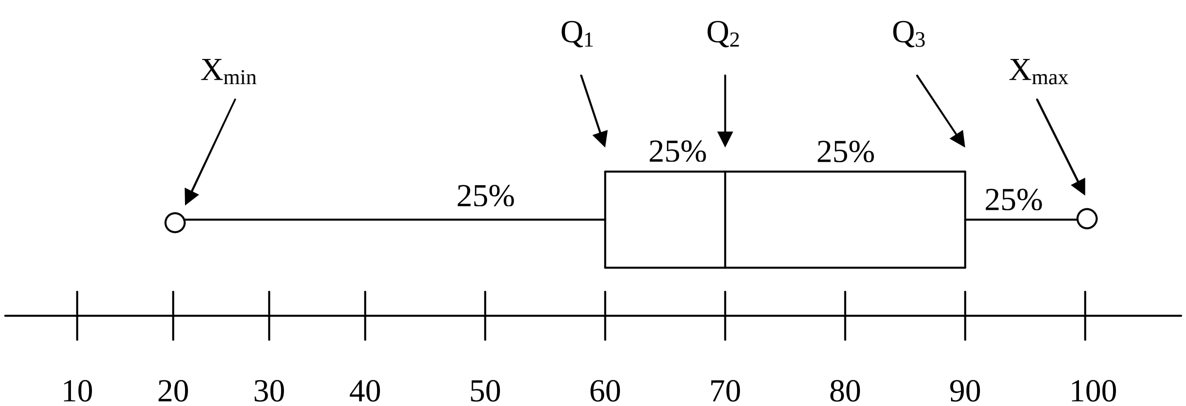
$Q_1 = 60$

$Q_2 = 70$

$Q_3 = 90$

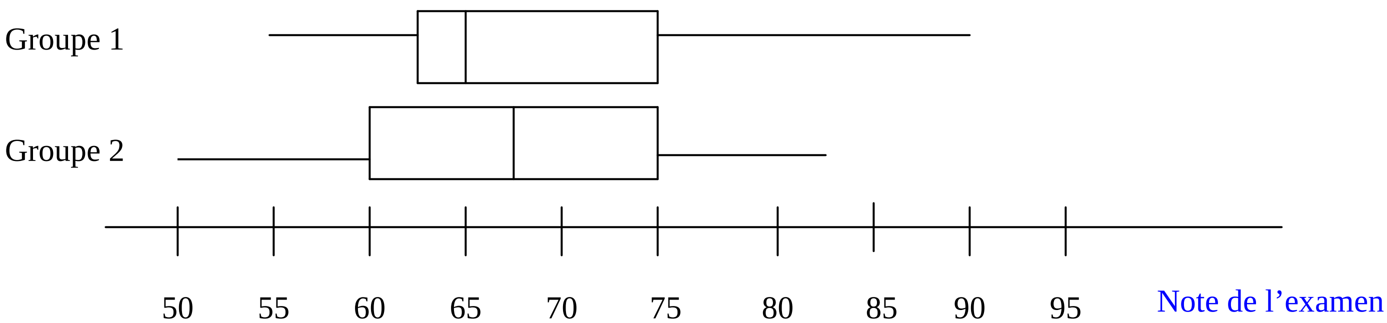
$X_{max} = 100$

Représentation du diagramme de quartile



Exemple 2 :

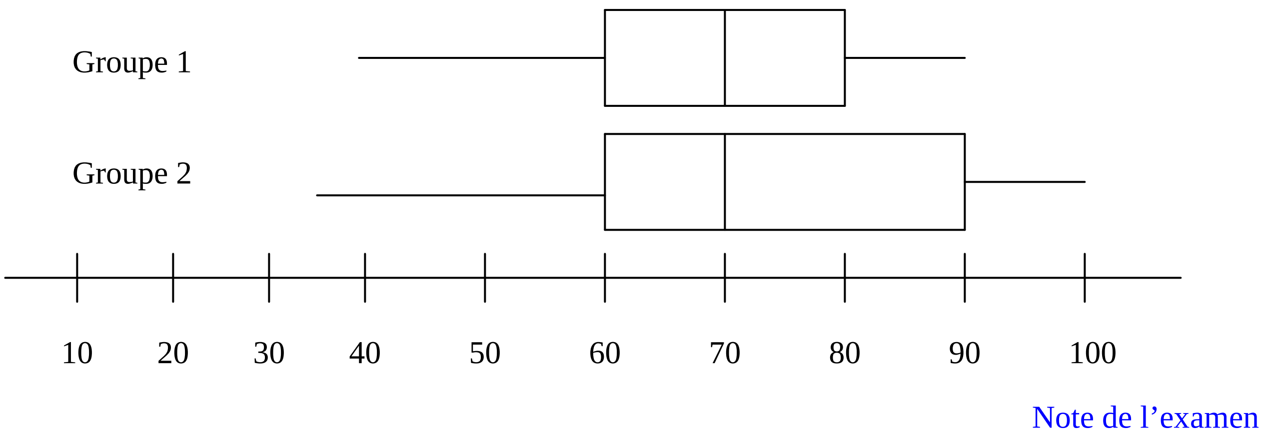
Voici deux diagrammes de quartiles représentant les résultats sur 100 à un examen de français dans deux groupes différents.



a) Dans quel groupe les notes sont-elles les moins dispersées?
   a. Groupe 2 car l’étendue est plus petite (32,5 contre 35 pour le groupe 1)
b) Dans quel groupe se trouve la médiane la plus élevée?
   a. Groupe 2 avec 67,5
c) Dans quel groupe se trouve l’élève qui a le mieux réussi?
   a. Groupe 1 avec 90.
d) La note de passage est de 60. Quel est le pourcentage approximatif des élèves qui ont réussi dans le groupe 2?
   a. Le diagramme de quartiles est séparé en 4 parties à peu près égales. Donc, environ 75%.
e) Le pourcentage des élèves qui ont réussi est-il plus élevé dans le premier groupe?
   a. Oui car le premier quartile est à environ 62,5. Donc le pourcentage de réussite est probablement supérieur à 75%
f) Le pourcentage des élèves qui ont obtenus plus que 75 est-il plus élevé dans le premier groupe?
   a. Non car le troisième quartile dans les deux groupes est 75. Ainsi, dans les deux groupes, il y a 25% des élèves qui ont obtenus plus que 75.

Exercice de consolidation :

Voici deux diagrammes de quartiles représentant les résultats sur 100 à un examen de français dans deux groupes différents.



Questions générales par rapport au diagramme de quartiles ci-dessus

1. Combien de données, en pourcentage, sont inférieures à Q_3 dans le groupe 1?
   a. 75%
2. Dans quel groupe les données sont plus homogènes?
   a. Dans le groupe 1 car l'étendue est la plus petite (E = 50).
3. Dans quel groupe les données sont plus hétérogènes?
   a. Dans le groupe 2 car l'étendue est plus grande (E = 65 environ). Ainsi, les données sont plus dispersées donc plus variées.
4. Dans quel groupe y a-t-il le plus de réussite?
   a. Les deux groupes ont 75% des élèves qui ont réussi.
5. Dans quel groupe l'étendue interquartile est la plus petite?
   a. Dans le groupe 1 avec une EI = 20, car pour le groupe 2, nous avons EI= 30.

Exemple avec données éloignées :

Voici une distribution de données :

10   40   60   65   65   75   80   90   95   100

$x < Q_1 - 1{,}5(Q_3 - Q_1) = 60 - 1{,}5(30) = 15$

Donc toutes les données en bas de 15 ne sont pas incluses dans le diagramme de quartiles. Ainsi, on ne considère pas le 10

$x > Q3 + 1{,}5(Q_3 - Q_1) = 90 + 1{,}5(30) = 135$

Donc toutes les données en haut de 135 ne sont pas incluses dans le diagramme de quartiles. La distribution est correcte à ce niveau.

$X_{min} = 40$

$Q_1 = 60$

$Q_2 = 70$

$Q_3 = 90$

$X_{max} = 100$