

Statistique – Visualisation de données

Somme

python : `sum()`

Usage : `sum(liste)`

Entrée : une liste de nombres

Sortie : un nombre

Exemple : `sum([4,8,3])` renvoie 15

Minimum/maximum

python : `min()`

Usage : `min(liste)` ou `min(a,b)`

Entrée : une liste de nombres ou bien deux nombres

Sortie : un nombre

Exemple :

- `min(12,7)` renvoie 7
- `min([10,5,9,12])` renvoie 5

Variance et écart-type

La **variance** d'une série de données (x_1, x_2, \dots, x_n) est définie comme la moyenne des carrés des écarts à la moyenne. C'est-à-dire :

$$v = \frac{1}{n}((x_1 - m)^2 + (x_2 - m)^2 + \dots + (x_n - m)^2)$$

où m est la moyenne de (x_1, x_2, \dots, x_n) .

Par exemple, pour la série $(6, 8, 2, 10)$, la moyenne est $m = 6.5$, la variance est

$$v = \frac{1}{4}((6 - 6.5)^2 + (8 - 6.5)^2 + (2 - 6.5)^2 + (10 - 6.5)^2) = 8.75$$

L'**écart-type** d'une série (x_1, x_2, \dots, x_n) est la racine carrée de la variance :

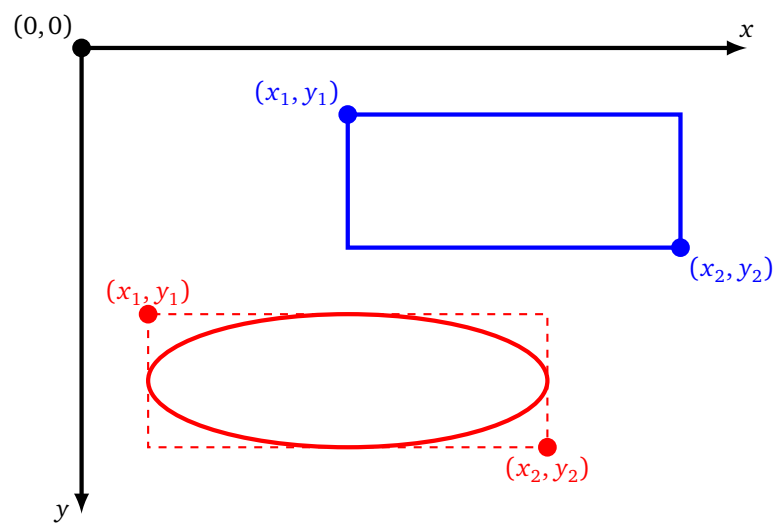
$$e = \sqrt{v}$$

où v est la variance.

Exemple :

$$e = \sqrt{v} = \sqrt{8.75} = 2.95 \dots$$

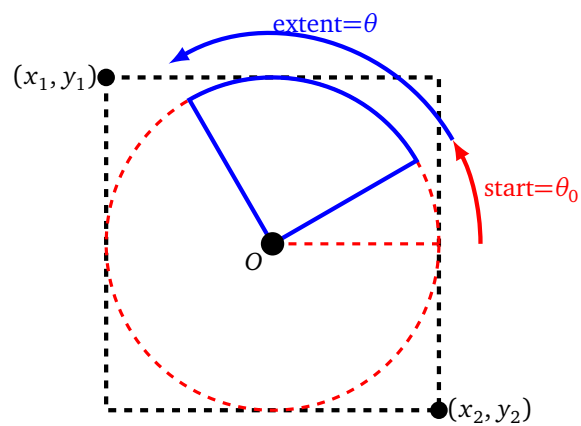
Graphiques avec tkinter



- `create_rectangle(x1,y1,x2,y2)`
- `create_oval(x1,y1,x2,y2)`
- `create_text(x,y,text="Mon texte")`

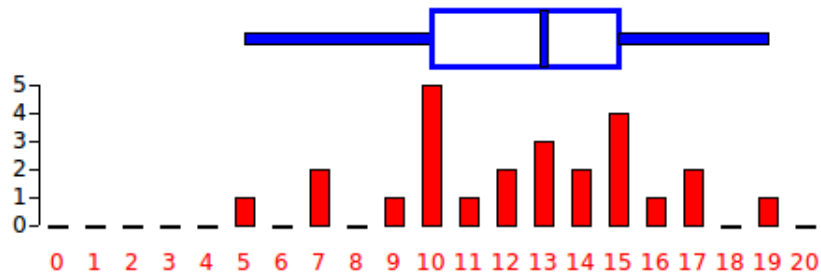
Tracer un arc

```
create_arc(x1,y1,x2,y2,start=debut_angle,extent=mon_angle)
```

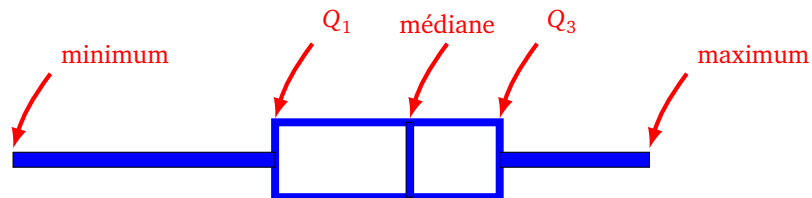


L'option `style=PIESLICE` affiche un secteur au lieu d'un arc.

Diagramme en boîte



Un *diagramme en boîte* (appelé aussi *boîte à moustaches*) est un graphique qui représente les principales caractéristiques d'une série statistique : minimum, maximum, médiane et quartiles. Le schéma de principe est le suivant :



Médiane et quartiles

Par définition, la moitié des valeurs est inférieure ou égale à la médiane, l'autre moitié est supérieure ou égale à la **médiane**.

Rappels. On note n la longueur de la liste et on suppose que la liste est ordonnée (du plus petit au plus grand élément).

- **Cas n impair.** La médiane est la valeur de la liste au rang $\frac{n-1}{2}$.
Exemple avec `liste = [12, 12, 14, 15, 19]` :
 - la longueur de la liste est $n = 5$ (les indices vont de 0 à 4),
 - l'indice du milieu est l'indice 2,
 - la médiane est la valeur `liste[2]`, c'est donc 14.
- **Cas n pair.** La médiane est la moyenne entre la valeur de la liste au rang $\frac{n}{2} - 1$ et au rang $\frac{n}{2}$.
Exemple avec `liste = [13, 14, 19, 20]` :
 - la longueur de la liste est $n = 4$ (les indices vont de 0 à 3),
 - les indices du milieu sont 1 et 2,
 - la médiane est la moyenne entre `liste[1]` et `liste[2]`, c'est donc $\frac{14+19}{2} = 16.5$.

Quartiles

Les quartiles répartissent les valeurs en : un quart en-dessous de Q_1 , un quart entre Q_1 et Q_2 , un quart entre Q_2 et Q_3 , un quart au-dessus de Q_3 . Pour le calcul, on utilise que :

- Q_2 est simplement la médiane de la liste entière (supposée ordonnée),
- Q_1 est la médiane de la sous-liste formée de la première moitié des valeurs,
- Q_3 est la médiane de la sous-liste formée de la seconde moitié des valeurs.

Effectifs

Les résultats d'une classe sont collectés sous la forme suivante d'un effectif par note :

```
effectif_notes = [0,0,0,0,0,1,0,2,0,1,5,1,2,3,2,4,1,2,0,1,0]
```

Le rang i va de 0 à 20. Et la valeur au rang i indique le nombre d'élèves ayant eu la note i .

Écris une fonction `notes_vers_liste(effectif_notes)` qui prend en entrée un effectif de notes et renvoie la liste des notes. Pour notre exemple la fonction doit renvoyer `[5,7,7,9,10,10,10,10,10,10,...]`.