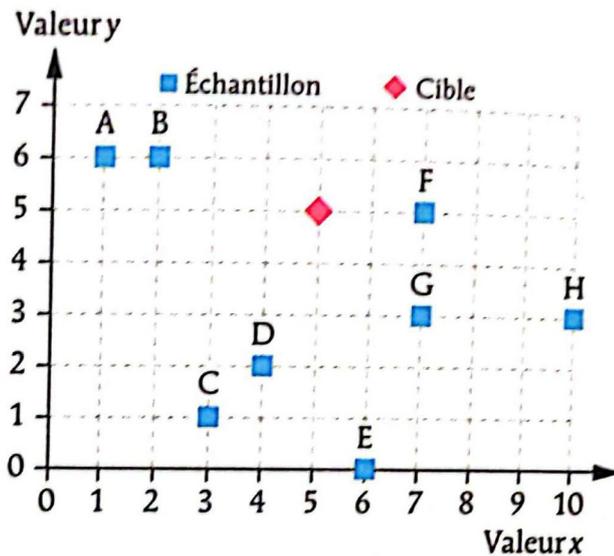


## Exercices

### Exercice 1

On donne la représentation d'un échantillon (points bleus nommés A, B...) et d'une cible (point rouge).



1. En utilisant la distance euclidienne, quels sont les deux plus proches voisins de la cible ?

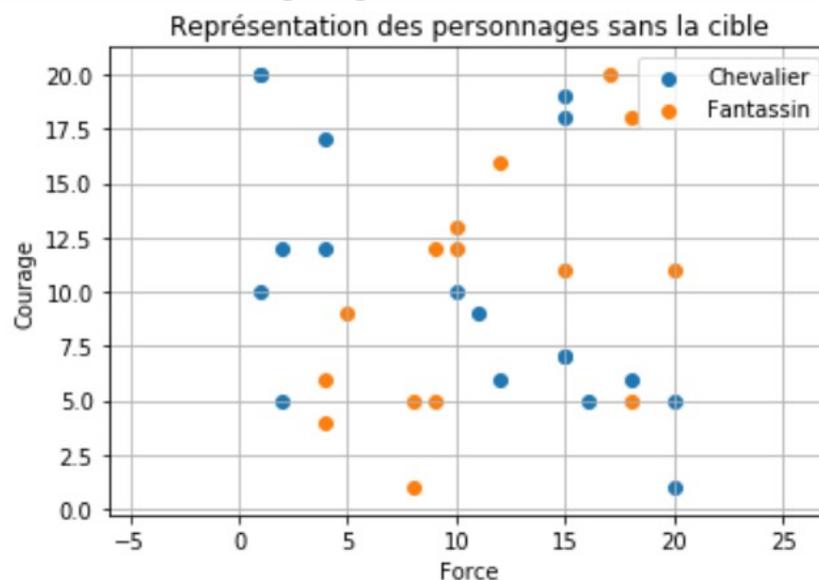
2. On suppose que les valeurs de y ne sont pas pertinentes. En utilisant la distance entre les valeurs de x ( $d = |x_1 - x_2|$ ), quels sont les plus proches voisins de la cible ?

### Exercice 2 : Jeu de rôle

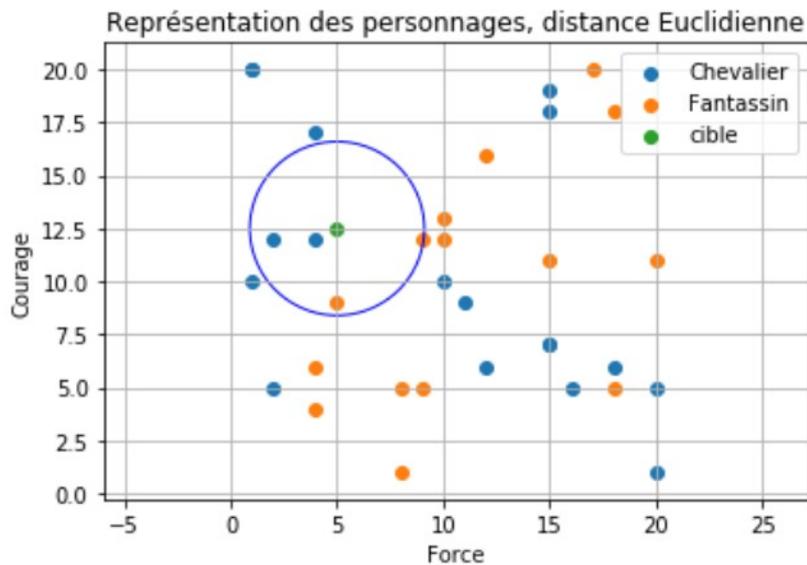
Considérons un jeu de rôle médiéval avec des chevaliers et des fantassins.

Le jeu génère des soldats aléatoirement et il faut leur donner une spécialité d'après deux de leurs caractéristiques (Courage et force) et celles de soldats déjà catégorisés en chevalier ou fantassins.

Voici la représentation des soldats pris pour référence :

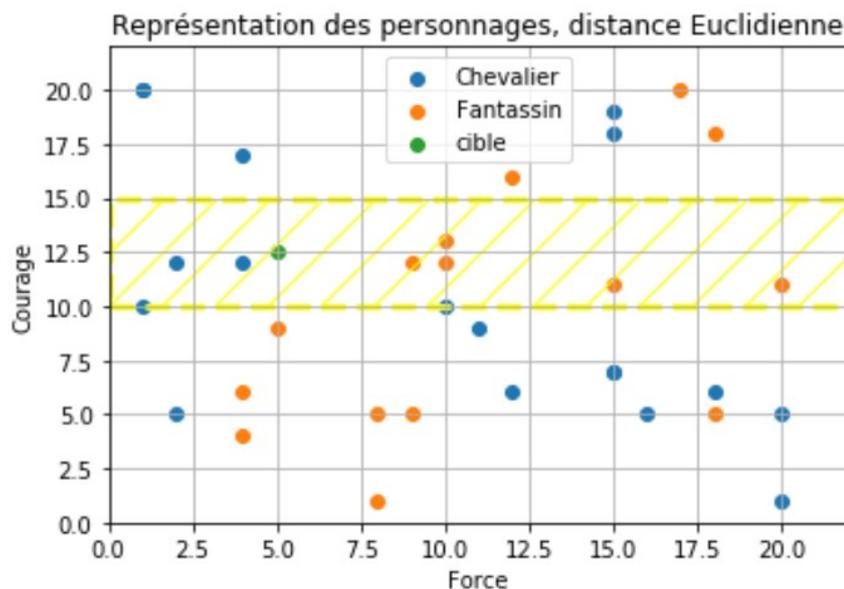


Nous introduisons une cible de force = 5 et de courage = 12,5.  
On choisit  $k = 4$  et la distance est schématisée par un disque :



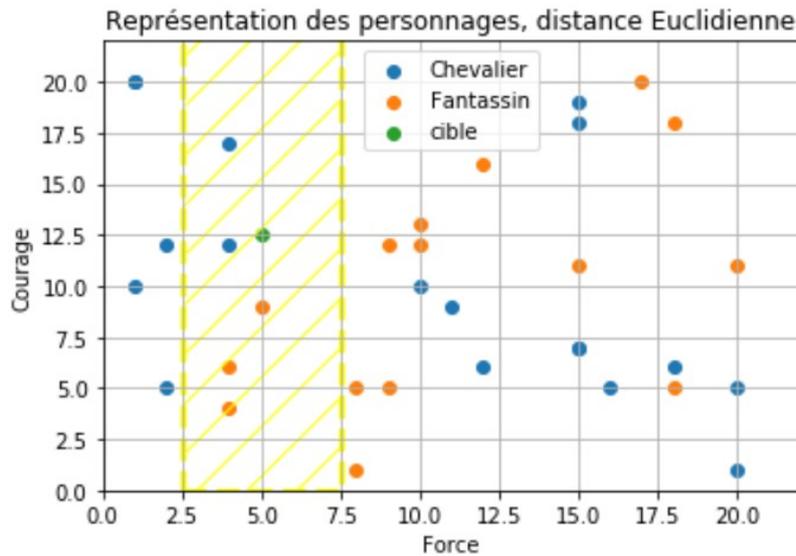
1. Quels sont les  $k$  plus proches voisins de notre donnée cible ?
2. Quel est le problème ?
3. Donner trois valeurs de  $k$  permettant de décider de la catégorie de notre donnée cible ?
4. Donner pour ces trois valeurs de  $k$  la catégorie de la donnée cible.

On choisit  $k = 9$ . Pour la distance, on décide que les valeurs de la force n'ont pas d'importance. La distance dépend donc uniquement du courage. On peut représenter graphiquement cette condition :



5. Quelle est la catégorie de notre donnée cible ?

On choisit  $k = 5$ . Pour la distance, on décide que les valeurs du courage n'ont pas d'importance. La distance dépend donc uniquement de la force. On peut représenter graphiquement cette condition :

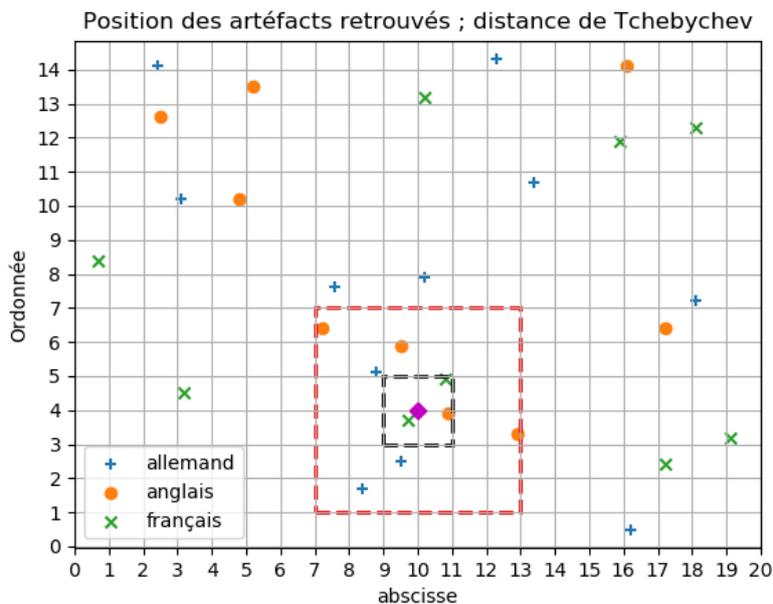


6. Quelle est la catégorie de notre donnée cible ?

**Exercice 3 : Champ de bataille**

Sur un champ de bataille de la Première Guerre Mondiale un mémorial a été construit. Afin de réaliser une extension, des fouilles préventives ont été réalisées par l'INRAP (Institut National de Recherches Archéologiques Préventives). Au cours de ces fouilles, différents objets ou éléments de squelettes humains ont été trouvés. L'étude de ces découvertes a permis d'identifier la nationalité de nombreux artefacts retrouvés : soit allemand, anglais ou français. Le plan ci-dessous représente la zone de fouille et la position des éléments dont l'origine a été identifiée. L'unité est le mètre.

Un élément d'un squelette a été retrouvé en (10; 4). Il est représenté par un losange couleur magenta sur le plan. L'objectif est de déterminer une origine probable pour cet élément de squelette avant de le déposer dans un ossuaire.



La distance qui sera prise en compte est la distance dite de Tchebychev. Pour cet exercice, l'ensemble des points se trouvant à une distance  $r$  d'un point  $I$  correspond au pourtour du carré, de centre  $I$ , de côtés parallèles aux axes et de longueurs  $2r$ .

Sur le graphique ci-dessus, le carré dessiné :

- en rouge correspond ainsi à l'ensemble des points se trouvant à 3 mètres.
- en noir correspond ainsi à l'ensemble des points se trouvant à 1 mètre.

1. A quelle valeur de  $k$  correspond le carré noir ?
2. Quelle serait l'origine du squelette en considérant cette valeur de  $k$  ?
3. On choisit  $k=9$ . Quelle serait l'origine du squelette en considérant ce  $k$  ?
4. On choisit  $k=11$ . Quelle serait l'origine du squelette en considérant ce  $k$  ?
5. Peut-on savoir à coup sûr, en prenant une valeur de  $k$  inférieure au égale à 11, si le combattant dont on a trouvé un squelette était un combattant de la Triple-Entente (France + Royaume-Uni + Russie) ou de la Triple-Alliance (Allemagne + Autriche-Hongrie + Italie) ? Répondre par plusieurs arguments.

#### Exercice 4 : Les pokemons

Le fichier Pokemons.csv contient environ 400 Pokemons, répartis selon 18 types.

Voici un nouveau Pokemon : Hippodocus qui possède les caractéristiques suivantes :

points de vie = 108  
attaque = 112  
défense = 118  
vitesse = 47

1. Quel type pouvez-vous attribuer à ce Pokemon ?
2. Varier les valeurs de  $k$  et les différents types de distances. Le résultat reste-il stable ?
3. Observer la table "pokemon", surtout les effectifs des différents types. Pensez-vous que la table soit efficace pour entrainer une intelligence artificielle à la reconnaissance des types pokemon ?