

Dernière mise à jour	Informatique	C. GAUDY/D. DEFAUCHY Site web
23/08/2021	1 - Recherche séquentielle - Dictionnaire	TD 1-2 – Comptage par dictionnaire

Informatique

1

Recherche séquentielle - Dictionnaire

TD 1-3

Comptage par dictionnaire

Dernière mise à jour	Informatique	C. GAUDY/D. DEFAUCHY Site web
23/08/2021	1 - Recherche séquentielle - Dictionnaire	TD 1-2 – Comptage par dictionnaire

Exercice 1: Comptage par dictionnaire

Question 1: Ecrire une fonction `Mot_to_dico(texte)` qui prend en argument une chaîne de caractère `texte` et renvoie un dictionnaire dont les clés sont les lettres qui apparaissent dans le texte et les valeurs le nombre d'occurrence de ces lettres

Vérifier :

```
>>> Mot_to_dico('ACCTAGCCCTA')
{'A': 3, 'C': 5, 'T': 2, 'G': 1}
```

Vous avez à votre disposition le fichier « 1-3 - TD - Comptage par dictionnaire - Extrait.py » dont [voici le lien](#).

Question 2: Récupérer (copier-coller) la chaîne de caractères extrait dans le fichier et utiliser la fonction `Mot_to_dico` pour retrouver le nombre de « a » dans l'extrait

Pour la suite, on appelle « mot » de k lettres un ensemble de k caractères successifs pris dans une chaîne de caractères.

Question 3: Ecrire une fonction `Compte_mots_2_lettres(texte)` qui, étant donnée une chaîne de caractères `texte`, renvoie un dictionnaire dont les clés sont les mots de 2 lettres de `texte` et les valeurs les occurrences de ces mots

Vérifier :

```
>>> Compte_mots_2_lettres('ACCTAGCCCTA')
{'AC': 1, 'CC': 3, 'CT': 2, 'TA': 2, 'AG': 1, 'GC': 1}
```

Question 4: Ecrire une fonction `Compte_mots_k_lettres(k, texte)` qui, étant donnée une chaîne de caractère `texte` et un entier k , renvoie un dictionnaire permettant de dénombrer tous les sous-mots de k lettres de `texte`.

Question 5: Utiliser la fonction précédente afin de déterminer le nombre de fois où apparaît « CTA » dans la séquence d'ADN suivante, que vous pouvez récupérer dans le fichier ouvert précédemment

```
'ACCTAGCCATGTAGAATCGCCTAGGCTTTAGCTAGCTCTAGCTAGCTG'
```