

Statistiques sur Covid-19

On désire faire des statistiques sur la pandémie actuelle.

Pour chaque pays, on recense les personnes qui ont eu un résultat (Outcome). Le résultat peut être Death 'D' (mort) ou Recovery 'R' (guéri).

On peut organiser les données dans un tableau $T[1..n]$ d'arbres de recherche binaire d'enregistrements dans lequel $T[I]$ est l'arbre associé au pays I.

Une personne est décrite par l'enregistrement suivant : Nom, Prénom, Age, Résultat

Les arbres de recherche binaire sont ordonnés selon l'age. De cette manière, l'arbre de recherche binaire peut contenir plusieurs exemplaires de l'age.

1. A partir de données générées aléatoirement créer la structure.

Pour chaque pays, le nombre de personnes (m) et les différents champs des enregistrements sont tous générés aléatoirement.

2. Afficher la structure par pays et ordonnée selon l'age

En vue de faire des statistiques globales, on regroupe tous les arbres dans un même arbre de recherche binaire ordonné cette fois-ci selon l'age et le résultat. De plus, l'age n'est pas dupliqué.

Chaque nœud de l'arbre est alors le triplet (Age, Occurrence, Résultat) où Occurrence désigne le nombre de personnes avec l'age Age. Les noms et prénoms sont ainsi éliminés.

3. Créer la nouvelle structure à partir de l'ancienne.

4. Afficher la nouvelle structure

5. Développer sur la nouvelle structure un seul algorithme efficace permettant de répondre aux requêtes suivantes:

- % de 'D' tels que l'age est dans l'intervalle [50, 60]
- % de 'D' tels que l'age est supérieur à 70
- % de 'R' tels que l'age est inférieur à 30
- Etc

Partie Khawarizm

Exprimez tous les modules nécessaires puis les mettre en œuvre sous l'environnement Khawarizm

Prenez $n = 10$, $m \leq 20$.

Partie C

(a) Demander la traduction automatique vers C et tester le programme résultant.

(b) Sur le programme C obtenu :

Récupérer sur un fichier TXT les données réelles sur les pays (ND=Nombre de résultats 'D', NR=Nombre de résultats 'R') à partir d'internet. Ce fichier contiendra alors 3 colonnes.

(Vous pouvez le faire à partir du site <https://www.worldometers.info/coronavirus/>)

Les ages, noms et prénoms restent aléatoires.

Afin de ne pas modifier la structure initiale, rajouter un tableau de pays Pays[1..n]. T[I] représente alors Pays [I]. Limiter n à 100 en considérant uniquement les pays les plus atteints par la maladie.

- * Dans le code C, au lieu de générer m aléatoirement, le prendre du fichier TXT (ND + NR). De plus, il faudra générer ND personnes avec le résultat 'D' et NR personnes avec le résultat 'R'.
- * Rajouter des modules qui améliorent l'affichage des structures.
- * Rajouter un module qui affiche de manière conviviale les résultats des requêtes.

Note importante : Le fichier TXT doit être référencé dans le programme **sans le chemin**, ce qui implique qu'il doit figurer dans le répertoire du programme.

N.B

1. Le tableau d'arbre est déclaré comme suit : *T un tableau(10) de arb de (chaine, chaine, entier, car)*
2. L'arbre global est déclaré comme suit : *arb de (entier, entier, car)*
3. Vous pouvez utiliser les fonctions
 - *Longchaine (S)* : donne la longueur du mot (S)
 - *Caract(S, I)* : accède au Ième caractère du mot S.
 - *Aleanombre(N)* : génère un entier entre 0 et N-1.
 - *Aleachaine(N)* : génère une chaîne de caractères entre 1 et N caractères.

Il est inutile d'essayer de bien présenter les résultats avec Khawarizm. Il n'est pas fait pour ce but.

Notation

Khawarizm : 15 pts

C : 5 pts

Passage au langage C

Il est peu probable que lors de la traduction, le programme obtenu ne tourne pas. Dans ce cas, l'étudiant devra faire l'effort pour chercher et corriger l'erreur. Les principales erreurs sont dues principalement aux initialisations implicites de variables dans l'environnement Khawarizm.

Toute erreur reconnue pertinente et corrigée par l'étudiant sera récompensée par un Bonus. Dans ce cas, nous envoyez le code Z pour lequel la traduction ne fonctionne pas et le code C corrigé en nous indiquant les changements. Si une erreur de traduction ne peut être corrigée, veuillez nous envoyer le code Z. L'erreur sera corrigée pour vous.

Attention au plagiat !!!

Date de remise : 31 mai 2020 avant Minuit

- * Les TP seront remis selon un formulaire qui vous sera transmis par la suite
- * Tout TP non remis selon l'échéancier fixé entraîne un retrait de 2 points par jour sur un total de 3 jours tolérés.
- * **Un TP non fait entraîne un retrait de 2 points sur la moyenne générale!**