

Groupes sanguins

Fin janvier 2019, un groupe d'amis a décidé de faire un don du sang.



Voici une tableau regroupant les données personnelles que les médecins leur ont demandé avant leur prise de sang :

Numéro_Donneur	Nom	Prénom	Date_de_Naissance	CP Ville
1	Roux	Emmanuelle	13/01/03	62610
2	Masson	Emeline	13/03/04	62100
3	Astier	Cécile	01/03/03	62100
4	Boulot	Céline	29/06/03	62137
5	Boulot	Juliette	29/06/03	62137
6	Campagne	David	16/05/03	62610
7	Bernard	Jean-Pierre	12/05/03	62100
8	Lefebvre	Claire	21/02/02	62100
9	Mayeux	Eric	13/06/02	62137
10	Garcia	Michel	20/06/03	62137
11	Naïdji	Antoine	13/04/03	62730
12	Van Helsen	Baptiste	03/03/02	62610

On appelle descripteur le nom associé aux données d'une colonne.

Un identifiant est un descripteur permettant d'identifier de manière unique un objet (ou une personne dans le cas présent).

- Quels sont les descripteurs utilisés ici ?
- Certains de ces descripteurs sont-ils un identifiant pour ces données ? Si oui, lesquels ?
- Classer ce tableau par ordre alphabétique des noms.

Nom	Prénom	Numéro_donneur	Date de Naissance	CP Ville

Après leur passage, un médecin a récapitulé les résultats des analyses dans un tableau en les anonymisant. L'un des amis a réussi à en prendre une photographie avec son téléphone.

Age	Sexe	Ville	Groupe sanguin
16	F	Coulogne	A
16	G	Ardres	A
17	F	Calais	A
16	F	Ardres	AB
16	G	Marck	AB
16	G	Coulogne	B
17	G	Coulogne	B
17	G	Ardres	B
15	F	Calais	B
16	G	Calais	O
16	F	Calais	O

- Quels sont les attributs utilisés ici ?
- Certains attributs sont-ils des identifiants ?

Voici une petite liste de codes postaux des alentours de Calais.

Ville	Code postal
Ardres	62610
Audruicq	62370
Calais	62100
Coulogne	62137
Guines	62340
Marck	62730
Sangatte	62231

En croisant les données, retrouver le groupe sanguin de chaque individu.

Vous pourrez vous servir du tableau ci-dessous à compléter :

Age	Année Naiss.	Sexe	Ville	Code postal	Nom-Prénom	groupe sanguin
16		F	Coulogne			A
16		G	Ardres			A
17		F	Calais			A
16		F	Ardres			AB
16		G	Marck			AB
16		G	Coulogne			B
17		G	Coulogne			B
17		G	Ardres			B
15		F	Calais			B
16		G	Calais			O
16		F	Calais			O

Remarques :

Lorsqu'un organisme collecte des données personnelles, même si elles sont anonymées, il est souvent possible, en croisant des données, de les désanonymiser. Une grande partie des individus d'une population est identifiable en ne connaissant que :

- leur date de naissance ou quelques dates précises (date d'entrée/sortie d'un hôpital par exemple).
- leur sexe.
- la zone géographique dans laquelle ils habitent.

Lors de la création d'une base de données confidentielles et sensibles, comme celle des dossiers médicaux, il faut donc choisir avec un grand soin les descripteurs si l'on souhaite qu'il soit impossible, ou du moins très difficile, de désanonymiser ces données, prendre très au sérieux la sécurité de la base et garantir que ces données ne pourront pas être détournée pour un *mauvais* usage.

Avec l'avènement du **R**èglement **G**énéral de la **P**rotection des **D**onnées, les organismes collectant nos données doivent décrire précisément quelles données sont sauvegardées dans leur base, pour quels objectifs ces données sont conservées et garantir leur sécurité. Mais ces précisions sont souvent difficiles d'accès : [Voici par exemple la charte de Doctissimo](#)

Quelques liens ... si vous souhaitez en savoir plus :

[A propos du dossier médical - source wikipédia](#)

[RGPD et secteur de la santé - source CNIL](#)

Licence CC-BY-SA

Mieszczak Christophe, formation académique SNT de l'académie de Lille