

Extrait Mines-Ponts 2016

Pour suivre la propagation des épidémies, de nombreuses données sont recueillies par les institutions internationales comme l'O.M.S. Par exemple, pour le paludisme, on dispose de deux tables :

La table palu recense le nombre de nouveaux cas confirmés et le nombre de décès liés au paludisme; certaines lignes de cette table sont données en exemple (on précise que isoest un identifiant unique pour chaque pays) :

nom	iso	annee	cas	deces
Bresil	BR	2009	309316	85
Bresil	BR	2010	334667	76
Kenya	KE	2010	898531	26017
Mali	ML	2011	307035	2128
Ouganda	UG	2010	1581160	8431
...				

La table demographie recense la population totale de chaque pays; certaines lignes de cette table sont données en exemple :

pays	periode	pop
BR	2009	193020000
BR	2010	194946000
KE	2010	4090900
ML	2011	14417000
UG	2010	33987000
..		

1. Au vu des données présentées dans la table palu, parmi les attributs nom, isoet annee, quels attributs peuvent servir de clé primaire ? Un couple d'attributs pourrait-il servir de clé primaire ? (On considère qu'une clé primaire peut posséder plusieurs attributs). Si oui, en préciser un.

2. Ecrire une requête en langage SQL qui récupère depuis la table palu toutes les données de l'année 2010 qui correspondent à des pays où le nombre de décès dus au paludisme est supérieur ou égal à 1000.

On appelle taux d'incidence d'une épidémie le rapport du nombre de nouveaux cas pendant une période donnée sur la taille de la population-cible pendant la même période. Il s'exprime en général en nombre de nouveaux cas pour 100 000 personnes par année. Il s'agit d'un des critères les plus importants pour évaluer la fréquence et la vitesse d'apparition d'une épidémie.

3. Ecrire une requête en langage SQL qui détermine le taux d'incidence du paludisme en 2011 pour les différents pays de la table palu.

4. Ecrire une requête en langage SQL permettant de déterminer le nom du pays ayant eu le deuxième plus grand nombre de nouveaux cas de paludisme en 2010 (on pourra supposer qu'il n'y a pas de ex-aequo pour les nombres de cas).

Questions supplémentaires

Ecrire les requêtes SQL permettant d'afficher :

5. Le nombre de pays étudiés

6. Le nombre de ces de paludisme au Brésil, ainsi que le nombre total de décès.

7. le nombre de cas de paludisme par année civile de la plus récente à la plus ancienne

8. le nom et nombre de cas total de paludisme de chaque pays

9. le nom des pays dont le nombre de décès moyen annuels dus au paludisme est supérieur à 500, ainsi que la moyenne du nombre de cas et la moyenne du nombre de décès.

Extrait Mines-Ponts 2016 – Éléments de correction

1. Chaque pays pouvant apparaître (avec son code iso) pour plusieurs années, et chaque année pouvant être associée à plusieurs pays, aucun de ces attributs ne peut faire office de clé primaire.

Il est en revanche logique que chaque pays (ou chaque code iso) apparaisse une seule fois par année. Donc le couple (iso,annee) (par exemple) peut faire office de clé primaire.

2.

```
SELECT *  
FROM palu  
WHERE annee = 2010 AND deces >= 1000
```

3.

```
SELECT nom , 1. * cas / pop * 100000  
FROM palu JOIN demographie  
ON iso = pays AND annee = periode  
WHERE annee = 2011  
(rappel : on multiplie 'cas' par 1. pour forcer le calcul en flottant plutôt que la division euclidienne entière)
```

4.

```
SELECT nom  
FROM palu  
WHERE annee = 2010  
EXCEPT  
SELECT nom  
FROM palu  
WHERE cas = (      SELECT MAX(cas)  
                  FROM palu  
                  WHERE annee = 2010 )  
ORDER BY cas DESC  
LIMIT 1
```

Questions supplémentaires

5.

```
SELECT COUNT(DISTINCT(pays))  
FROM demographie
```

6.

```
SELECT COUNT(cas), COUNT(deces)  
FROM palu,  
JOIN demographie  
ON pays = iso  
WHERE nom="Bresil"
```

7.

```
SELECT annee, SUM(cas)  
FROM palu  
GROUP BY annee  
ORDER BY annee DESC
```

8.

```
SELECT SUM(cas)  
FROM palu p  
JOIN demographie  
ON pays = iso  
GROUP BY pays
```

9.

```
SELECT nom, AVG(cas), AVG(deces)  
FROM palu p  
JOIN demographie  
ON pays = iso  
GROUP BY pays  
HAVING AVG(deces)>=500
```