



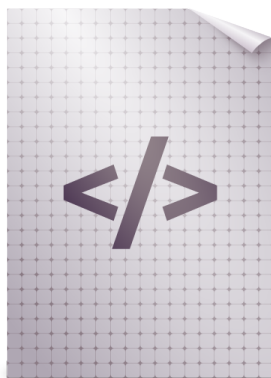
INSA Lyon
20, avenue Albert Einstein
69621 Villeurbanne Cedex

LIVRABLE DE PROJET

Informatique Décisionnelle

« Mortality Data Pack »

du 7 avril au 17 avril 2014



Hexanôme :

Guillaume ABADIE
Thierry CANTENOT
Juliette COURLET
Rémi DOMINGUES
Adrien DUFFY-COISSARD
Ahmed KACHKACH

Enseignants :

M MIQUEL
A TCHOUNIKINE

Année scolaire 2013-2014

Sommaire

1	Analyse des données	1
1.1	Visualisation des données	1
1.2	Séparation des données de sexe et d'age	1
1.3	Incohérences dans les nombre de morts	1

1. Analyse des données

1.1 Visualisation des données

Requête utile pour voir la répartition des causes de décès selon le sexe :

```
SELECT sexage.SEX, causes.LABEL, COUNT(*) AS Number
FROM london INNER JOIN
      sexage ON london.SEXAGE = sexage.SEXAGE INNER JOIN
      causes on london.CAUSE = causes.CAUSE
GROUP BY sexage.sex, causes.label
```

1.2 Séparation des données de sexe et d'âge

Vu le format des données de la table SexAge, dissocier l'effet du sexe et de l'âge s'avérerait hardu.

Nous avons donc modifié la table en créant deux nouvelles colonnes : Sex, nvarchar qui contiendra le sexe (M/F), et Age, nvarchar qui contiendra la tranche d'âge (<1, 10-14, ...).

Nous avons alors lancé la requête suivante afin de séparer les données de "sexage" sur ces deux colonnes :

```
UPDATE sexage
SET Sex=LEFT(label, 1), Age=RIGHT(label, LEN(label) -
1);
```

1.3 Incohérences dans les nombre de morts

En analysant la table des morts de chaque région, nous avons remarqué que la somme des morts par cause de niveau 2 était inférieure au nombre de morts par cause de niveau 1 (càd de toutes les causes de mort).

Cette incohérence peut être montrée sur la table "westmids" (par exemple) avec la requête suivante :

```
SELECT causes.NIVEAU, SUM(westmids.DEATHS) AS TotalDeaths
FROM westmids INNER JOIN causes ON westmids;CAUSE = causes.
    CAUSE
GROUP BY causes.NIVEAU
HAVING (causes.NIVEAU = 1) OR (causes.NIVEAU = 2)
```

Ceci est probablement dû à une erreur lors du remplissage des formulaires (cause de niveau 2 non-remplie).

La solution qu'on propose est de rajouter une nouvelle cause de mort de niveau 2 représentant une cause indéfinie, afin de ne pas sous-estimer le nombre de morts en ignorant ceux dont la cause de niveau 2 n'est pas renseignée.

```
METTRE LES REQUETES QUI VONT BIEN ICI
```