# Le corpus de données

La base de données est fournie par l’institut national des études démographiques (INED) et concerne différents indicateurs démographiques de 196 pays. Ces indicateurs sont des estimations pour l’année 2011 obtenues à partir d’une étude « World Population Prospects. Nations Unies » datant de 2009.

# Signification des caractères

Chaque individu du corpus de données représente un pays qui va être décrit selon 11 variables dont cinq sont quantitatives :

* POP : Population totale (en milliers)
* TNAT : Taux de natalité
* TMORT : Taux de mortalité
* EV : Espérance de vie
* TMORTINF : Taux de mortalité infantile
* NBENF : Nombre d’enfants par femme
* TCR : Taux de croissance
* POP65 : Population ayant plus de 65 ans (en milliers)
* T65 : Taux de population ayant plus de 65 ans (T65)

et deux sont qualitatives :

* CONT : Continent

AFRIQUE – AMERSUD – AMERNORD – EUROPE - OCEANIE – ASIE

* REG : Région

AMERSUD : Amérique du sud et caraibes

AMERNORD : Amérique du nord

AMERCENT : Amérique centrale

AFOR : Afrique Orientale

AFCEN : Afrique centrale

AFSEP : Afrique septentrionale

AFAUS : Afrique australe

AFOCC : Afrique occidentale

EUROEST : Europe de l'est

EUROPNORD : Europe du nord

EUROSUD : Europe du sud

EUROUEST : Europe de l'ouest

AUSTZEL : Australie Nouvelle zélande

ASIEST : Asie de l'est

ASISUDEST : Asie du sud-est

ASICENTSUD : Asie du centre et du sud

ASIOUEST : Asie de l'ouest