## Milestone 1: Project description

* The names of all project members : Nicolas Binuani, Rodolphe Calvet, Sully Dilou, Delin JIA, Zakaria TIZAF
* The topic to be addressed, the datasets to use, where the data come from, and their formats.

**Dataset :** Health\_Europe\_dataset\_somediseases

**Source :** <http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=hlth_cd_anr&lang=en>

**Format :** csv

* A description of the problem
  + Who are the users?
    - OMS - Doctors - Pharmacy companies who want to follow the evolution of some disease depending on age, country or gender and apply or modify health plans.
  + What are their backgrounds?
    - Health domain expertise and maybe some politicians
  + What are they trying to understand from the data?
    - The evolution of some diseases in each country and during a decade to validate or not some health policies taken (example with ALCOVE european program for Alzheimer disease 2011)
  + Is your visualization aimed primarily at *exploring* or *communicating* the data?
    - The first goal is to communicate on the evolution of some diseases in Europe.
    - After that, we can also use it to explore some new insights and change health european programs.
* A description of the data
  + What are the characteristics of the data, the attributes, the size of the dataset, etc.

The dataset gives the total number of deaths by disease, gender, age, and year from 2001 to 2010 in Europe.

It contains 190000 rows and 6 columns.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | GEO | SEX | AGE | Disease | Number |
| 2001 | Belgium | Female | From 25 to 29 years | Pneumonia | 261 |
| ... | ... | ... | ... | ... | ... |

Columns Time(Year) and Value (Number) are integers, the other columns present textual data.

