

# The new French information system on asylum : from a fragile approach to an agile approach

Parallel session 1D : *Innovative uses of administrative data for statistical purposes*

International Forum on Migration Statistics, 15-16 January 2018

Victor Ng, Attaché statisticien de l'Insee  
DSED, Ministère de l'Intérieur, France

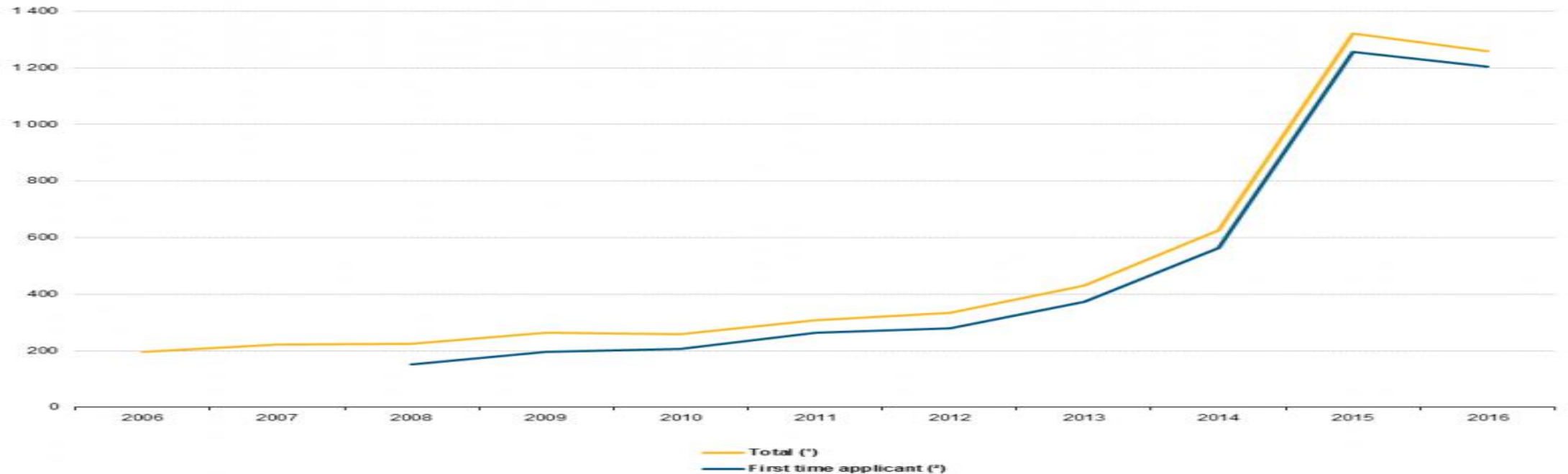


# Overview

- The asylum in France
- The new information system on asylum
- The statistics on asylum
- Solving technical issues
- Further perspectives

# Contextualizing asylum

- 2015 : unseen European migrant crisis
- Over 1,200,000 first applicants



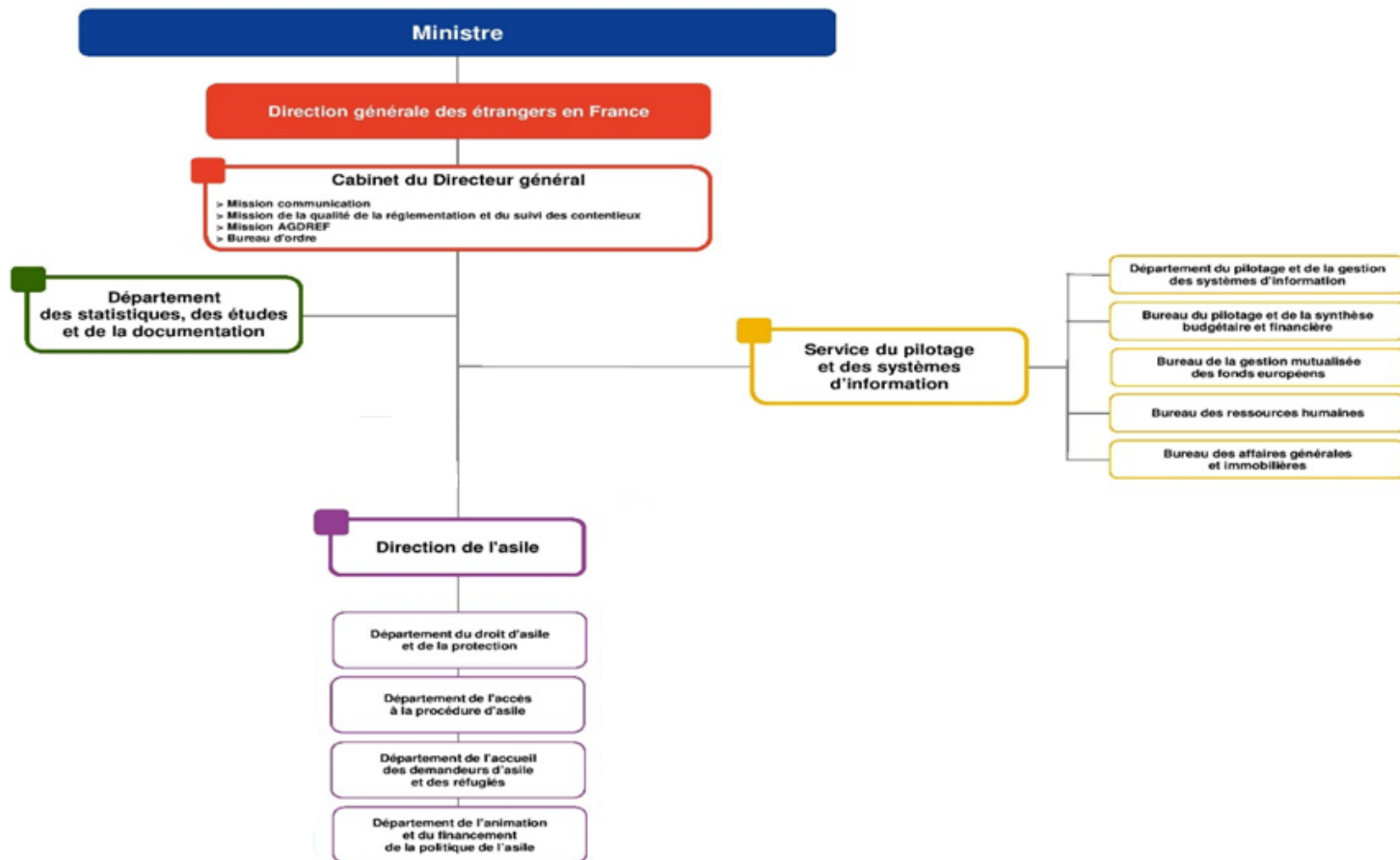
(\*) 2006 and 2007: EU-27 and extra-EU-27.

(\*) 2006 and 2007: not available.

Source: Eurostat (online data codes: migr\_asyctz and migr\_asyappctza)

# Birth of a new information system on asylum

- Necessity for a European settlement
- Transcription into French law 2015-925, 29<sup>th</sup> July 2015
- Birth of a new and effective information system
- Within the French Ministry of Interior:
  - Contracting owner: Direction of asylum
  - Project manager: Service of IT (and its private consultants)
  - Statistical support: Department of statistics



# Statistics on asylum

- Birth in a period of crisis and emergency
- Priority to the management of the asylum seekers...
- ... but necessity to count and to calculate relevant indicators to know who they are
- Prior to this information system and still existing: info collected by the OFPRA

# Towards a Big Data structured information system

## Characteristics of the migrant crisis

- Quick spread amongst European states and quick evolution of asylum policies
- Large number of registered asylum seekers/applications
- Diverse profiles, motivations or migratory routes

## Expected qualities for the information system

- Velocity
- Volume
- Variety

# Two more V's

## Information system

- Variability
- Veracity

## Goal

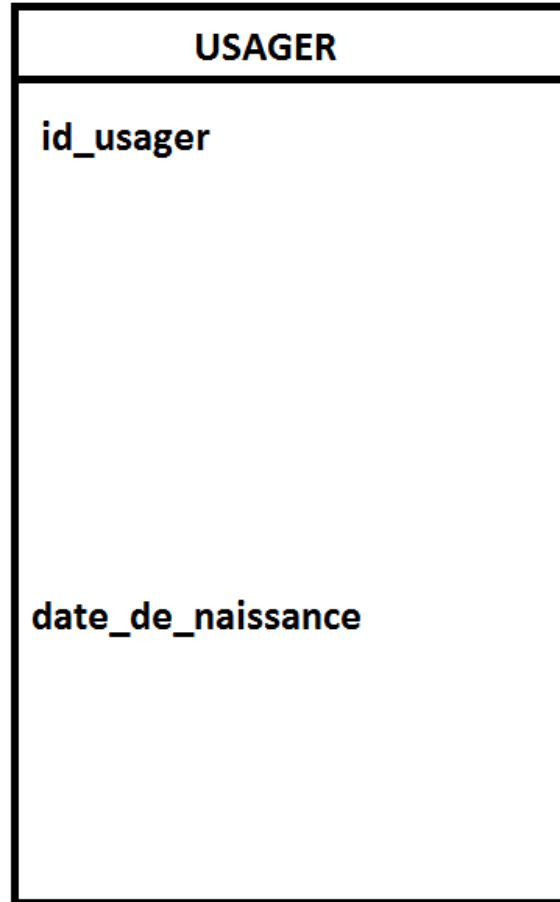
- Responsiveness to changes and reactivity to requests
- Achieving a good level of relevance, accuracy and reliability



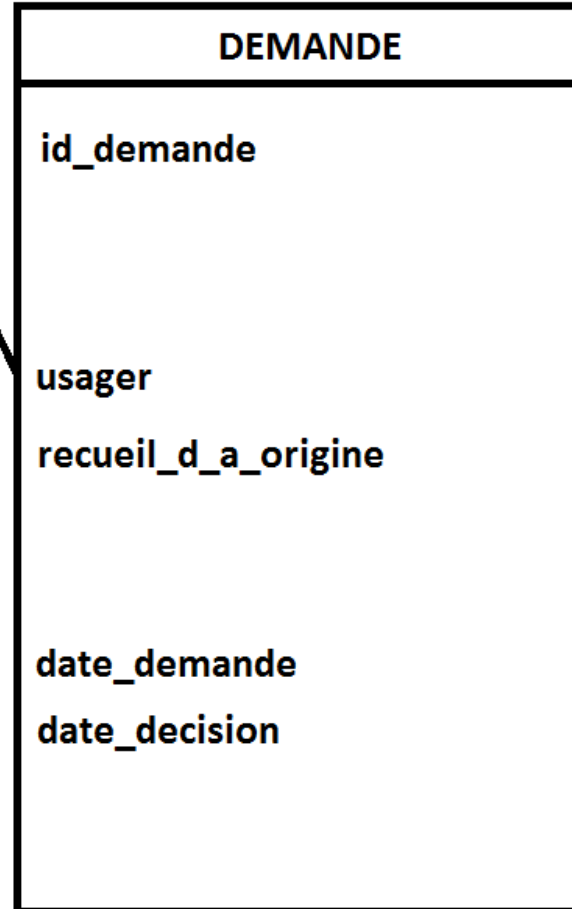
# About the extractions

- Big Data structured : collections with JSON format variables
- Difficult use of classical tools or softwares
- 5 CSV files delivered weekly
  - site.csv: inventory of one-stop-shops and prefectures
  - usager.csv: 1 line per registered user of the asylum system
  - demande\_asile.csv: 1 line per registered asylum application
  - recueil\_d\_a.csv: 1 line per each group of users indicating the family link between them
  - [originally not provided] droit.csv: 1 line per each attestation of rights delivered to an applicant

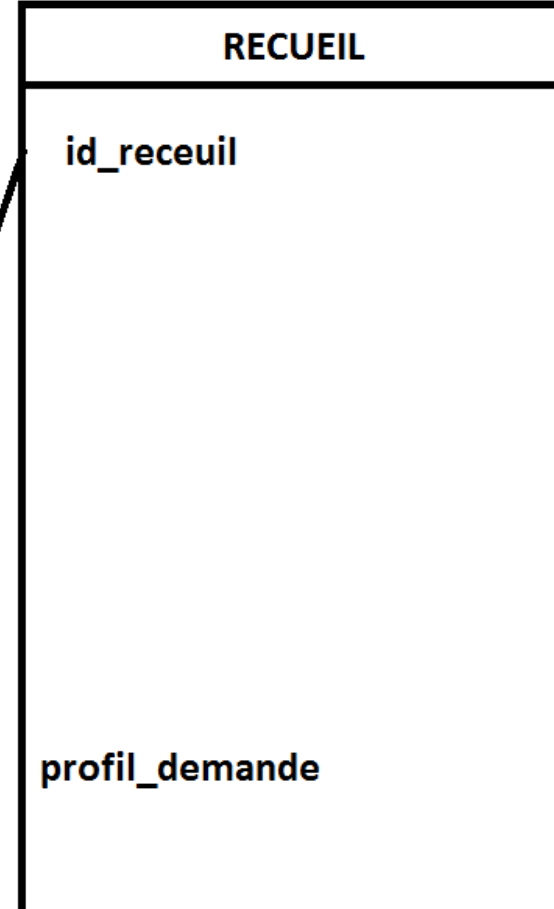
(User)



(Application)



(Links)



# 'Sometimes it's funny what a statistician collects'

- Expectation: 1 id per person and errorless information in variables
- Reality: 1 id per user, 1 id per asylum application, 1 id per family link, 1 id per attestation of rights...

With possible duplicates indicating that a user would have applied multiple times after a period of time (that is not really legal...)

And with inconsistencies and slugs in some variables

- Statistician's job: getting information out of this datasource and make it as clear as possible for the policy maker
- That implies a step-by-step work:
  - understanding what is needed
  - naming and locating the variable(s)
  - defining the indicators and the way to compute them
  - carrying out the computation
  - giving a feedback

# From fragility...

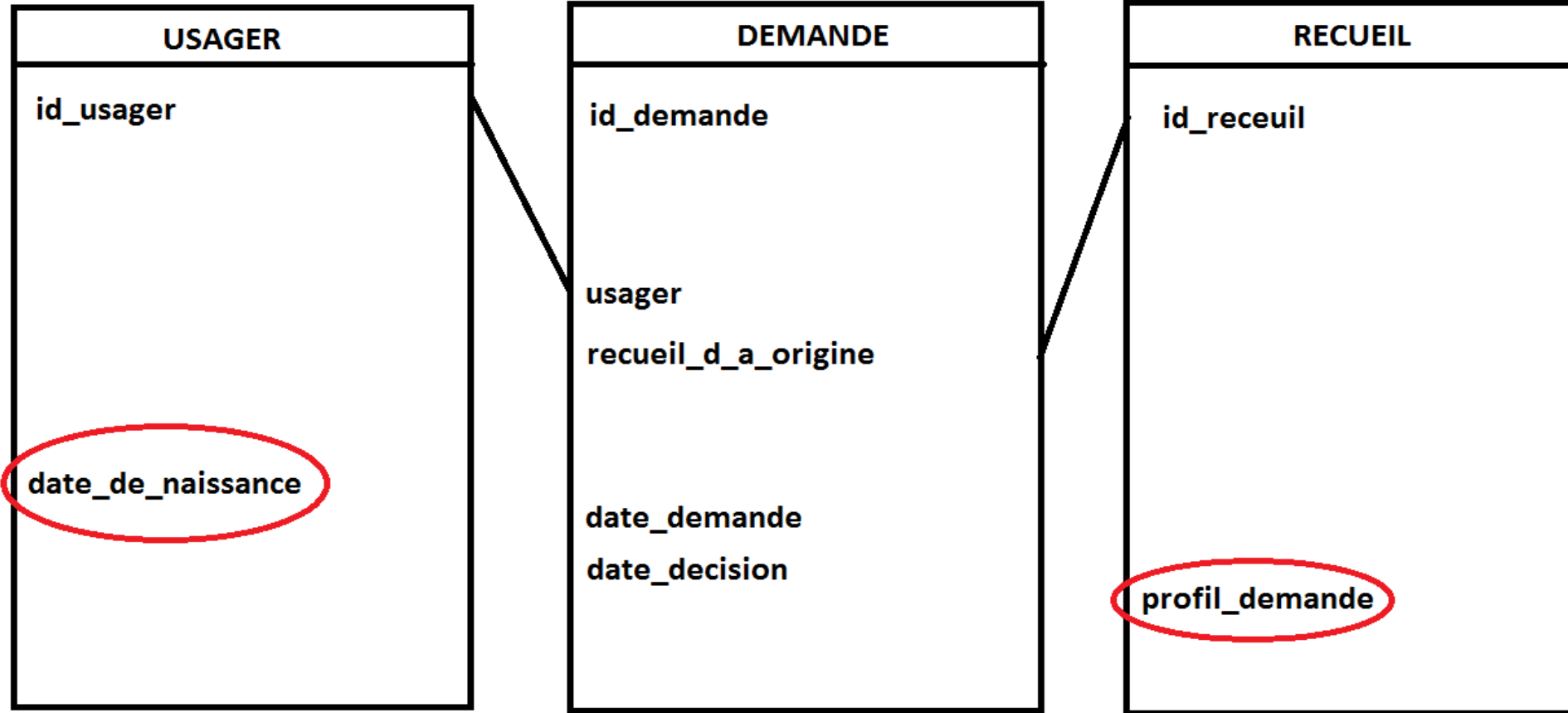
- DBMS : Microsoft Access
  - Inexistence of error log
  - Unconvenient or unpleasant to use: manual and sequential execution
  - Lack of rapidity, optimality or statistical quality
- Weekly or monthly reports to local responsables for asylum: due to the growing volume of data, there could be a lack of timeliness
- Unanswered questions (not limited):
  - How to compute the number of minors?
  - How to deal with inconsistencies, slags or mismatches?
  - What is the proper way to compute the duration between two 'events' for a user or an application?
- We had to address those gaps of knowledge

# ... to agility (or sort of...)

- Agility in ways to find solutions
- Statistical software : SAS
- Minors, episode 1: an unperfect but acceptable (and accepted) rectification
  - Algorithm based first upon calculated age then upon registered profile
  - 2016: reasonable at first glance but with a closer look, too few minors
- Minors, episode 2: the discovery of an existing variable
  - enfants\_presents\_au\_moment\_de\_la\_demande
  - 2016: about 14 000 accompanying minors (close to OFPRA)
- Minors: explanations in the next two slides
- Mismatches with permits files
  - Explanations?

# Algorithm for minors rectification

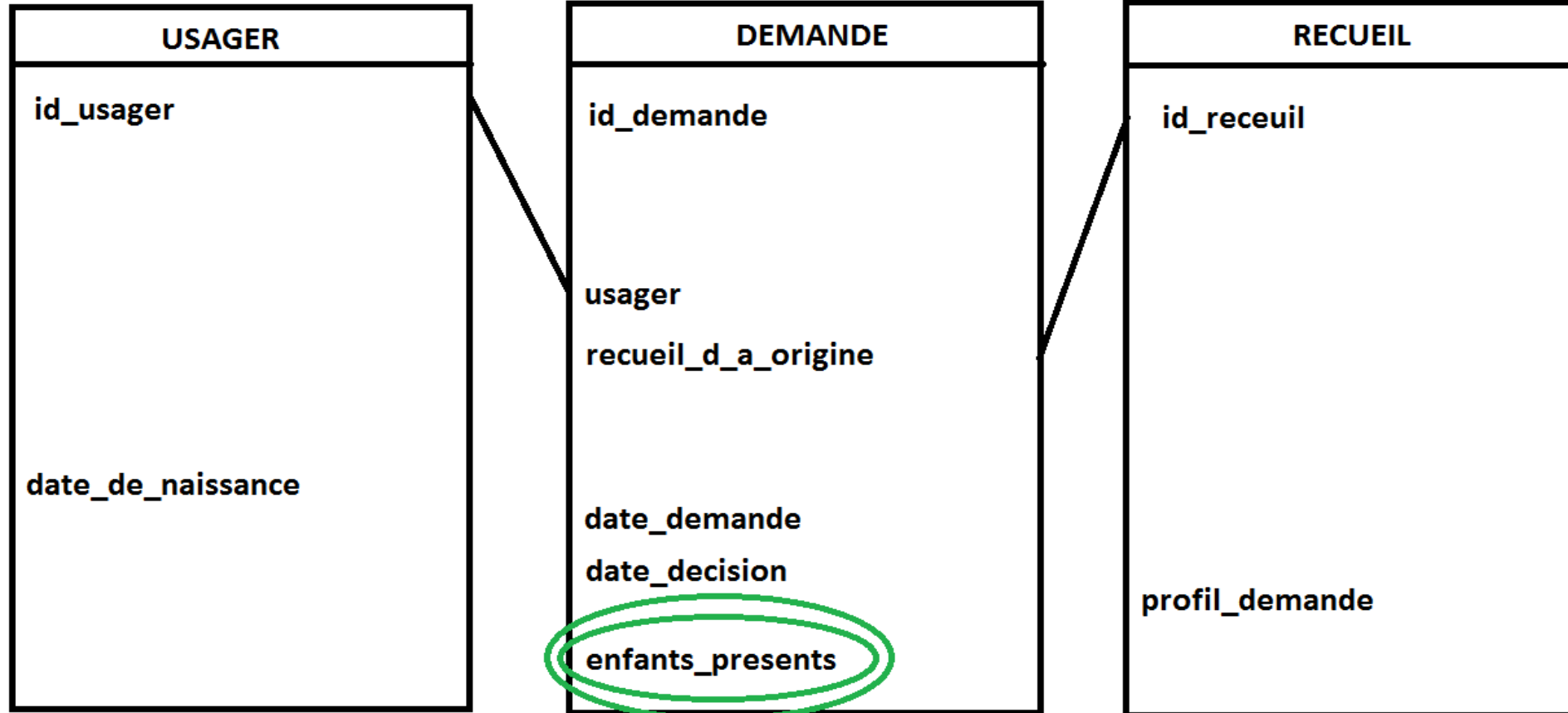
- Case 1: Calculated age under 18
  - If the profile is already accompanied minor or unaccompanied minor, the profile is maintained.
  - If the profile is lonely adult, it is rectified as unaccompanied minor.
  - If the profile is family, it is rectified as accompanied minor.
  - If the profile is empty, it is rectified as accompanied minor.
- Case 2: Calculated age above 18
  - If the profile is already family or lonely adult, the profile is maintained.
  - If the profile is unaccompanied minor, it is rectified as lonely adult.
  - If the profile is accompanied minor, it is rectified as family.
  - If the profile is empty, it is rectified as family.



# Counting accompanying minors

- For each application, variable containing a list
  - Empty list: []
  - Or list of ids : e.g. [41, 42]
- Name of the variable is very interesting:  
enfants\_presents\_au\_moment\_de\_la\_demande
- Findings:
  - Two different application do not share any common id contained in this variable
  - The total number of ids matches (more or less) with the number given by OFPRA (national determining authority)



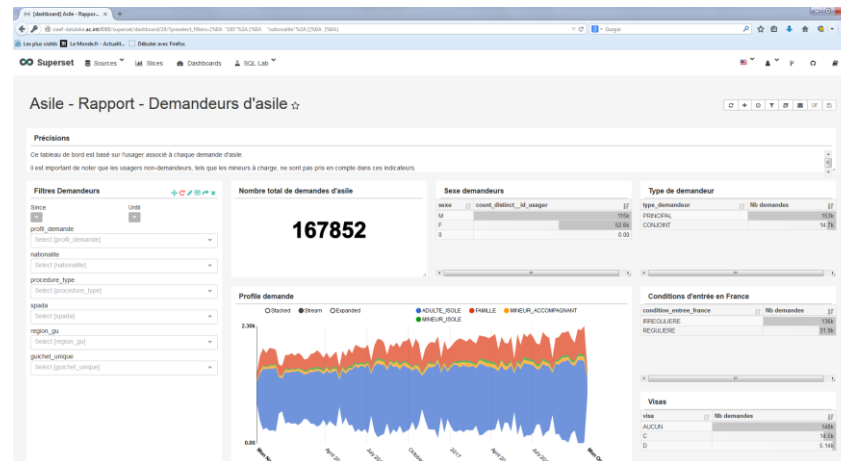


# Cruising speed

- Meetings or informal discussions between:
  - The Direction of Asylum and the private IT consultants
  - The private IT consultants and the statisticians
  - The Direction of Asylum and the Department of Statistics
  - Or the three
- Perpetual succession of trial and error processes
  - E.g. as we saw, the pursuit of a definition for the accompanying minors
  - New questions: e.g. What about Dublin procedure?
- Statisticians provide help to Direction of asylum to ask for new controls of variables in order to improve quality of basic data
- Final aim: improving relevance, accuracy, reliability, quality, etc

# 'Better data. Better lives' : the perspectives

- Improving quality of the datasource
- Datalake and Dashboards (Superset)



- Half-yearly official publication: mid-term?
- Prospective studies: mid-term?

**Thank you for your attention!**

**We welcome any question you may have...**

**Now...**

Or later : [statistiques-dgef@interieur.gouv.fr](mailto:statistiques-dgef@interieur.gouv.fr)