



## ***Building a population and housing census from administrative data.***

## ***The experience of the 2021 Census in Spain***

***Antonio ARGÜESO***  
***Director General for Population Statistics***  
***INE-Spain***



## *Recent evolution of the population and housing Censuses in Spain*

**2001:** Classic (door-to-door) census.

Staff: 45,000 enumerators contracted for 4 months

Overall Budget ~ USD 250 Million

**2011:** Population: A combination between admin data and a survey to 9% of population.

Buildings and dwellings: 80% were visited, 20% using admin data

Staff: 4,500 enumerators hired for 6 months

Budget: ~ USD 85 million

**2021:** Population and housing Census made entirely from admin data

(+ A complementary survey addressed to 1% of population to cover some non-core topics)

Staff: 15 people (Census)

Budget: Census: USD 0 euros (external contracts) + USD 2 Million (survey)

The project has been complemented by a survey to 1% of population

Why?

- Some **mandatory** information of **dwellings** is missing in admin data: water supply system, toilet facilities, heating... (but needed at NUTS-2 o 3 only)
- Some **traditionally collected** information of population and dwellings is also missing: languages spoken, commuting, daily cares, accesibility of buildings
- A survey is useful for improving imputation methods (same-sex couples, flat mates or couples , tenure status...)

➤ The job is almost done

- Current situation: ~ 12 people working on the project.
- Population Census and Compl. Survey published in Dec 2022.
- Housing Census scheduled for last week June 2023



Figure 2 – Census methods used in the EU-27 (2001-2021)



Source: EPRS, based on UNECE and NSI data.

Are you thinking of a register-based Census? 3 things you need:

1

A population register as backbone of the whole operation

Police Records? Tax collection Agency? Health system?

...A good dwelling register is also needed! And more difficult to build than the one on population

2

Good administrative data covering many fields

Population characteristics, education, economic activity,...

3

Legal provisions to get access to them

- Spanish statistical Law (1989) allows access to registers
- European Regulations (223/2009 and Census Regl.)

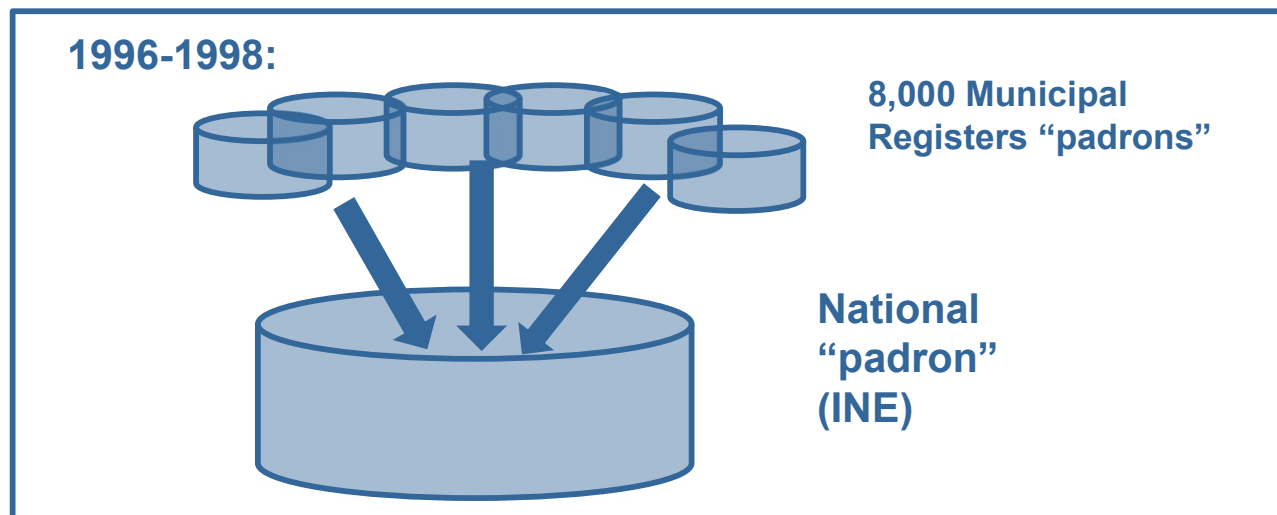
...A population register as backbone of the whole operation (named “*Padrón*”)

In Spain we do have a “different” population register to start with

Why “different”?

1. It is managed by INE (a statistical office), not by the Police, nor Public administr. Ministry, as in other countries...
2. INE has access to direct identifiers (National ID, Passport No,...)

It was **built in 1996** from 8.100 municipal registers (people get registered at municipal offices) but the central database is managed by INE...



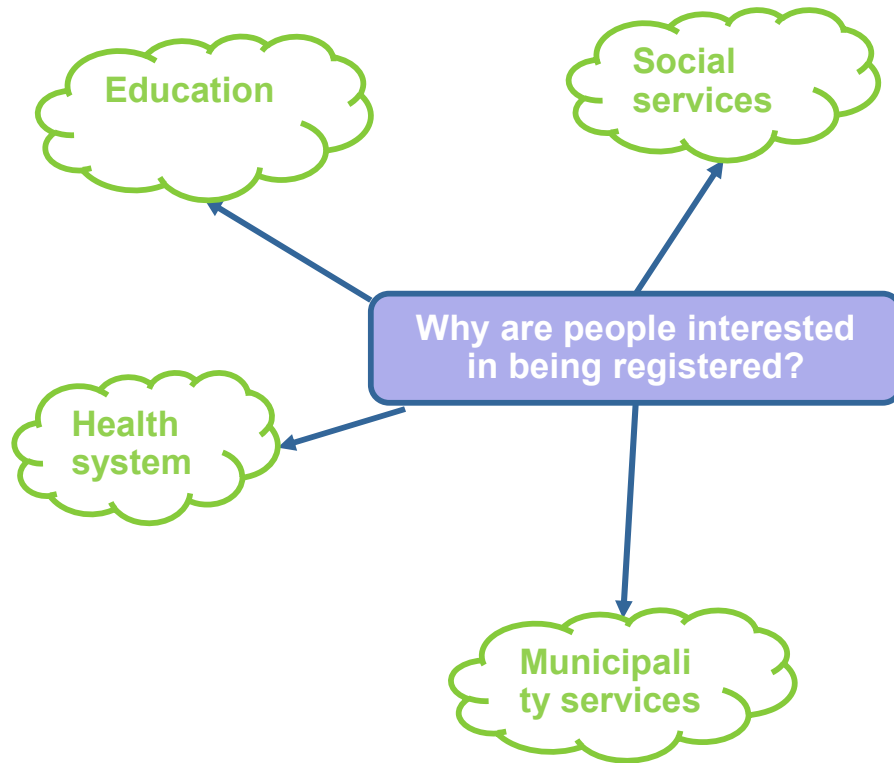
=> From 1996 to 2021: Building a register-based Census took 20+ years

## ...A population register as backbone of the whole operation

Padrón, the population Register. How does it work?



...People go to their municipality to be registered. Why?



- ✓ Because it is mandatory (by National Law)
- ✓ It is a and needed and a proof of residence for many purposes
- ✓ Even for residents without a residence permit

Why are people interested in getting deregistered (when they leave the country)?

NO BENEFITS

- ✗ not many deregistrations made by citizens (20%)
- ✗ Deregistrations made by municipalities (expiration of registrations) or by new occupants



The typical problem of register-based censuses!  
(how to deal with people leaving the country)



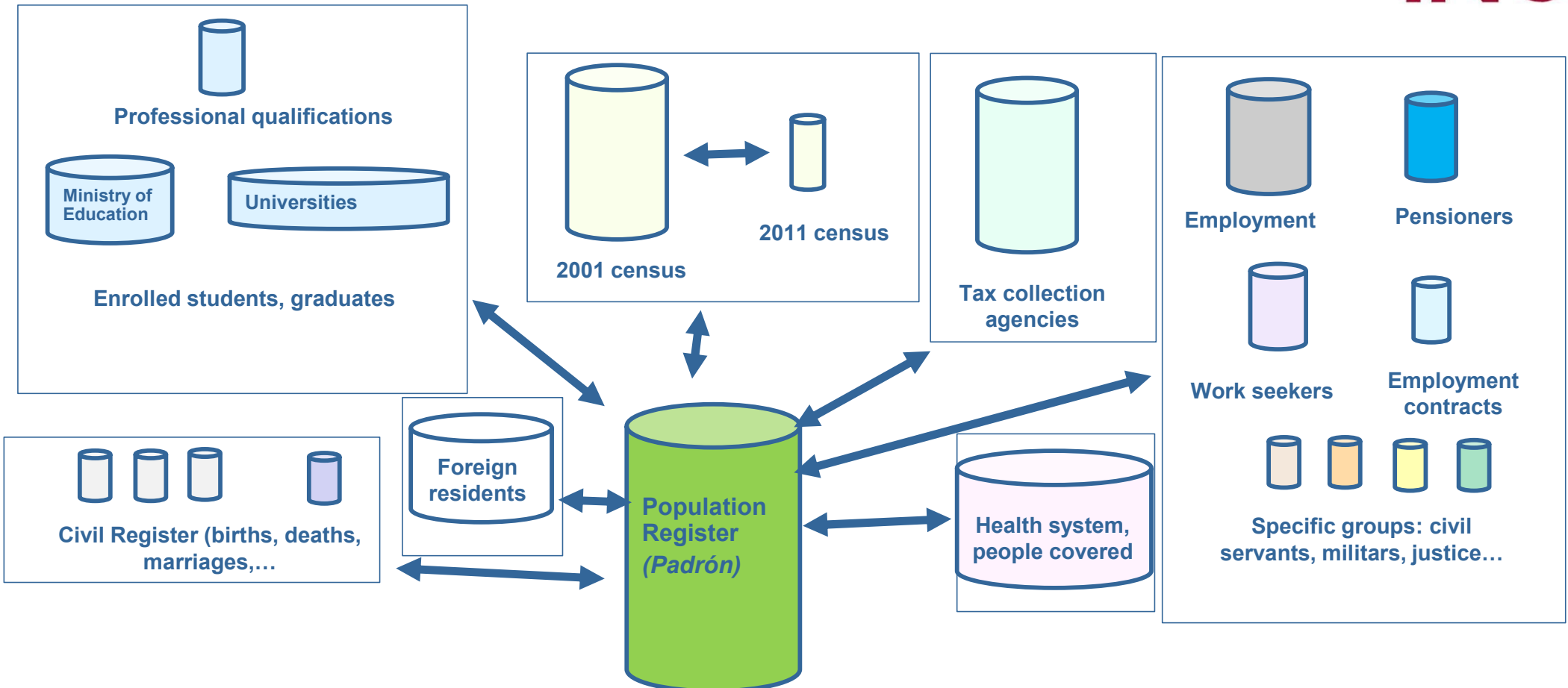
2

Good administrative data covering many fields

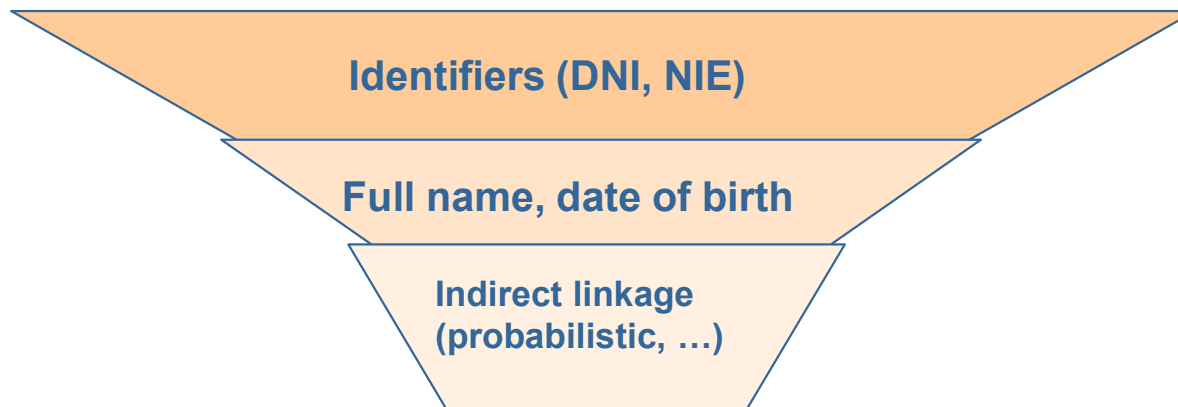
Because the population register contains only a few variables for every person

2. Adding variables to the Population Register (columns)

	identific ation	sex	age	nationality						
1		...	...	...	...	...	...	...	...	...
2		...	...	...	...	...	...	...	...	...
3		...	...	...	...	...	...	...	...	...
4		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
·		...	...	...	...	...	...	...	...	...
47M		...	...	...	...	...	...	...	...	...




**Objective:** linking at individual level population register with all sources



Adding variables: an specific approach for each variable

An example: marital status

Many sources, but none of them provide the whole picture!



		Marital status of the population				
		single	married	widowed	Divorced / separated	Total
winning source	Tax collection agency	15,2	37,1	0,7	4,1	57,1
	Civil register	0,8	6,0	0,0	0,3	7,1
	social Security	0,0	0,0	5,5	0,0	5,5
	Register of Foreign Residents	3,1	1,0	0,1	0,1	4,3
	Imputation	25,9	0,0	0,1	0,0	26,0
	<b>Total</b>	<b>45,0</b>	<b>44,1</b>	<b>6,3</b>	<b>4,5</b>	<b>100</b>

## The Housing Census

- *Padrón* contains addresses for occupied dwellings but we need to have all of them, not just the occupied ones
- Addresses are not perfect. *Padrón* is not geo-referenced



Cadastre is a well managed, updated, complete source used for taxation. It provides a unique ID for all “properties”

- But:
- We need to explore these (very complex) sources to translate from “cadastre” concepts to the demographic ones (from “residential unit” to “dwelling”)
  - We need to link addresses in cadastre with the ones in *Padron* (it was seen as a big challenge back in 2015)



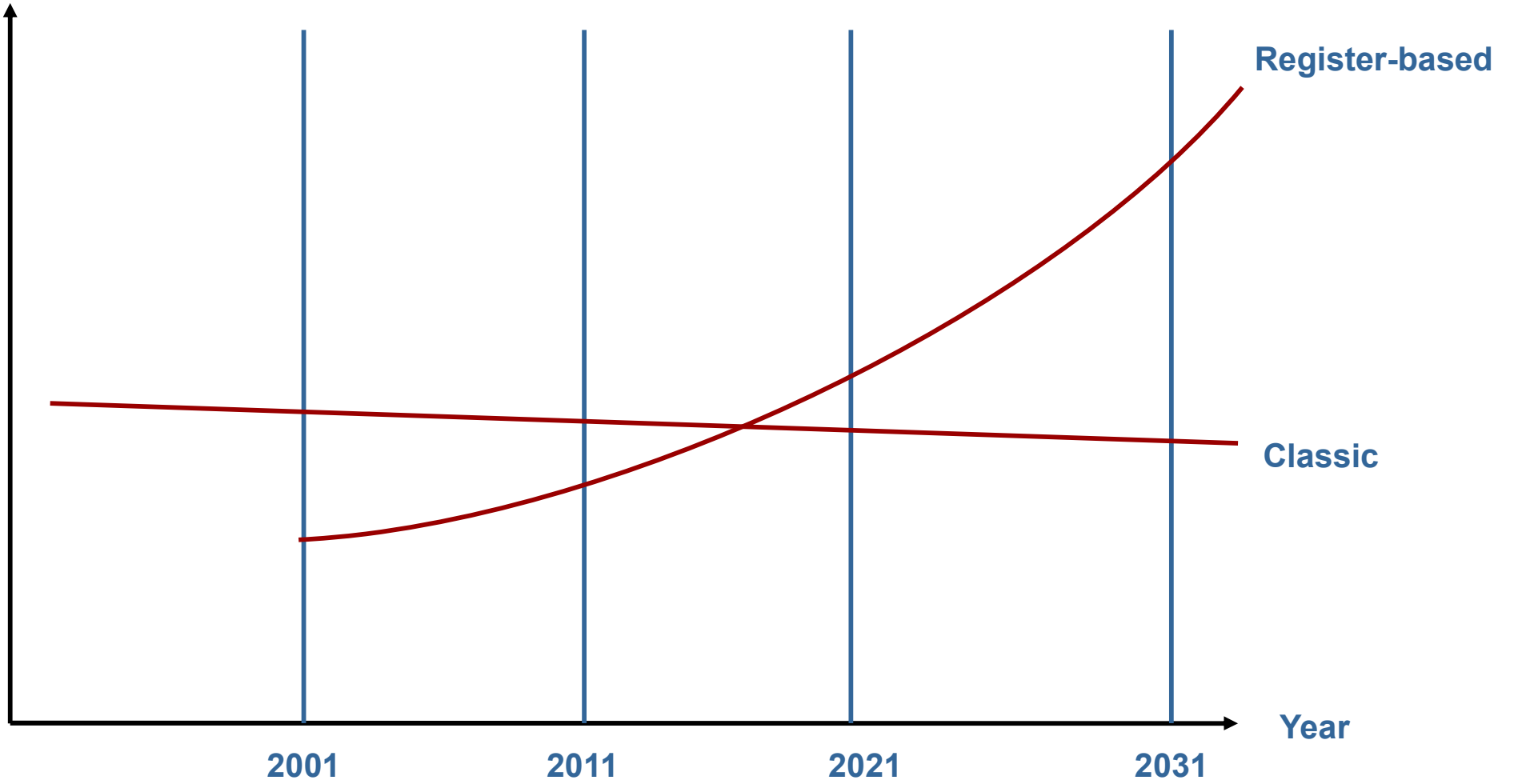
**100% of the 27 million dwellings have “XY coordinates”**  
**98% of *padrón* addresses are perfectly linked**

**Some final remarks:**

- **Administrative data are not perfect, but a classic census is (by far) a lower quality solution**
- **A register-based census is not an overnight decision: it can be done in 2021 because a true population register was implemented in 1996, 25 years earlier!**
- **This approach is the first step into a new (continuous) system of social statistics (next session)**



Quality





INē



IN  
e

Instituto  
Nacional de  
Estadística