

# 7 France

## Twentieth Century Census Microdata

Jean-Louis Rallu  
Laurent Toulemon

### Introduction

French censuses are conducted by the Institut National de la Statistique et des Études Économiques (INSEE), which carries out all stages from collection of data to publication of reports. As a charter member of the IPUMS International project, INSEE is cooperating fully in the integration enterprise. While French census microdata have rarely been available to academics in the past, once the integration of French census microdata is complete, five national censuses—beginning with that of 1962—will be made available to researchers who agree to sign a standard non-disclosure agreement.

Since 1962, French censuses have used a definition of legal population intermediate between ‘*de facto*’ and ‘*de jure*.’ A system of double-counting is used to avoid difficulties linked with the two usual definitions and a check is done to ensure that a single census form exists for each individual. Population at local levels is calculated without double-counting.

---

*Dr. Jean-Louis Rallu is a research fellow in the Population and Development program of the Institute National D'étude Démographiques (INED) in Paris. His research has focused on demographic characteristics of the Pacific Islands; and on migration, work and economic development. Dr. Laurent Toulemon is a research fellow in the Fertility, Family and Sexuality program of the INED. He has published widely on changing patterns of fertility and family structure in France.*

For more information on this project, see <http://www.ined.fr/index.html>.

This system is based on two lists of the household form: list A and list B. List A enumerates all persons living in the household, including those temporarily absent, except those who are in the conditions enumerated at the top of list B:

- military servicemen
- regular soldiers abroad
- children in boarding schools and students living outside the household, except those abroad
- children in care of a nurse outside of the household
- children working and living at their workplace
- ill persons in sanatoriums
- old or disabled persons in a retirement home, a special care-center or a psychiatric hospital
- other persons on long term absence

No individual form is completed for persons in list B. Their individual form is completed at the place where they actually stayed at the census date. Persons enumerated in institutions such as barracks, jails, boarding schools, hostels, sanatoriums, retirement homes and hospitals (resident for 3 months or more)—that is, most of the persons included in list B—are what is called *population comptée à part*. They can be accounted at both places, as well as other double countings enumerated in list B, but totals at districts or higher levels exclude all double-countings.

## Availability of Data

Computer readable microdata are available for censuses from 1962 on. French censuses have been held at irregular intervals: in 1962, 1968, 1975, 1982, 1990 and 1999, in March or April. Data are for mainland France, including Corsica. Overseas departments (DOM) and territories (TOM) are not included. DOMs were not enumerated simultaneously with mainland France before 1982, and TOMs are still enumerated separately.

A machine readable codebook will be available for the first time for the 1999 census. LEDA (INSEE software used to

process censuses) code list and books are available for previous censuses.<sup>1</sup>

## Sampling

Various samples are drawn from census data (see Table 7-1). Most published tables come from a 1/20 sample<sup>2</sup>, but detailed tables, for instance on migration, use a 1/4 sample. The 1/100 samples are drawn from the 1/20 sample dataset.

**Table 7—1. Existing files of 1962 to 1990 censuses.**

<b>File Name</b>	<b>Year</b>	<b>File Characteristics</b>
RP62	1962	1/20
RP68	1968	1/20 1/4 (GTS format currently being set to LEDA format) except Corsica Selected variables of 1/1 except Corsica
RP75	1975	1/20, 1/5 Selected variables of 1/1 except Corsica 1/1000
RP82	1982	1/20 1/4 Selected variables of 1/1 except Corsica 1/100 1/1000
RP90	1990	1/4 Selected variables of 1/1 1/24 1/100

<sup>1</sup> For old censuses, only a copy or two remain and they do not leave the INSEE census unit.

<sup>2</sup> 1/24 for 1990.

## Privacy Law

Legal constraints on the release of census data are linked with the *Loi Informatique et Liberté* and the *Commission nationale de l'informatique et des libertés*. A 1/100 anonymous (without individual identification) sample which would not include information at levels below departments (there are 96 departments in France<sup>3</sup>) would be in agreement with the law.<sup>4</sup>

Except for the 1962 census, due to condition of preservation of the tapes, the overall feasibility of integration is good.

## Data Design

INSEE uses proprietary software, LEDA, for processing census data. The hierarchical structure of the data file is made up of six-levels: GEO, district, building, household (includes information on housing conditions), family, and individual.

The list of variables for the 1990 census, by above levels (including constructed variables) totals 475.

Use of census data in conjunction with GIS has been carried out by INSEE and other institutes such as IGN (*Institut Géographique National*).

## Comparability of Available Information through Time

Comparability of data available for successive censuses varies according to the questions asked as well as the fields of information that were computerized. Most important information (occupation, household composition) is double-

---

<sup>3</sup> Overseas Departments: Martinique and Guadeloupe in the Carribean, French Guyana in South America and Reunion in the Indian Ocean will not be included in microdata. Separate censuses were held in mainland France and overseas departments at different dates until 1975. Overseas Territories still have separate censuses and will also not be included.

<sup>4</sup> For 1999, CNIL would also allow 'commune' to be available at the 1/100 level to enable analysis by communes typology (for instance rural/urban).

coded when major changes occurred between two censuses, so that comparisons remain possible with previous census.

## Geographic Variables

At the department level, administrative boundaries show high consistency. However, former departments 75 and 78 (*Seine* and *Seine et Oise*) were separated into several departments (78, 91, 92, 93, 94, 95) in 1964. Comparability between 1968 and 1962 can be obtained by regrouping data for communes.

Rural-urban typology exists from 1954. The definition of urban area is stable: built up areas with less than 200 meters between houses and consisting of 50 persons or more constitute an "agglomeration of population." Such units with a population of 2,000 or more constitute urban communes. If this population is situated in several communes, these communes are "urban," except those where most of the population is situated outside of the "agglomeration of population." An urban unit can be a town (a single commune agglomeration of population) or a multi-commune agglomeration of population.

Boundaries of metropolitan areas have changed frequently, but it is possible to compare similar areas at successive censuses, using data on communes included in those areas (*Recensement général de la population de 1990 – Composition communale des unités urbaines. Population et délimitation, 1990, and – Evolutions démographiques 1975-1982-1990, INSEE*).

Various administrative units have been created during the 1962-1990 period: including ZPIU (industrial or urban population area), and ZE (employment area).

## Housing Variables

The figure for principal residence *résidence principale* is equal to the figure of households or heads of household.<sup>5</sup> Housing type and construction type were a single question up to 1982. Some categories could therefore be difficult to follow

---

<sup>5</sup> In 1962, mobile homes (boat, barges or caravans) were included in ordinary households (*menages ordinaires*), but they were separated from 1968.

between 1982 and 1990, although the most important ones are consistent.

Questions on water and energy have been replaced by questions on housing condition. The electricity question was dropped from 1962, the gas question from 1982 and the water question from 1990. However, when available, this information is highly consistent. Questions on household facilities such as bath-shower, and to a lesser extent heating and toilet facilities, vary greatly.

## Individual Variables

With respect to individual variables, age can always be recalculated from the full date of birth (y, m, d). Marital status refers to legal marital status (4 categories: single, married (or re-married), widowed or divorced) for the whole period.

Changes in family definition have occurred: until 1982, children (un-married or without spouse) below age 25 were included in families; from 1990, this age-limit was removed. "Head of household" was changed to a reference person (*personne de référence*) in 1982, and a new procedure was used to choose the reference person.<sup>6</sup> Note that up to 1975, grandchildren, sons and daughters in law, step children and adopted children were included among 'children' of the head of the household.

### *Education*

Age-left-school is available from 1968. Education categories have changed due to reforms in curriculum and new diplomas. Major changes are between 1968 and 1975, and 1982 and 1990. General education and technical-vocational education were separate questions until 1990. The former is more consistent than the latter until 1990. Consistency becomes rather low between 1982 and 1990 when general and vocational diplomas are included in one seven category question, whereas altogether 12 categories existed in 1982. Major groups, on the other hand, are consistent. Note also that no box was provided

---

<sup>6</sup> A similar procedure can be applied to 1975 data (see coding schemes below).

for “no diploma” until 1999, so that it is not clear if persons coded “without diploma” are really such or just non-responses.

### *Economic*

Economic variables are consistent as regards labour force participation and occupational status. Note that unemployment in French censuses is not comparable to the ILO definition.<sup>7</sup> Occupational categories changed in 1982; similar categories with those of 1975 are also available for 1982. New aspects have been considered from 1990: full/part time, new types of trainees, etc.

### *Fertility*

French censuses do not include information on fertility by means of a question on ever-born children.<sup>8</sup> Census data on families do not include children who left, married children, whenever they would still be household members, and children above a certain age, varying with census (see Coding schemes section below).

### *Migration*

Two questions provide information on migration: birth place (in 1/100 data files birth place would be at department level only) and previous residence. Previous residence is always as of January 1 of the year preceding the census. As censuses are held at irregular intervals, the duration varies and comparability is low at both national and international levels.

## Coding Schemes

This section reviews the coding schemes of two variables which differ significantly across countries.

### *Head of Household*

Household-family analysis is carried out in two steps. First, families are constituted. A family consists of at least two

---

<sup>7</sup> Yearly labour force surveys provide unemployment rates in both census and ILO definitions.

<sup>8</sup> A family survey is carried out on a large sample (200 000 or more) at every census since 1962, except 1968.

persons including:

- a married or un-married couple and eventually their children (or children of one of the spouses)
- a lone person and his/her children (one-parent family)

Adults have to be 15 years of age or above and children have to be single with no un-married spouse present. Until the 1982 census, children had to be below 25. The data file of the 1990 census includes a variable similar to families in the 1982 census. Second, the reference person (formerly, head of household) is chosen according to a procedure established in 1982. Before 1982, the household-head was the person designated as such (self designation or in agreement with members of the household) and listed on the first line of the household members list. The criteria used are family, fathers, labour force participation and age and are applied as follows:

- If the household includes at least one couple-family, the reference person is the oldest father among the couple-families participating in the labour force; if none is participating in the labour force, it is the oldest father
- If the household includes at least one single-parent family, the reference person is the oldest adult in one-parent families, participating in the labour force; if none is in the labour force, it is the oldest adult
- If the household does not include any family, the reference person is the oldest adult participating in the labour force, except lodgers, and salaried men or women; if none is participating in the labour force, it is the oldest
- Links between reference person and other household members (LPRM) are recalculated from links to the person at the top of the household members list (L1) and the link between the latter and the reference person. Another variable (L2) shows adults and children of the two first families in the household.<sup>9</sup>

---

<sup>9</sup> Source: INSEE, *Recensement général de la population de 1990, Guide d'utilisation*, Tome 5.

The link between the reference person of the household and the reference person of the family is also used to characterize the family structure of the household. From 1982 on, lodgers and salaried men or women are excluded from consideration.

## Occupation

Occupation is recorded as an open-ended question (blank box). PROF (occupation) is coded from the word list of occupations (PCS – see annex 2) for employed labour force. PROFX is a coded value. PROF is a recoded value. A hotdeck is used to infer occupation of employed labour force when it is not reported. Former occupation (PROFA) is known for unemployed persons and persons who worked previously.

- Socio-economic category (CS), is the first two digits of PCS; it consists of 86 categories
- CS24 and CS8 are regrouped socio-economic categories. CS8 is the first digit of CS
- Employment status (STAT – 4 categories: wage earner, own account, employer, family helper) and (STATD – 8 categories, with 5 types of wage earners: private and various public) is also available
- POSP (12 categories) is occupational status. SITP (7 categories) and EMPL (10 categories) provide information on situation in employment: various trainees, short/long-term contracts, etc

## Conclusion

French categories are different from those used in English-speaking countries and also from those used by Southern European countries. However, there are extensive possibilities for comparisons of French census data with other countries. Educational levels and diplomas are rather similar, particularly for university education. Household and family structures are difficult points, but comparative studies have been carried out within the European Community which can provide frameworks for extended comparisons. The same applies to occupational and socio-economic categories.

## Appendix 1. Niveau menage (*Household Level*)

### TYPMD Type de menage detaille (*Detailed Household Type*)

---

#### Homme vivant seul (Lone male)

- 011 Activ (Labour force)
- 012 Inactive (Non-Labour force)

#### Femme vivant seule (Lone female)

- 021 Activ (Labour force)
- 022 Inactive (Non-Labour force)

#### 030 Autre Menage sans famille (Other Household without family)

#### Famille principale monoparentale (Main family is one-parent family)

##### Home actif (Labour force male)

- 101 Avec 1 enfant (with 1 child)
- 102 Avec 2 enfants (with 2 children)
- 103 Avec 3 enfanats ou plus (with 3 or more children)

##### Homme inactif (Non-labour force male)

- 111 Avec 1 enfant (with 1 child)
- 112 Avec 2 enfants (with 2 children)
- 113 Avec 3 enfanats ou plus (with 3 or more children)

##### Femme active (Labour force female)

- 201 Avec 1 enfant (with 1 child)
- 202 Avec 2 enfants (with 2 children)
- 203 Avec 3 enfanats ou plus (with 3 or more children)

##### Femme inactif (Non-labour force female)

- 211 Avec 1 enfant (with 1 child)
- 212 Avec 2 enfants (with 2 children)
- 213 Avec 3 enfanats ou plus (with 3 or more children)

## Appendix 1. Niveau menage (*Household Level*)

### TYPMD Type de menage detaille (*Detailed Household Type*) (*continued*)

---

Famille principale couple (*Main family is couple family*) - **continued**

Homme et femme actifs (*Labour force male and female*)

Homme de moins de 40 ans (*Male below 40*)

300	Sans enfant ( <i>no children</i> )
301	1 enfant ( <i>1 child</i> )
302	2 enfants ( <i>2 children</i> )
303	3+ enfants ( <i>3+ children</i> )

Homme de plus de 40 ans (*Male above 40*)

310	Sans enfant ( <i>no children</i> )
311	1 enfant ( <i>1 child</i> )
312	2 enfants ( <i>2 children</i> )
313	3+ enfants ( <i>3+ children</i> )

Homme actif et femme inactive

(*Labour force male and non-labour force female*)

Homme de moins de 40 ans (*Male below 40*)

400	Sans enfant ( <i>no children</i> )
401	1 enfant ( <i>1 child</i> )
402	2 enfants ( <i>2 children</i> )
403	3+ enfants ( <i>3+ children</i> )

Homme de plus de 40 ans (*Male above 40*)

410	Sans enfant ( <i>no children</i> )
411	1 enfant ( <i>1 child</i> )
412	2 enfants ( <i>2 children</i> )
413	3+ enfants ( <i>3+ children</i> )

Homme inactif et femme active

(*Non labour force male and labour force female*)

500	Sans enfant ( <i>no children</i> )
501	1 enfant ( <i>1 child</i> )
502	2 enfants ( <i>2 children</i> )
503	3+ enfants ( <i>3+ children</i> )

Home et femme inactifs (*Non labour force male and female*)

600	Sans enfant ( <i>no children</i> )
601	1 enfant ( <i>1 child</i> )
602	2 enfants ( <i>2 children</i> )
603	3+ enfants ( <i>3+ children</i> )

---

## Appendix 2. Niveau famille – Family Level

**LPRMF**    **Lien de la personne de référence de la famille à la personne de référence du ménage** (*Link between the reference person of the family and the reference person of the household*)

Modalités <i>Values</i>	Libellés <i>Title</i>
1	Personne de référence du ménage ( <i>Reference person of the household</i> )
3	Enfant de la personne de référence du ménage ou de son conjoint ( <i>Child of the reference person of the household or of his/her spouse</i> )
4	Petit-enfant de la personne de référence du ménage ou de son conjoint ( <i>Grand-child of the reference person of the household or of his/her spouse</i> )
5	Ascendant de la personne de référence du ménage ou de son conjoint ( <i>Father, mother, grand-parent of the reference person of the household or of his/her spouse</i> )
6	Autre parent de la personne de référence du ménage ou de son conjoint ( <i>Other relative of the reference person of the household or of his/her spouse</i> )
7	Ami de la personne de référence du ménage ou de son conjoint ( <i>Friend of the reference person of the household or of his/her spouse</i> )

### Appendix 3. Niveau individu (*Individual Level*)

#### LPRM      Lien à la personne de référence du ménage (Link to the reference person of the household)

<b>Modalités</b> <i>Categories</i>	<b>Libellés</b> <i>Title</i>
PRM	Personne de référence du ménage ( <i>Reference person of the household</i> )
CONJ	Conjoint de la personne de référence du ménage ( <i>Spouse of the reference person of the household</i> )
ENF	Enfant de la personne de référence du ménage ou de son conjoint ( <i>Child of the reference person of the household or of his/her spouse</i> )
PENF	Petit-enfant de la personne de référence du ménage ou de son conjoint ( <i>Grand-child of the reference person of the household or of his/her spouse</i> )
ASC	Ascendant de la personne de référence du ménage ou de son conjoint ( <i>Father, mother, grand-parent of the reference person of the household or of his/her spouse</i> )
AUTP	Autre parent de la personne de référence du ménage ou de son conjoint ( <i>Other relative of the reference person of the household or of his/her spouse</i> )
AMI	Ami de la personne de référence du ménage ou de son conjoint ( <i>Friend of the reference person of the household or of his/her spouse</i> )
PENS	Pensionnaire, sous-locataire, enfant en nourrice sans lien de parenté avec la personne de référence ou son conjoint ( <i>Lodger, roomer, nursing child without kinship link to the reference person nor to his/her spouse</i> )
DOME	Domestique ou salarié logé ( <i>Maid or lodged salaried-man</i> )

## Appendix 4. Niveau individu (*Individual Level*)

### CS8 Socio-Economic Status (8 Categories)

<b>Modalités</b> <i>Categories</i>	<b>Libellés</b> <b>Title</b>
1	Agriculteurs exploitants ( <i>Farmers</i> )
2	Artisans, commerçants, chefs d'entreprise ( <i>Craftsmen, tradesmen, company managers</i> )
3	<b>Cadres, professions intellectuelles supérieures ...</b> ( <i>Executive managers, professionals, scientists....</i> )
4	Professions intermédiaires ( <i>Middle level occupations</i> )
5	Employés ( <i>Employees</i> )
6	<b>Ouvriers</b> ( <i>Production workers, incl. agricultural workers</i> )
7	<b>Retraités</b> ( <i>Pensioneers</i> )
8	Autres personnes sans activité professionnelle ( <i>Other non labour force</i> )

## Bibliography

INSEE, (various years), *Recensement Général de la population*, various volumes: *Population, Ménages, Familles, Activité, Unités géographiques: départements, communes, ...*INSEE, Paris.