



Microdata from the Survey on Population in Relation to Activity List  
of file variables

## CONTENTS

<b>1. Introduction .....</b>	<b>1</b>
<b>2. Criteria for selection of variables.....</b>	<b>2</b>
2.1 Criteria of confidentiality .....	2
2.2 Confidentiality criteria.....	2
2.3 Quality criteria .....	3
<b>3. Use of microdata file.....</b>	<b>4</b>
<b>4. References .....</b>	<b>4</b>
<b>APPENDIX 1. List of variables of the microdata file .....</b>	<b>5</b>

## **1. Introduction**

The Survey on Population in Relation to Activity (PRA) statistical operation provides constant statistical information about the volume and characteristics of the main groups, which can in turn be used to classify the population of the Basque Country according to participation in distinct economic activities, as well as their situation changes.

The files of the Survey on Population in Relation to Activity constitute a dissemination product aimed at users with experience in the analysis and treatment of microdata, or files containing information at registry or unit level. This format provides added value to users, permitting them to carry out data exploitation and analyses that, owing to obvious limitations, cannot be covered by circulation in the form of tables, publications and reports.

A microdata file is provided, whose units are people, and a number of the household they belong to is facilitated (randomly generated), along with the relationship between the members of the household.

## **2. Criteria for selection of variables**

This file contains a selection of variables collected in the survey for the selected registry. The selection of variables was based on criteria for both confidentiality and quality, which are detailed below:

### **2.1 Criteria of confidentiality**

This content refers to data that are especially protected by the personal data protection act and that require express consent from the surveyed individual for them to be handed over. In this survey it is considered that there are no variables of this type.

### **2.2 Confidentiality criteria**

Maintaining the statistical secrecy of informants is one of the main premises that govern statistical activity, as reflected in a number of documents such as the European Statistics Code of Practice, in its PRINCIPLE 5: Statistical confidentiality and data protection, according to which "The privacy of data providers, the confidentiality of the information they provide, its use only for statistical purposes and the security of the data are absolutely guaranteed".

Therefore, the dissemination of any individual information collected under statistical secrecy must be carried out anonymously and in such a way that it cannot be associated with any specific individual, not even via the content or disaggregation of the disseminated variables.

The fact that it involves a sample of the population increases the level of protection of the file, but even so there must be assurance that there is no revelation of the statistical secret.

Therefore, and in order to guarantee this safeguard, the following restrictions have been applied to the PRA individuals file:

- Under no circumstances are direct identification variables or data of a personal nature (national identity document, name, surname, addresses, telephone numbers, etc.) to be included.
- The maximum geographical disaggregation included is desegregation by province and also by capitals/rest.
- The age of the person surveyed has been aggregated into five-year groups.
- The identification number of each household on the file is not maintained from one quarter to another, in a way that it is not possible to identify the same household on two files for different quarters.

The data protection methods used are basically aggregation and recodification of data. Neither elimination nor disturbance of values has been used.

### **2.3 Quality criteria**

The main limitation in any survey by sample is due to the availability of information solely for the sample units and not for the total objective population. The sampling design of the survey will shed light on the level of representativeness that can be achieved with the information collected.

In the case of the PRA, it involves a sample of dwellings stratified into 20 regions, plus the 3 provincial capitals. It is, in addition, a balanced sample (CUBE method) for the following variables: distribution of population by age and sex in the regions and capitals, family sizes, presence on administrative employment and unemployment registers and, lastly, nationality (in two modalities, Spanish and foreign).

For the main variables and crosses of the survey the level of representation required is provinces and capitals. However, the large number of variables and the high level of detail that can be achieved with the variables included in this file do not guarantee the statistical significance of all the operations or analyses that can be performed on this data.

The conclusions derived from studies or analyses performed on this data are the responsibility of the end user. However, to give an idea of the quality of the information for a specific cross or disaggregation, we recommend going to the sampling error tables for the survey:

[http://es.eustat.eus/estadisticas/tema\\_57/opt\\_1/tipo\\_1/ti\\_Encuesta\\_de\\_poblacion\\_en\\_relacion\\_con\\_la\\_actividad/temas.html#el](http://es.eustat.eus/estadisticas/tema_57/opt_1/tipo_1/ti_Encuesta_de_poblacion_en_relacion_con_la_actividad/temas.html#el)

and the corresponding sampling errors calculation report ([http://www.eustat.es/document/datos/Calculo\\_errores\\_PRA\\_c.pdf](http://www.eustat.es/document/datos/Calculo_errores_PRA_c.pdf))

where recommendations on the acceptable levels of sampling error appear.

### **3. Use of microdata file**

The PRA individuals microdata file is in CSV format and is structured around the thematic areas collected by the survey for this statistical unit.

The data file is accompanied by its registry design (in Excel format) where the description of the file variables is included. In summary, the list of variables is provided in Annex I. One of the variables is the individual elevator of each person to be taken into account upon carrying out any tabulation of the data, in order to go from sampling to population units. The elevator is the weight of each unit in the sample, that is, the number of people it is representing in the population according to the design of the sample used.

### **4. References**

Survey on Population in Relation to Activity (PRA).

[http://www.eustat.eus/estadisticas/tema\\_37/opt\\_0/ti\\_Poblacion\\_en\\_relacion\\_con\\_la\\_actividad\\_PRA/temas.html](http://www.eustat.eus/estadisticas/tema_37/opt_0/ti_Poblacion_en_relacion_con_la_actividad_PRA/temas.html)

## **APPENDIX 1. List of variables of the microdata file**

<b>Num</b>	<b>Variable</b>	<b>Description</b>
1	NUMH	Household number
2	AENC	Year surveyed
3	TENC	Reference quarter
4	TERH	Province
5	MUNI	Capital
6	SEXO	Gender
7	LNAC	Place of birth
8	AGE	Age
9	NACI	Nationality
10	LEST	Level of education completed
11	ENRE	Formal education system
12	FOCU	Professional training
13	SJUB	Retirement status
14	SILH	Household duties situation
15	EMPTP	Part-time employment
16	JRED	Reason for reduced working hours
17	BUSQ	Job search
18	RBUSQ	Reasons for seeking employment
19	JORB	Working hours sought
20	GBUSQ	Carry out employment seeking activities
21	FBUSQ	Main employment seeking method
22	MSBUSQ1	Months seeking employment
23	DISP	Availability
24	PRA1	Relation to activity (ILO)
25	PRA2	Relation to activity
26	PROF	Main occupation
27	RACT	Main activity
28	SPRO	Main professional situation
29	SECT	Main institutional sector
30	CONTR	Type of contract
31	HTRA	Hours worked
32	PARE	Relationship
33	ELEV	Elevator