



2006 Demographic Survey Microdata

Description of the File

CONTENTS

1. Introduction	2
2. Criteria for selection of variables.....	2
2.1 Confidentiality criteria	2
2.2 Quality criteria	3
2.2.1 Relevance	3
2.2.2 Compartability.....	3
2.2.3 Accuracy and reliability	3
3. Record design.....	4
4. Elevators	4
5. Description of variables	5
ANNEX I	6

1. Introduction

The microdata file from the 2006 Demographic Survey (hereinafter ED 2006) is a dissemination product aimed at users experienced in analysing and processing microdata. This format provides added value to the user, allowing him or her to use and analyse data which, due to obvious limitations, the current standard dissemination in the form of tables, publications and reports cannot encompass.

Although the scope and purpose of the ED 2006 goes far beyond what is included in this file (analysis of make up of households, fertility, ties, etc.), a product is produced with great potential for use that includes the most important socio-demographic variables of the families and individuals in the Basque Country and covers areas of great interest, such as: characteristics and setting of the households, job market, educational level, use of the Basque language, migration, etc.¹

2. Criteria for selection of variables

Specifically, and due to the complex structure and dimension of this survey, a sub-set of variables already prepared for treatment relating to the housing and individuals included in the survey has been selected. This selection has been carried out based on criteria of both confidentiality and quality of information which is detailed below.

2.1 Confidentiality criteria

Maintaining statistical secrecy for those supplying information is one of the main guiding principles in statistical activity. Therefore, the dissemination of any individual information collected under the obligation of statistical secrecy should be anonymous and carried out in such a way that none of the identities of the statistical units can be deduced, not even via the content or disaggregation of the published variables.

Therefore, to safeguard this guarantee, the following restrictions have been applied to the microdata file to be disseminated:

- In no event does it include direct identification variables or personal details (ID No., name, surname, address, telephone No., etc.).
- The maximum geographical disaggregation included is the ZONE (grouping of districts) and provincial capitals.
- The age of the person interviewed is presented in five-yearly bands given its highly identifiable nature in combination with the other demographic variables. Other references to age or time periods (e.g. year of moving to the municipality, year house was built, etc) are also presented in five- or ten-yearly bands.
- Standard groupings of variables which characterise the family and the housing are used to avoid the identification of "unusual" units (e.g. very large families, large houses, etc.)
- Housing identifiers that can isolate individuals belonging to a single household are not included. The basic statistical unit included in the file is the individual and a variable is provided that enables the record for the person of reference to be filtered, so that the file can be used at the level of housing.

¹ For further information, consult the survey questionnaire and documents at:
www.eustat.es/estadisticas/idioma_c/tema_165/opt_0/tipo_3/ti_Familias/temas.html

- No sensitive subject area has been included. Said subject areas include all those protected by the Protection of Personal Data Act 15/1999, of 13 December, such as health, race, religion, ideology, finances, etc.

2.2 Quality criteria

Another aspect to be taken into account in the analysis and processing of statistical products is data quality. The quality of the statistical product can be defined using the specific factors enacted by EUROSTAT: relevance, accuracy and reliability, timeliness and punctuality, coherence and comparability, accessibility and clarity.

In this case and to select the variables to be included in the ED 2006 microdata file, we have focused on aspects such as relevance (relevance or interest of the data for the final user) and the accuracy and reliability of the information.

2.2.1 Relevance

Given the great quantity of variables collected by the survey, we have attempted to synthesise and include that are supposed to be of the greatest interest and relevance for the general survey users. Thus, the following considerations have been made:

- For the first time in this survey, variables relating to the amenities of the housing are included, along with the social and physical environment that are of growing interest for the user in recent years.
- The most important variables in the areas of job market, education level and use of the Basque language are included, along with the main migratory variables. These are of great interest for the generic user.

2.2.2 Comparability

- Self-classification variables² have not been included to avoid problems of objectivity and comparability in the estimates arising from those variables (e.g. unemployment, inactivity data, etc.).

2.2.3 Accuracy and reliability

In any sampling survey, the main limitation consists of the fact that information is available solely for the sample units and not for the whole of the target population. The sampling design will shed light on the level of representativity that we can attain with information collected.

In the case of the ED 2006, the size of the sample (25,000 dwellings and 66,822 individuals) ensures a more than acceptable representativity at greater geographical disaggregation than in the other sampling surveys prepared by EUSTAT: Even though the sampling errors are not systematically calculated for this sample, a table of theoretical maximum errors calculated based on the sample design of the survey is below. Specifically, this design stratifies the sample by province and size of the municipality, by systematically selecting the units within each stratum:

² A self-classification variable is one when the person being surveyed defines himself or herself in a specific category according to his/her subjective perception.

Maximum Variation Coefficients (VC) per stratum to provide dwellings in the stratified sample. ED 2006.

	Occupied Housing Sample	Sample Framework Population	VC (%)
Total Basque Country	25,000	2,082,587	0.65%
1, + 100,000 inhabitants	9,862	745,201	1.04%
2, 40,000-100,000	3,841	376,688	1.61%
3, 10,000-40,000	6,485	560,466	1.25%
4, Up to 10,000 inhabitants	4,812	400,232	1.47%
Total Álava	5,539	286,387	1.33%
1, + 100,000 inhabitants	4,194	216,852	1.53%
2, 40,000-100,000	-	-	-
3, 10,000-40,000	366	18,931	5.18%
4, Up to 10,000 inhabitants	979	50,604	3.17%
Total Bizkaia	10,966	1,122,637	0.95%
1, + 100,000 inhabitants	3,418	349,972	1.70%
2, 40,000-100,000	3,127	320,087	1.78%
3, 10,000-40,000	2,447	250,579	2.01%
4, Up to 10,000 inhabitants	1,974	201,999	2.24%
Total Gipuzkoa	8,495	673,563	1.08%
1, + 100,000 inhabitants	2,250	178,377	2.10%
2, 40,000-100,000	714	56,601	3.72%
3, 10,000-40,000	3,672	290,956	1.64%
4, Up to 10,000 inhabitants	1,859	147,629	2.31%

If we take into account that an acceptable sample error level allows variation coefficient values between 10% and 15%, we can state that the accuracy achieved by this survey is excellent with respect to the sampling strata. However, it should be pointed out that the detail that can be reached with the possible crosses of the variables included in this file does not guarantee the statistical significance of all the uses or analyses that can be carried out with this data. Therefore, the conclusions reached from the studies or analyses carried out with this data are the responsibility of the final user.

3. Record design

The EDV 2006 microdata file is in text format, with fixed columns and structured around the subject areas covered by the survey for housing, families and individuals. The description of the variables, the rate, the possible states and the exact positions in the file are described in the "Diseño_registro_ED_2006.xls" Excel file, which is attached together with the microdata file.

4. Elevators

In the EDV 2006, the post-stratification and adjustment methods used mean that the calculated elevator is unique for housing and individuals. In other words, the elevator provided is adjusted for all the occupied housing in the Basque Country and for all the individuals in the population. Given that

the file contains as many records as surveyed individuals³, a single record per dwelling will be needed to obtain the adjustment by dwellings and be able to use the file at that file. Therefore, a variable is provided to perform this filter and obtain the housing sample (PARE=1).

5. Description of variables

The literal descriptions of the coding and classifications used for each of the variables included in the file are included. See Annex I.

³ The survey gathers information about all the individuals that reside in the dwellings in the sample at the time of carrying out the survey.

ANNEX I

DESCRIPTION OF THE MICRODATOS_ED_2006.txt FILE**TERH. Province of residency**

- 01 Araba
- 20 Gipuzkoa
- 48 Bizkaia

ZONR. Residency Zone

- 1 Llanada alavesa
- 2 Rest of Alava
- 3 48011 Bilbao
- 4 Duranguesado
- 5 Gernika-Bermeo+Plentzia-Mungia+Markina-Ondarroa
- 6 Arratia-Nerv+Encartaciones
- 7 Donostialdea+Bajo Bidaso
- 8 Alto Deba+Bajo Deba+Urola Costa
- 9 Tolosaldea+Goierri

MUNR. Municipality of residency (only provincial capitals)

- 1 Bilbao
- 2 Vitoria-Gasteiz
- 3 Donostia-San Sebastián
- 4 Other municipalities

ACAL. Has hot water

- 1 Yes
- 2 No

ACON. Year dwelling built

- 1 ≤ 1900
- 2 1901 - 1940
- 3 1941 - 1950
- 4 1951 - 1960
- 5 1961 - 1970
- 6 1971 - 1980
- 7 1981 - 1990
- 8 1991 - 2000
- 9 ≥ 2001

CALE. Has heating

- 1 Central
- 2 Individual
- 3 Separate appliances
- 4 None

GAST. Has gas distributed by piping

- 1 Yes
- 2 No

NHAB. Number of rooms

- 1 1 - 3 rooms
- 2 4 rooms
- 3 5 rooms
- 4 6 rooms
- 5 7 or more rooms

REGT. Ownership system

- 1 Owned
- 2 Rented
- 3 Other system

SUPF. Usable surface area of the building

- 1 ≤ 30 m²
- 2 31 – 60 m²
- 3 61 – 90 m²
- 4 91 – 120 m²
- 5 121 – 150 m²
- 6 151 – 180 m²
- 7 ≥ 181 m²

TELF. Has a landline

- 1 Yes
- 2 No

TFAM. Size of the family

- 1 1 person
- 2 2 people
- 3 3 to 5 people
- 4 6 to 9 people
- 5 10 or more people

TIPF. Family type

- 1 Single member
- 2 Composite
- 3 Nuclear without children
- 4 Nuclear with children
- 5 Single parent (father or mother)
- 6 Extended
- 7 Pluri-nuclear

NUBD. Number of bathrooms

- 1 None
- 2 1 bathroom
- 3 2 bathrooms
- 4 3 or more bathrooms

TELFM. Has a mobile telephone

- 1 Yes
- 2 No

REFR. It has a fridge

- 1 Yes
- 2 No

PARK. Car park

- 1 Underground car park
- 2 Surface car park
- 3 No car park.

RUIDX. Noise in the dwelling

- 1 Yes
- 2 No

CONTA. Pollution or bad smells in the dwelling

- 1 Yes
- 2 No

LIMP. Streets not very clean

- 1 Yes
- 2 No

MCOM. Poor accessibility

- 1 Yes
- 2 No

PVERD. Few green areas

- 1 Yes
- 2 No

DELIN. Delinquency around the dwelling

- 1 Yes
- 2 No

ECIV. Legal civil status

- 1 Single
- 2 Married
- 3 Widowed
- 4 Divorced, separated

FAMI. Type of family coexistence

- 1 With parents
- 2 Own nuclear family
- 3 Single-member family
- 4 Composite member family
- 5 Other situation

PARE. Situation in the family

- 1 Householder
- 2 Other relationship

SEXO. Sex

- 1 Male
- 2 Female

SFAM. Family Status

- 1 Main person
- 2 Group head
- 3 Members of the family group
- 4 Isolated individuals

TANU. Size of the nucleus

- 1 1 person
- 2 2 people
- 3 3 people
- 4 4 to 5 people
- 5 6 or more people

TNUC. Type of nucleus

- 1 Isolated individuals
- 2 1st nucleus: only couple
- 3 1st nucleus: couple with children
- 4 1st nucleus: father or mother with children
- 5 2nd and 3rd nuclei: only couple
- 6 2nd and 3rd nuclei: couple with children
- 7 2nd and 3rd nuclei: father or mother with children

EDAD. Five-year age bracket

- 1 0 – 4 years old
- 2 5 – 9 years old
- 3 10 – 14 years old
- 4 15 – 19 years old
- 5 20 – 24 years old
- 6 25 – 29 years old

- 7 30 – 34 years old
- 8 35 – 39 years old
- 9 40 – 44 years old
- 10 45 – 49 years old
- 11 50 – 54 years old
- 12 55 – 59 years old
- 13 60 – 64 years old
- 14 65 – 69 years old
- 15 70 – 74 years old
- 16 75 – 79 years old
- 17 80 – 84 years old
- 18 85 years old and over

EMPA. Electoral roll status

- 1 Registered in the dwelling
- 2 Not registered in the dwelling

CANA- Autonomous community of birth

- 1 Andalusia
- 2 Aragón
- 3 Asturias
- 4 Cantabria
- 5 Castilla-León
- 6 Castilla-La Mancha
- 7 Catalonia
- 8 Comunidad Valenciana
- 9 Extremadura
- 10 Galicia
- 11 Madrid
- 12 Navarra / Nafarroa
- 13 Basque Country / Euskadi
- 14 Rioja (La)
- 15 Other Autonomous Communities
- 16 Foreigner

LNAC. Place of Birth (Basque Country)

- 1 Araba
- 2 Bizkaia
- 3 Gipuzkoa
- 4 Other provinces
- 5 Foreigner

ZONN. Birth Zone

- 1 Llanada alavesa
- 2 Rest of Spain
- 3 48011 Bilbao
- 4 Duranguesado
- 5 Gernika-Bermeo+Plentzia-Mungia+Markina-Ondarroa
- 6 Arratia-Nerv+Encartaciones
- 7 Donostialdea+Bajo Bidaso
- 8 Alto Deba+Bajo Deba+Urola Costa
- 9 Tolosaldea+Goierri

NACI1. Nationality (continent)

- 1 Europe
- 2 North America
- 3 Central America
- 4 South America
- 5 Asia
- 6 Africa
- 7 Oceania

RELA1. Activity status (studying)

- 1 Studies
- 2 Does not study

RELA2. Activity status (housework)

- 1 Does housework
- 2 Does not do housework

RELA3. Activity status (retirement)

- 1 Retired, pensioner
- 2 Not retired, pensioner

RELA5. Activity status (employment)

- 1 Employed
- 2 Unemployed
- 3 Inactive

RELA8. Relation to activity (unemployed)

- 1 Unemployed looking for first job
- 2 Unemployed but has worked previously.

CATR. Employment Autonomous Community

- 1 Andalusia
- 2 Aragón
- 3 Asturias
- 4 Cantabria
- 5 Castilla-León
- 6 Castilla-La Mancha
- 7 Catalonia
- 8 Comunidad Valenciana
- 9 Extremadura
- 10 Galicia
- 11 Madrid
- 12 Navarra / Nafarroa
- 13 Basque Country / Euskadi
- 14 Rioja (La)
- 15 Other Autonomous Communities
- 16 Foreigner

ZONT. Employment zone

- 1 Llanada alavesa
- 2 Rest of Alava
- 3 48011 Bilbao
- 4 Duranguesado
- 5 Gernika-Bermeo+Plentzia-Mungia+Markina-Ondarroa
- 6 Arratia-Nerv+Encartaciones
- 7 Donostialdea+Bajo Bidaso
- 8 Alto Deba+Bajo Deba+Urola Costa
- 9 Tolosaldea+Goierri

CSEC. Socio-economic Status

- 1 Higher groups
- 2 Middle professions and jobs
- 3 Farming and craft classes
- 4 Skilled workers
- 5 Unskilled workers
- 6 Dependent inactives

LHOG2 – Housework (dedication)

- 1 Regularly, always does it
- 2 Regularly, at weekends

- 3 Occasionally
- 4 Rarely
- 5 Never

LTRA – Place of work (Basque Country)

- 1 Araba
- 2 Bizkaia
- 3 Gipuzkoa
- 4 Other provinces
- 5 Foreigner

PROF1 – Profession (CON to 1 digit)

- 1 Directors, managers
- 2 Professionals, scientists
- 3 Support technicians
- 4 Administration Employees
- 5 Shop assistants, waiters
- 6 Farmers, fishermen
- 7 Skilled workers
- 8 Machinery operators
- 9 Unskilled workers

SPRO – Professional status

- 1 Entrepreneurs
- 2 Self-employed
- 3 Cooperative members
- 4 Employees
- 5 Temporary workers
- 6 Family support
- 7 Other status

ANAF – Literacy

- 1 Yes
- 2 No

DEST – Time spent studying

- 1 Full-time
- 2 Average
- 3 Little

EKEN – Level of understanding Basque

- 1 High
- 2 Medium
- 3 Low

EKEN – Level of writing Basque

- 1 High
- 2 Medium
- 3 Low

EKEN – Level of speaking Basque

- 1 High
- 2 Medium
- 3 Low

EKEN – Level of reading Basque

- 1 High
- 2 Medium
- 3 Low

EKNG – Overall level of Basque

- 1 Basque speakers
- 2 Near Basque speaker
- 3 Non-Basque speaker

LENM –Mother tongue

- 1 Basque
- 2 Spanish
- 3 Both
- 4 Other

NIVI. Education level (key groups)

- 1 Illiterate
- 2 No studies
- 3 Pre-school and primary
- 4 Vocational Training
- 5 Secondary
- 6 Upper vocational training
- 7 Higher education

ALLE. Year when arrived to the current town of residence (five-year bracket)

- 1 ≤1950
- 2 1951-1955
- 3 1956-1960
- 4 1961-1965
- 5 1966-1970
- 6 1971-1975
- 7 1976-1980
- 8 1981-1985
- 9 1986-1990
- 10 1991-1995
- 11 1996-2000
- 12 2001-2005
- 13 2006-2007

CAPR. Autonomous community of origin

- 1 Andalusia
- 2 Aragón
- 3 Asturias
- 4 Cantabria
- 5 Castilla-León
- 6 Castilla-La Mancha
- 7 Catalonia
- 8 Comunidad Valenciana
- 9 Extremadura
- 10 Galicia
- 11 Madrid
- 12 Navarra / Nafarroa
- 13 Basque Country / Euskadi
- 14 Rioja (La)
- 15 Other Autonomous Communities
- 16 Foreigner

ZONP. Zone of origin

- 1 Llanada alavesa
- 2 Rest of Spain
- 3 48011 Bilbao
- 4 Duranguesado
- 5 Gernika-Bermeo+Plentzia-Mungia+Markina-Ondarroa
- 6 Arratia-Nerv+Encartaciones
- 7 Donostialdea+Bajo Bidaso

- 8 Alto Deba+Bajo Deba+Urola Costa
- 9 Tolosaldea+Goierri

DURR – Length of residency (five-year bracket)

- 1 0 – 4 years
- 2 5 – 9 years
- 3 10 – 14 years
- 4 15 – 19 years
- 5 20 – 24 years
- 6 25 – 29 years
- 7 30 – 34 years
- 8 35 – 39 years
- 9 40 – 44 years
- 10 45 – 49 years
- 11 50 – 54 years
- 12 55 – 59 years
- 13 60 – 64 years
- 14 65 – 69 years
- 15 70 – 74 years
- 16 75 – 79 years
- 17 80 – 84 years
- 18 85 years and over

EDMG1 – Age when migrated (five-year brackets)

- 1 0 – 4 years old
- 2 5 – 9 years old
- 3 10 – 14 years old
- 4 15 – 19 years old
- 5 20 – 24 years old
- 6 25 – 29 years old
- 7 30 – 34 years old
- 8 35 – 39 years old
- 9 40 – 44 years old
- 10 45 – 49 years old
- 11 50 – 54 years old
- 12 55 – 59 years old
- 13 60 – 64 years old
- 14 65 – 69 years old
- 15 70 – 74 years old
- 16 75 years old and over

LPRO. Place of origin (Basque Country)

- 1 Araba
- 2 Bizkaia
- 3 Gipuzkoa
- 4 Other provinces
- 5 Foreigner

LR10. Place of residency 10 years earlier (Basque Country)

- 1 Araba
- 2 Bizkaia
- 3 Gipuzkoa
- 4 Other provinces

CAMB- Changes of residency

- 1 1 change
- 2 2 changes
- 3 3 or more changes

NACI. Nationality

- 1 Spanish

- 2 Spanish and other
- 3 Other

SEDE. Sedentary

- 1 Yes
- 2 No