

Ficheiro de Uso Público - Censos 2021

Introdução

Este documento descreve a metodologia de constituição da amostra de microdados relativa a indivíduos e alojamentos dos Censos 2021 que é disponibilizada como Ficheiro de Uso Público.

O PUF dos Censos 2021 contempla duas tabelas amostrais, uma de indivíduos residentes e outra de alojamentos familiares e coletivos contendo algumas variáveis do edifício, ambas com 5% de registos e uma variável comum de ligação.

Amostra de microdados

A elaboração desta amostra surge da necessidade de informação censitária, por parte dos utilizadores, em níveis cada vez mais desagregados e sob a forma de ficheiros de microdados que viabilizem a realização de cruzamentos e análises à medida das suas necessidades.

A prática internacional que tem vindo a ser implementada em vários países, colocando à disposição dos utilizadores amostras de microdados dos Censos e os avanços técnicos e tecnológicos em matéria de “software” para tratamento e salvaguarda da confidencialidade, tornam atualmente possível a disponibilização de ficheiros de microdados com segurança no cumprimento das normas e legislação em vigor sobre confidencialidade dos dados individuais.

O INE, IP, disponibiliza um ficheiro de microdados que inclui duas tabelas amostrais, uma de indivíduos residentes e outra de alojamentos familiares e coletivos contendo algumas variáveis do edifício, ambas com 5% de registos e uma variável comum de ligação.

Metodologia de construção da amostra

De forma a garantir a representatividade de todos os tipos de alojamento efetuou-se uma seleção por etapas.

Etapas 1 – Para cada NUTS II, definiu-se uma amostra de 5% do total de alojamentos familiares, com base num processo de seleção sistemática com um passo fixo a inicializar no 1º alojamento de cada NUTS II, após ordenação. Para tal, os alojamentos foram ordenados por uma variável “ID_ORDA” que resulta da concatenação das variáveis NUTS1 + NUTS2 + NUTS3 + DT_COD + MN_COD + FR_COD + SEC_COD + SS_COD + N_EDIF + N_ALOJ;

Etapas 2 – A amostra de indivíduos residentes nos alojamentos familiares será composta por todos os indivíduos pertencentes aos alojamentos selecionados durante a etapa 1;

Etapas 3 – Para cada NUTS II, definiu-se uma amostra de 5% de cada um dos seguintes grupos populacionais: indivíduos residentes em alojamentos coletivos; corpo diplomático e pessoal das forças armadas em missão oficial no estrangeiro; população embarcada; população sem-abrigo. A seleção dos indivíduos pertencentes aos grupos populacionais indicados foi efetuada de modo sistemático, para cada NUTS II, sendo os registos ordenados por uma variável “ID_ORDA” que resulta da concatenação das variáveis NUTS1 + NUTS2 + NUTS3 + DT_COD + MN_COD + FR_COD + SEC_COD + SS_COD + N_EDIF + N_ALOJ + N_AGR + N_NUCL + N_IND;

Etapas 4 – Selecionaram-se os alojamentos correspondentes aos indivíduos residentes selecionados na etapa 3.

Após a realização destas 4 etapas, efetuou-se a junção dos registos em duas tabelas, uma de alojamentos (resultantes da etapa 1 e da etapa 4) e uma de indivíduos residentes (resultantes da etapa 2 e da etapa 3).

Para a constituição das duas tabelas foram selecionadas as variáveis mais relevantes, cumprindo a anonimização recomendada e as regras relativas ao segredo estatístico.