

Centro de Estudos da Metrópole
Notas sobre a nova divulgação dos dados do Censo Demográfico de 2010
(25/05/2016)

- Trata-se da versão mais recente dos microdados da amostra do Censo Demográfico de 2010, divulgada pelo IBGE em 11/03/2016. Traz eventuais correções de informações, além de novas variáveis. Em especial, foram adicionadas as seguintes informações:
 - **V1005 – Situação do setor censitário**
 - **V5040 - INDICADORA DE FAMÍLIA**
 - **V5090 - TIPO DE COMPOSIÇÃO FAMILIAR DAS FAMÍLIAS ÚNICAS E CONVIVENTES PRINCIPAIS**
 - **V5100 - TIPO DE COMPOSIÇÃO FAMILIAR DAS FAMÍLIAS CONVIVENTES SECUNDÁRIAS**
 - **V5130 - ORDEM LÓGICA NA FAMÍLIA**
 - **V6462 - Qual era a ocupação que exercia no trabalho que tinha?** *(Codificada conforme a CBO-Domiciliar, utilizada no Censo de 2000 e nas PNADs de 2002 a 2014)*
 - **V6472 - Qual era a atividade principal do empreendimento em que tinha esse trabalho?** *(Codificada conforme a CNAE-Domiciliar, utilizada no Censo de 2000 e nas PNADs de 2002 a 2014)*
- O banco de dados traz registros (linhas/casos) de pessoas, emigração e mortalidade. As informações sobre domicílios figuram como variáveis nos casos.
- **IMPORTANTE:** Criamos e adicionamos ao banco de dados a variável **tipo_registro**, que informa o tipo de registro de cada linha: **Pessoas, Emigração ou Mortalidade. É recomendável selecionar apenas o tipo desejado em cada análise particular.** Em geral, informações específicas de Pessoas apenas trazem informações válidas para os casos relevantes; e o mesmo ocorre com Emigração e Mortalidade. No entanto, algumas variáveis de identificação (como UF, Código do Município etc) e as variáveis de domicílio trazem informações para todos os casos. Incluir todos os casos altera substantivamente os resultados, produzindo erros.
- Optamos, desta vez, por manter as variáveis de "Marca de imputação", que informam quando uma informação foi gerada e atribuída a partir de algum método de imputação de *missings*. Isso ocorre, por exemplo, no caso de domicílios fechados e por consequência do tratamento de informações discrepantes (ver as notas técnicas do IBGE: 1- "Estudo e tratamento dos rendimentos", 2 – "Notas Metodológicas - Microdados da Amostra-2010"). Na versão divulgada pelo Centro de Estudos da Metrópole anteriormente, havíamos deletado as colunas relativas às variáveis que indicavam a imputação, com vistas a tornar o arquivo menor. Com elas, agora, o arquivo de dados possui aproximadamente 10gb. Apesar de ocupar mais espaço em disco, o SPSS não experimenta pior desempenho nas análises.
- Não há labels nas variáveis cujas informações são descritas pelos arquivos auxiliares (ex.: ocupação, atividade econômica, cursos do ensino superior, municípios etc.). Essas informações podem ser obtidas na pasta "Anexos Auxiliares"