

Compatibilização dos Censos Demográficos

PUC-Rio / Departamento de Economia

Este documento descreve as variáveis obtidas a partir do pacote `datazoom_censo`, que compatibiliza as informações contidas nos microdados dos Censos Demográficos de 1970 a 2010, e discute os procedimentos empregados para obtê-las.

Apesar de os sucessivos Censos Demográficos (Censos) serem semelhantes quanto aos temas gerais abordados, há pouca consistência entre os respectivos conjuntos de microdados. Isto se deve, em primeiro lugar, às mudanças ocorridas no conjunto de informações coletadas: inclusão e exclusão de variáveis, alterações no nível de detalhe com o qual os quesitos são investigados etc. Em geral, esta evolução se dá no sentido de ampliar o conjunto de informações pesquisadas para domicílios e pessoas. Nesse sentido, o conteúdo obtido em anos mais recentes pode ser usado para reconstituir a informação do tipo coletada nos anos anteriores.

Além disso, de ano para ano ocorrem alterações substanciais na forma de registro das informações, como a mudança dos nomes de todas variáveis e a reordenação de categorias de diversos quesitos. Assim, mesmo quando informação idêntica está disponível para anos distintos, é necessário certo esforço para obter dados comparáveis.

A compatibilização realizada pelo pacote `datazoom_censo` consiste na construção, a partir dos microdados dos Censos 1970, 1980, 1991, 2000 e 2010, de variáveis que podem ser imediatamente comparadas entre os anos. As bases de dados geradas pelo pacote foram elaboradas buscando abranger todas as informações que pudessem ser obtidas, de forma conceitual muito próxima, para domicílios ou pessoas em pelo menos dois dos anos considerados. Entre outras ações, o procedimento definiu variáveis com nomes informativos de seus conteúdos, e uniformizou o registro das categorias (mesma forma de registrar respostas do tipo “sim/não”, mesmos códigos para registrar estado de residência e de nascimento etc.).

A seguir, são apresentadas as variáveis construídas, juntamente com detalhes sobre os procedimentos adotados para obtê-las e, quando necessário, qualificações ao grau de compatibilidade obtido.

Identificação de domicílios

Frequentemente é preciso associar a determinado indivíduo as características do domicílio em que ele reside. Para tanto, é fornecida, no arquivo de domicílios e de pessoas, a

variável 'id_dom', que identifica as observações no arquivo de 'domicílios' permitindo relacioná-las aos respectivos residentes no arquivo de 'pessoas'. Variáveis de identificação para domicílios encontram-se prontamente disponíveis a partir do Censo 1980. Nesses anos 'id_dom' recebeu o valor destas variáveis, mas foi necessário construir a identificação para 1970 (para detalhes, veja o quadro abaixo).

Observação: a variável 'id_dom' simplesmente permite, com respeito aos dados de um determinado ano, associar os domicílios a seus moradores. Não é possível identificar um mesmo domicílio ou pessoa em anos diferentes.

Organização dos dados e identificação dos domicílios nos Censos de 1970 a 2010

Com respeito à separação entre informações de domicílios e pessoas e à forma de relacionar estes dois tipos de unidades, há várias diferenças entre os arquivos de microdados originais dos sucessivos censos.

Nos Censos 1970 e 1980, cada arquivo de dados (que corresponde a uma UF) contém informações tanto sobre os domicílios da amostra quanto sobre os indivíduos residentes. Cada linha do arquivo diz respeito a um indivíduo, e os dados de um domicílio aparecem (repetidamente) como parte do registro de cada um de seus residentes. Quanto à identificação, o Censo 1980 apresenta explicitamente a variável "identificação do domicílio", que se repete para todos os moradores, mas o mesmo não ocorre para 1970.

A construção de 'id_dom' para 1970 utiliza a identificação do domicílio implícita na estrutura dos dados. Mais especificamente, os moradores de um mesmo domicílio encontram-se em linhas consecutivas, e em cada domicílio particular o primeiro indivíduo é o chefe da família principal ou única. Assim identifica-se o domicílio de residência de um indivíduo através das seguintes regras: (1) se o indivíduo reside em domicílio particular e é chefe da família principal ou única, então ele está num domicílio diferente daquele em que reside o indivíduo da linha anterior; (2) se o indivíduo reside em domicílio coletivo e na linha anterior há um residente de domicílio particular, então novamente temos que estão em domicílios diferentes; (3) nos demais casos, o domicílio é o mesmo da linha anterior.

Em 1991 permanece o uso de arquivos contendo simultaneamente as informações de domicílios e pessoa;, no entanto os domicílios passam a ser representados em linhas separadas. As linhas de domicílios e de pessoas diferem quanto ao conteúdo, mas possuem uma parte comum, que contém (entre outras informações) a indicação do tipo de registro e a "identificação do questionário" (que permite fazer a correspondência entre cada domicílio e seus moradores).

Finalmente, em 2000 e 2010, os microdados de pessoas e domicílios se apresentam em arquivos separados. O "número de controle", informado em ambos os arquivos, identifica o domicílio ao qual um indivíduo pertence.

Compatibilização de dados de domicílio

C.1. Situação do domicílio

Em 1970, os domicílios do Censo eram classificados, com respeito à situação, em três categorias, conforme se localizassem em “área urbana”, “suburbana” ou “rural”, conforme delimitadas pela administração municipal. Nos anos seguintes, foi eliminada a classificação “suburbana”, que havia sido incorporada legalmente à área urbana. Além disso, a situação do domicílio passou a combinar o elemento legal (localização no quadro urbano ou rural do município) com outros geográficos (contiguidade, presença de serviços, infraestrutura).

Nos arquivos compatibilizados, utilizando apenas os Censos 1991 e 2000, é possível manter a classificação mais desagregada:

```
* sit_setor = 1 - Área urbanizada de vila ou cidade
*           2 - Área não urbanizada de vila ou cidade
*           3 - Área urbanizada isolada
*           4 - Rural - extensão urbana
*           5 - Rural - povoado
*           6 - Rural - núcleo
*           7 - Rural - outros aglomerados
*           8 - Rural - exclusive os aglomerados rurais
```

A compatibilização com os dados de 1980 exige a agregação das categorias 1 e 2 como “Urbana – vila ou cidade” e 4 a 7 como “Aglomerado rural”:

```
* sit_setor_B = 1 - Vila ou cidade
*             2 - Urbana isolada
*             3 - Aglomerado rural
*             4 - Rural exclusive os aglomerados
```

Finalmente, na compatibilização com 1970 e 2010, só é possível distinguir a área urbana e a rural. A distinção entre as duas áreas está prontamente disponível de 1980 em diante, enquanto para os dados de 1970 foi necessário agregar a categoria “suburbana” à situação urbana.

```
* sit_setor_C = 1 - Urbana
*             0 - Rural
```

C.2. Espécie do domicílio

Foi mantida a classificação adotada nos sucessivos Censos. Como somente em 1980 era possível avaliar se um domicílio coletivo era permanente ou improvisado, a compatibilização agrega estas duas categorias.

```
* especie = 0 - particular permanente
*          1 - particular improvisado
*          2 - coletivo
```

C.3. Material das paredes

Este item foi investigado apenas em 1980, 1991 e 2010. Em 1980 e 1991, as alternativas são as mesmas, apesar de possuírem códigos diferentes. Os códigos de 1980 foram alterados para se adequar aos de 1991. Em 2010, as alternativas são mais desagregadas, mas não há maiores problemas em agrupá-las para o formato encontrado nos anos anteriores: 'alvenaria com ou sem revestimento' = 'alvenaria'; 'taipa revestida ou não revestida' = 'taipa'; 'madeira aproveitada' = 'material aproveitado'. A alternativa 'sem parede' foi considerada *missing*.

```
* paredes = 1 Alvenaria
*         = 2 Madeira aparelhada
*         = 3 Taipa não revestida
*         = 4 Material aproveitado
*         = 5 Palha
*         = 6 Outro
```

C.4. Material da cobertura

Essa variável existe apenas para 1980 e 1991, com uma diferença nos códigos das alternativas. Além disso, em 1980, há a alternativa 'telha de amianto', e, em 1990, a alternativa é 'telha de cimento-amianto'.

```
*cobertura = 1 laje de concreto
*          = 2 telha de barro
*          = 3 telha de amianto
*          = 4 zinco
*          = 5 madeira aparelhada
*          = 6 palha
*          = 7 material aproveitado
*          = 8 outro material
```

C.5. Tipo de domicílio

Desde o Censo de 1991, é possível distinguir três tipos de domicílio: "casa", "apartamento" e "cômodo". No arquivo com os anos 1991 e 2000 esta informação se apresenta da seguinte forma:

```
* tipo_dom = 1 - casa
*           2 - apartamento
*           3 - cômodo
```

Em 1980 só existiam as categorias "casa" e "apartamento". Como esta última era atribuída aos domicílios que posteriormente seriam classificados como cômodos (IBGE, 1980, p. 32), a compatibilização com o ano de 1980 requer a agregação de 2 e 3.

```
* tipo_dom_B = 1 - casa
*             2 - apartamento (ou cômodo)
```

Em 1970 não se averiguou o tipo do domicílio. Já em 2010, duas particularidades. Para domicílios particulares permanentes, há a alternativa 'oca ou maloca' para áreas indígenas,

que foi codificada como 'casa'. E diferentemente dos outros anos, investigou-se o tipo de domicílio para domicílios improvisados ('tenda', 'trailer', etc.) ou coletivos ('asilo', 'penitenciária', etc.). Esses foram considerados *missing*.

Em 1991 e 2000, adicionalmente ao tipo de domicílio, é investigado o tipo de setor em que este se localiza, permitindo identificar os chamados aglomerados subnormais (favelas e outras formas de ocupação desordenada).

```
* subnormal = 0 - não
*             1 - sim
```

Por consistência, esta variável foi construída apenas para os domicílios particulares permanentes dos tipos "casa" e "apartamento" (mesmo estando disponível para a totalidade da amostra em 2000), uma vez que em 1991 a informação necessária somente é aferida para este conjunto.

C.6. Condição de ocupação e aluguel

Em todos os anos considerados, os domicílios são divididos, quanto à condição de ocupação, em "próprio", "alugado", "cedido" e "outra condição". Além disso, desde 1980 diferencia-se "cedido pelo empregador" de "cedido de outra forma". Para estes anos, temos:

```
* cond_ocup = 1 - próprio
*             2 - alugado
*             3 - cedido por empregador
*             4 - cedido de outra forma
*             5 - outra condição
```

Para a compatibilização com o Censo 1970, as categorias 3 e 4 são agregadas:

```
* cond_ocup_B = 1 - próprio
*              2 - alugado
*              3 - cedido
*              4 - outra condição
```

Para 2000 e 2010, a condição de ocupação "próprio" é desagregada em duas categorias, indicando se já pagou pelo domicílio ou se continua pagando. Temos:

```
* cond_ocup_C = 1 - próprio, já pago
*              2 - próprio, ainda pagando
*              3 - alugado
*              4 - cedido por empregador
*              5 - cedido de outra forma
*              6 - outra condição
```

Exceto em 1991, é possível distinguir o domicílio próprio já pago daquele em aquisição. Assim, para 1970, 1980, 2000 e 2010, é apresentada:

```
* dom_pago = 0 - Domicílio próprio em aquisição
*           1 - Domicílio próprio já pago
```

Em 1991 e 2000, é avaliada também a propriedade do terreno, para o caso dos imóveis próprios; assim, está disponível:

```
* terr_proprio = 0 - não
*                1 - sim, terreno é próprio
```

O aluguel mensal pago pelo domicílio, fornecido nos Censos 1980, 1991 e 2010, é armazenado em 'aluguel'. Para 1970, esta variável está disponível apenas em intervalos de valores nominais. A variável 'aluguel_70' foi mantida na base compatibilizada, para uma eventual tentativa de compatibilização por parte do usuário.

C.7. Abastecimento de água

Para todos os anos, foi possível construir a seguinte variável:

```
* abast_agua = 1 - rede geral com canalização interna
*              2 - rede geral sem canalização interna
*              3 - poço ou nascente com canalização interna
*              4 - poço ou nascente sem canalização interna
*              5 - outra forma
```

Para 2000 e 2010, é investigada a existência de água canalizada no domicílio:

```
* agua_canal = 1 - Canalizada em pelo menos um cômodo
*              2 - Canalizada só na propriedade ou terreno
*              3 - Não canalizada
```

C.8. Instalações sanitárias

Em todos os anos foi aferida (para os domicílios particulares permanentes) a existência de acesso a alguma instalação sanitária, ainda que comum a mais de um domicílio, o que permitiu construir a variável:

```
* sanitario = 0 - não tem acesso
*              1 - tem acesso a sanitário
```

Além disso, em todas as ocasiões foi possível classificar o tipo de escoadouro das instalações sanitárias (para os domicílios que tinham acesso) nas seguintes categorias:

```
* tipo_esc_san = 1 - Rede geral
*               2 - Fossa séptica
*               3 - Fossa rudimentar
*               4 - Outro escoadouro
```

Estas correspondem exatamente às categorias de 1970 e 1980. Posteriormente houve ampliações da classificação, que são idênticas em 2000 e 2010 (para 1991, houve ampliação, mas não da forma encontrada nos anos seguintes.). Portanto, apenas para 2000 e 2010, tem-se:

```
* tipo_esc_san_B = 1 - Rede geral
*                 2 - Fossa séptica
*                 3 - Fossa rudimentar
*                 4 - Vala
*                 5 - Rio, lago ou mar
```

* 6 - Outro

Adicionalmente, em 1980 e 1991, a exclusividade da instalação sanitária foi investigada, sendo possível obter:

* sanitario_ex = 0 - não tem acesso a instalação sanitária exclusiva
* 1 - tem acesso a instalação sanitária exclusiva

Finalmente, desde 1991, é registrado o número de banheiros do domicílio (cômodo com chuveiro ou banheira e aparelho sanitário). Esta informação é usada na construção de:

* banheiros = 0 - não tem
* 1 a 4 - número de banheiros
* 5 - cinco ou mais banheiros

C.9. Destino do lixo

Somente pesquisado a partir de 1991. As mesmas categorias são empregadas em 2000 e 2010.

* dest_lixo = 1 - Coletado por serviço de limpeza
* 2 - Colocado em caçamba de serviço de limpeza
* 3 - Queimado (na propriedade)
* 4 - Enterrado (na propriedade)
* 5 - Jogado em terreno baldio ou logradouro
* 6 - Jogado em rio, lago ou mar
* 7 - Tem outro destino

C.10. Iluminação elétrica

Em todos os anos considerados, investiga-se a presença de iluminação elétrica, registrada por:

* ilum_eletr = 0 não tem
* 1 tem

Adicionalmente, em 1980 e 1991, avalia-se a presença de medidor do consumo de energia exclusivo do domicílio, mas o quesito não é claro quanto à classificação do domicílio que possui medidor coletivo, isto é, quando o medidor registra o consumo de mais de um domicílio. Essa diferenciação é explícita em 2010. Para este último ano, considerou-se o medidor coletivo como medidor exclusivo do domicílio.

* medidor_el = 0 não tem
* 1 tem

C.11. Bens de consumo duráveis

O Censo avalia a existência de diversos tipos de bens de consumo duráveis (e em alguns casos a quantidade ou características destes) nos domicílios. Como a captação destes itens não foi uniforme ao longo do tempo, nem todas as informações coletadas em determinado ano

podem ser acompanhadas nos dados dos recenseamentos seguintes. Os itens detalhados a seguir abrangem as informações que podem ser obtidas igualmente em pelo menos um par de anos cobertos pelo Portal.

Fogão – Este item não está presente nos Censos 2000 e 2010. Nos demais anos, a presença de equipamento para cozinhar é investigada juntamente com o tipo de energia utilizado. Enquanto em 1991 foi investigado o combustível (ou fonte de energia) usado nos domicílios que tivessem qualquer tipo de equipamento para cozinhar, quer fosse fogão ou fogareiro, em 1970 o quesito foi desconsiderado onde havia apenas este último. Assim os dados desses anos não são compatíveis. No entanto, em 1980 é possível avaliar separadamente a presença de fogão e de fogareiro, o que permite construir, para 1970 e 1980 as variáveis:

```
* fogao = 0 - não tem
*         1 - tem

* comb_fogao = 1 - gás
*             2 - lenha
*             3 - carvão
*             4 - outro
*             0 - não tem fogão
```

Além disso, para 1980 e 1991 é possível observar:

```
* fogao_ou_fog = 0 - não tem
*               1 - tem (fogão ou fogareiro)

* comb_cozinha = 1 - gás
*               2 - lenha
*               3 - carvão
*               4 - outro
*               0 - não tem fogão nem fogareiro
```

Rádio – Este item tem sido pesquisado de maneira uniforme entre 1970 e 2010. Temos:

```
* radio = 0 - não tem
*         1 - tem
```

Geladeira – Em 1970, 1980 e 2010 é investigada simplesmente a presença de geladeira, enquanto em 2000 afere-se a presença de geladeira ou freezer, sem discriminação. Em 1991 é possível avaliar, separadamente a presença de geladeira e de freezer. Assim, para 1970 a 1991 é possível comparar:

```
* geladeira = 0 - não tem
*             1 - tem
```

Para 1991 e 2000:

```
* gelad_ou_fre = 0 - não tem
*               1 - tem
```

Cabe ainda observar que, em 1991, só foram consideradas geladeiras elétricas (o que exclui aparelhos a gás ou querosene). Como nesse ano a investigação da presença de eletrodomésticos se restringe a domicílios com acesso à rede elétrica (ver comentário no final da seção), a distinção parece ser de pouca importância.

Telefone – Este quesito não foi investigado em 1970. Nos demais anos, diz respeito à existência de acesso a linha telefônica instalada, ainda que extensão de linha de outro domicílio. Em 2010, além de telefone fixo, foi investigada a existência de telefone celular. Apenas o telefone fixo foi considerado na construção da variável abaixo.

```
* telefone = 0 - não tem  
*           1 - tem
```

Televisão – Em 1980 e 1991, é possível avaliar, separadamente, a presença de TV em cores e em preto e branco. Para esses anos, são construídas:

```
* televisao, tv_pb, tv_cores = 0 - não tem  
*                               1 - tem
```

Em 1970, 2000 e 2010 não é feita essa distinção, e somente a primeira variável está disponível.

Automóvel – A presença de automóvel no domicílio é investigada em todos os anos, sendo excluídos automóveis de carga, mas incluídos os utilitários. Em 1970 e 2010, somente é avaliada a presença de automóvel para uso particular. Em 1980 registra-se disponibilidade de automóvel particular e, apenas em caso negativo, de carro de trabalho. Em 1991 os dois tipos são registrados separadamente (além da quantidade dos particulares e se o de trabalho é próprio ou cedido pelo empregador). Finalmente, no Censo 2000 agregam-se os dois tipos. Em síntese, para os anos de 1980 e 1991 temos as duas variáveis abaixo, mas, para 2000, só a primeira, e para 1970 e 2010, a segunda.

```
* automovel, automov_part = 0 - não tem  
*                               1 - tem
```

Máquina de lavar – Desde 1991, investiga-se a presença de máquina de lavar roupa no domicílio. E desde 2000, de computador.

```
* lavaroupa = 0 - não tem  
*             1 - tem
```

```
* microcomp = 0 - não tem  
*             1 - tem
```

Observação: no Censo 1991 a investigação dos quesitos referentes a eletrodomésticos limitou-se a domicílios com iluminação elétrica. Portanto as variáveis ‘geladeira’, ‘gelad_ou_fre’, ‘televisao’, ‘tv_pb’ e ‘tv_cores’ do arquivo compatibilizado são ‘missing’ para os domicílios sem iluminação elétrica em 1991. Logo, para o cálculo de estatísticas comparáveis dessas variáveis nos demais anos, é necessária a restrição ao universo correspondente (ou seja, para os quais a variável ‘ilum_eletr’ vale 1).

C.12. Número de cômodos

Para todos os anos estão disponíveis as variáveis 'tot_comodos', que indica o número de cômodos do domicílio e 'tot_dorm', que representa o número de cômodos servindo como dormitório.

C.13. Renda domiciliar

Em todos os anos foram feitos cálculos dos rendimentos mensais de cada morador dos domicílios particulares. Estes, juntamente com a informação sobre a relação de cada pessoa como chefe do domicílio disponível a partir de 1980 (ver seção D.2.), permitem obter a renda domiciliar, que consiste na soma dos rendimentos individuais dos moradores, com exceção dos empregados domésticos e seus parentes e dos hóspedes e pensionistas. Este valor, que pode ser computado em 1980, 1991, 2000 e 2010, é indicado na variável 'renda_dom'. Para 1980, no entanto, é necessário construir essa variável a partir das rendas individuais. Desse modo, 'renda_dom' não está presente no arquivo de domicílios, mas somente no de pessoas.

C.14. Peso do domicílio

A variável 'peso_dom' registra o peso amostral do domicílio. Esta variável foi apenas renomeada para ter o mesmo nome em todos os censos.

C.15. Número de moradores do domicílio

As variáveis 'n_pes_dom', 'n_homem_dom' e 'n_mulher_dom' indicam o número total de moradores do domicílio, o número de moradores do sexo masculino e o número de moradores do sexo feminino, respectivamente. Para 2010, no arquivo de domicílios existe apenas a informação com respeito ao total de moradores. Assim, o número de moradores por sexo deve ser obtido por meio do arquivo de indivíduos. Isso faz com que essas variáveis não estejam presentes no arquivo de domicílios para este ano.

Compatibilização de dados de pessoa

D.1. Sexo

Para todos os anos, foi adotada a seguinte representação:

```
* sexo = 0 - feminino  
*       1 - masculino
```

D.2. Condição na família e no domicílio

Nos anos de 1980, 1991, e 2000, foi investigada a posição dos indivíduos tanto (i) relativamente à pessoa responsável pelo domicílio quanto (ii) à pessoa responsável pela

família. A cada ano, as alternativas são as mesmas para ambas as variáveis. Entre os anos, contudo, há variação no detalhamento das classificações e, de forma potencialmente mais problemática, no conceito de “família” adotado (com consequente impacto na forma em que os membros de um domicílio são separados em famílias conviventes). A definição de família é abordada no quadro ao final desta seção.

Neste quesito, o Censo 1991 é o mais completo. Suas categorias foram agregadas para se tornarem compatíveis com o Censo de 2000, resultando nas variáveis `cond_dom` e `cond_fam`, que apresentam a mesma codificação, mostrada a seguir.

```
* cond_*** = 1 - Pessoa responsável
*           2 - Cônjuge, companheiro(a)
*           3 - Filho(a), enteado(a)
*           4 - Pai, mãe, sogro(a)
*           5 - Neto(a), bisneto(a)
*           6 - Irmão, irmã
*           7 - Outro parente
*           8 - Agregado(a)
*           9 - Pensionista
*          10 - Empregado(a) doméstico(a)
*          11 - Parente do(a) empregado(a) doméstico(a)
*          12 - Individual em domicílio coletivo
```

Para a compatibilização com 1980 são necessárias mais reclassificações, que dão origem a ‘`cond_dom_B`’ e ‘`cond_fam_B`’, cuja codificação é apresentada abaixo. A mesma classificação se aplica para o Censo 1970, com duas ressalvas. Em primeiro lugar, nessa pesquisa não foi apurada a posição no domicílio, apenas na família, ou seja, temos somente ‘`cond_fam_B`’. Segundo, não existe a categoria “parente do(a) empregado(a) doméstico(a)”, e não há na documentação uma indicação clara de como foram classificadas as pessoas nesta situação. Para o Censo 2010, não há posição na família. Por outro lado, ambas as variáveis para condição no domicílio podem ser construídas.

```
* cond_***_B = 1 - Pessoa responsável
*             2 - Cônjuge, companheiro(a)
*             3 - Filho(a), enteado(a)
*             4 - Pai, mãe, sogro(a)
*             5 - Outro parente
*             6 - Agregado
*             7 - Hóspede, pensionista
*             8 - Empregado(a) doméstico(a)
*             9 - Parente do(a) empregado(a) doméstico(a)
*            10 - Individual em domicílio coletivo
```

O conceito de “família” nos Censos

Nos Censos 1970, 1980 e 2000, considerava-se como família: (i) a pessoa que morasse sozinha; ou (ii) duas ou mais pessoas ligadas por laço de parentesco ou dependência doméstica (relação de subordinação dos empregados domésticos e dos agregados à pessoa responsável pelo domicílio); ou ainda (iii) duas ou mais pessoas vivendo em domicílio particular e ligadas por normas de convivência. Em 1970 e 80, este último caso admitia no máximo cinco pessoas; quando havia mais de seis indivíduos sem laço de parentesco nem dependência doméstica, o domicílio era considerado coletivo.

*Por esta definição, domicílios onde não residiam familiares do indivíduo responsável, mas havia empregado doméstico com parentes **não** foram divididos em famílias conviventes. Isto ocorre porque o fato de haver empregado doméstico morador invalida o caso (i), enquanto o responsável, isolado da família do empregado, não pode compor uma grupo do tipo (ii) ou (iii). Neste caso, considerou-se uma única família no domicílio, composta por “responsável”, “empregado” e “parentes do empregado...”, o que justifica esta última categoria como possível “posição na família”, como adotado em 1980 e 2000 (mas não em 70, como comentado nesta seção). (Este problema parece quantitativamente pequeno; em 1980, esta situação afetava 230 pessoas em 157 domicílios no estado do Rio de Janeiro, onde foram pesquisadas no total 1.248.287 pessoas em 342.368 domicílios.) Enquanto em 80 esta é a única circunstância em que há “parentes de empregado(a)...” na família, em 2000 ocorrem casos em que as famílias dos empregados não são separadas da família do patrão mesmo esta última tendo mais de um membro.*

Em 1991, por outro lado, admitiu-se, excepcionalmente no caso citado acima, considerar uma família composta unicamente pela pessoa responsável e outra pelo empregado e seus familiares. Nesse ano, coerentemente, o quesito “posição na família” não inclui a categoria “parente de empregado(a) doméstico(a)”. Nos demais aspectos, o conceito é idêntico ao considerado em 1970 e 80.

D.3. Idade

Os microdados de todos os Censos considerados permitem avaliar a idade em anos completos (`idade`) e, para crianças com menos de um ano, a idade em meses (`idade_meses`). Além disso, com exceção de 1980, também é reportado se a idade foi presumida (isto é, se foi fornecida diretamente, em vez de calculada a partir da informação da data de nascimento), de acordo com a seguinte variável:

```
* idade_presumida = 0 - não
*                  1 - sim
```

D.4. Cor ou raça

Este quesito não foi investigado em 1970. Em 1991, 2000 e 2010, as alternativas são as mesmas. Em 1980, não há a categoria “indígena”, e o manual do entrevistador indica que se assinala a opção “pardo” no caso de pessoas que se declarem índias.

Assim, para dados de 1991, 2000 e 2010, temos:

```
* raca = 1 - branca
*       2 - preta
*       3 - amarela
*       4 - parda
*       5 - indígena
```

Os dados compatibilizados com 1980 por sua vez apresentam:

```
* racaB = 1 - branca
```

- * 2 - preta
- * 3 - amarela
- * 4 - parda

onde os indivíduos classificados como indígenas em 1991, 2000 e 2010 foram reclassificados como pardos.

D.5. Religião

O número de categorias cresceu a cada Censo. Desde 1980, entretanto, manteve-se a possibilidade de agregar as denominações nos grupos representados pela seguinte variável:

- * religiao = 0 - sem religião
- * 1 - católica
- * 2 - evangélica tradicional
- * 3 - evangélica pentecostal
- * 4 - espírita kardecista
- * 5 - espírita afro-brasileira
- * 6 - religiões orientais
- * 7 - judaica/israelita
- * 8 - outras religiões

Para tornar a informação compatível com a de 1970 foi necessário agregar ainda mais as categorias, fazendo-as corresponder às classificações disponíveis naquele ano:

- * religiao_B = 0 - sem religião
- * 1 - católica
- * 2 - evangélica
- * 3 - espírita
- * 4 - outra

D.6. Deficiência física e mental

Os Censos 1991, 2000 e 2010 investigaram o tema. Em 1991 foi investigada a presença de sete tipos de deficiência, para indivíduos afetados por apenas um deles, ou registrada a ocorrência de “mais de um”, neste caso sendo impossível avaliar quais são estes problemas.

Nos Censos 2000 e 2010, um questionário mais abrangente foi aplicado, com questões individuais para indicar a ocorrência de cada tipo de deficiência, além de avaliar a intensidade de algumas delas. Por exemplo, enquanto em 1991 somente se avaliava a presença de cegueira total, em 2000 há também dois níveis intermediários de “dificuldade de enxergar”. Por outro lado, o questionário de 2000 reúne os três tipos de paralisia e a ausência de membro (ou parte) numa mesma questão e aponta apenas a deficiência mais grave. Assim, se um indivíduo, além de paraplégico, também tenha perdido um membro, só a paralisia é indicada.

Por conta dessas diferenças, optou-se por não compatibilizar 1991 com os demais anos. Para 2000 e 2010, têm-se as seguintes variáveis que indicam o grau de dificuldade de enxergar, ouvir e caminhar:

- * dif_enxergar, dif_ouvir, dif_caminhar = 1 não consegue de modo algum
- * 2 grande dificuldade

* 3 alguma dificuldade
* 4 nenhuma dificuldade

Além disso, existe a variável indicadora de deficiência mental:

* def_mental = 1 tem deficiência mental
* 0 não tem

D.7. Naturalidade e migração

Condição de migrante – Desde 1980, é possível saber se o indivíduo sempre morou no mesmo município. Em 1970 não é possível distinguir os indivíduos que sempre residiram em um município daqueles que, tendo morado em outras localidades, voltaram a residir no município onde nasceram. Assim, nos dados de 1980, 1991, 2000 e 2010 foi construída:

* sempre_morou = 0 - não
* 1 - sim

Adicionalmente, em 1980 e 1991 é possível verificar se o indivíduo migrou entre as zonas urbana e rural do município. Esta é a informação contida na variável:

* onde_morou = 0 só na zona urbana
* 1 só na zona rural
* 2 nas zonas urbana e rural

Nacionalidade e naturalidade – Em todos os anos foi investigada a nacionalidade dos indivíduos, permitindo construir:

* nacionalidade = 0 - brasileiro nato
* 1 - brasileiro naturalizado
* 2 - estrangeiro

Também são fornecidas informações sobre o local de nascimento. As informações captadas em todos os anos permitem saber se um indivíduo nasceu no município e estado onde reside no momento da pesquisa. Registra-se então, respectivamente:

* nasceu_mun, nasceu_UF = 0 não
* 1 sim

Para quem não nasceu na UF de residência, é informada a unidade da federação de origem. E para quem não é brasileiro nato, o país de origem (ver códigos para países no anexo):

* UF_nascim = 11-53 código da UF de nascimento
* pais_nascim = 30-98 código do país estrangeiro

Finalmente, em 1991, 2000 e 2010, foi investigado ainda o ano em que cada estrangeiro fixou residência no país, registrado na variável 'ano_fix_res'.

Última migração/tempo de moradia – Em 1970, o tempo de moradia no município somente foi investigado para quem não nasceu no município de residência atual (na época da pesquisa).

Ou seja, ao contrário dos outros anos, em 1970, para aqueles que nasceram, migraram e eventualmente retornaram ao município natal, não foi coletada essa informação. Já em 1980, infelizmente, os intervalos de tempo de moradia são diferentes daqueles observados em 1970. A partir de 1991, a variável passa a ser contínua, sendo registrada em 'anos_mor_mun'. Assim, para todos os anos, exceto 1980, foi construída:

```
* t_mor_mun_70 = 0 - menos de 1 ano
*                1 - 1 ano
*                2 - 2 anos
*                3 - 3 anos
*                4 - 4 anos
*                5 - 5 anos
*                6 - 6 a 10 anos
*                7 - 11 anos ou mais
```

E, exceto para 1970:

```
* t_mor_mun_80 = 0 - menos de 1 ano
*                1 - 1 ano
*                2 - 2 anos
*                3 - 3 anos
*                4 - 4 anos
*                5 - 5 anos
*                6 - 6 a 9 anos
*                7 - 10 anos ou mais
```

Note que, para obter uma variável compatível com intervalos idênticos para todos os anos, basta recodificar as variáveis acima para criar a alternativa "6 anos ou mais". **Importante:** para obter uma estatística compatível para todos os anos, é necessário considerar 'nasceu_mun' igual a zero, de modo a obter a mesma amostra pesquisada em 1970. Caso 1970 seja desconsiderado, esse procedimento não se faz necessário.

Os procedimentos adotados para o tempo de moradia na UF de residência atual variam ao longo dos anos. Em 1970, há informação somente para quem não nasceu no município, mas a pessoa que nasceu e sempre morou na UF de residência atual é obrigada a escolher alguma alternativa, quer será aquela que representa sua idade.

Em 1980, para esse caso há a alternativa "nasceu", que deve ser assinalada para aqueles que nasceram e sempre moraram na UF. Esse procedimento é equivalente ao adotado em 2010, embora por meio da aplicação de um questionário distinto, isto é, em 2010, quem nasceu e sempre morou na UF de residência não é questionado quanto ao tempo de moradia na UF.

Em 1991, a pessoa indica o número de anos que mora na UF, mesmo tendo nascido e sempre morado na UF de residência. A diferença para 1970 é o fato de terem ou não nascido no município de residência: enquanto em 1970 somente aqueles que não nasceram no município são questionados, em 1991, aqueles que nasceram, mas nem sempre moraram, no município também respondem ao quesito. Em 2000, apesar das diferenças no questionário, o resultado é qualitativamente igual ao de 1991, ou seja, mesmo aqueles que nasceram e sempre moraram na UF devem indicar o tempo de moradia na UF de residência.

Como em 1970 a variável não é contínua, não é possível separar quem nasceu e sempre morou na UF de quem eventualmente migrou. Assim, optou-se por adaptar os Censos 1980 e 2010 aos demais, considerando a idade como o tempo de moradia na UF para quem não há informação. Além da variável contínua 'anos_mor_UF' para 1991, 2000 e 2010, foi criada para todos os anos, exceto 1980:

```
* t_mor_UF_70= 0 - menos de 1 ano
*              1 - 1 ano
*              2 - 2 anos
*              3 - 3 anos
*              4 - 4 anos
*              5 - 5 anos
*              6 - 6 a 10 anos
*              7 - 11 anos ou mais
```

E, exceto para 1970:

```
* t_mor_UF_80 = 0 - menos de 1 ano
*              1 - 1 ano
*              2 - 2 anos
*              3 - 3 anos
*              4 - 4 anos
*              5 - 5 anos
*              6 - 6 a 9 anos
*              7 - 10 anos ou mais
```

Vale lembrar que há duas variáveis por causa da diferença nos intervalos. Novamente, para incluir 1970, devem-se considerar apenas os indivíduos para os quais 'nasceu_mun' é igual a zero.

Também foram investigados o município, a UF e o país de residência anterior para quem migrou há menos de 10 anos. Porém, essa investigação variou ao longo do tempo. Apenas em 1991 e 2010 foram pesquisadas as três localidades. Em 1970 e 2000, não foi pesquisado o município. E em 1980, apesar de haver esse quesito no questionário, nos microdados não consta uma variável indicando o país de residência anterior. Note-se ainda que, em 1970, não vale a restrição de tempo de migração, ou seja, mesmo quem migrou há 10 anos ou mais respondem aos quesitos de local de residência anterior. Aliado ao fato de que somente respondem ao tema de migração aqueles que não nasceram no município, deve-se avaliar cuidadosamente a utilização das variáveis abaixo para 1970. Os códigos dos países são os mesmos utilizados em 1970, de tal forma que países que se desmembraram foram reagrupados nessa compatibilização.

```
* mun_mor_ant = códigos dos municípios
* UF_mor_ant = 11-53 código da UF em que morava
* pais_mor_ant = 30-98 país estrangeiro especificado
```

Em 1970, 1980 e 1991, foi investigada a situação (zona urbana ou rural) do domicílio no município onde o migrante morava anteriormente.

```
* sit_mun_ant = 0 zona urbana
*              1 zona rural
```

Local de moradia há 5 anos – Desde 1991 passou a ser investigado o local em que cada indivíduo residia exatamente 5 anos antes da data da pesquisa. Logo foi possível a construção das variáveis ‘mun_mor5anos’, ‘UF_mor5anos’ e ‘pais_mor5anos’, que informam o município, a UF e o país em que o indivíduo residia nessa data. Para 1991 e 2000, também foi averiguada a situação do domicílio em que as pessoas residiam cinco 5 anos antes, indicada em:

```
* sit_dom5anos = 1 zona urbana
*              0 zona rural
```

D.8. Educação

Em todos os anos, a alfabetização de cada indivíduo com cinco anos de idade ou mais foi indagada. No Censo 2000 esta informação foi solicitada para todos os indivíduos; para compatibilizar, essa informação para os menores de cinco anos foi considerada missing. É considerado alfabetizado o indivíduo que, na data da pesquisa, for capaz de ler e escrever pelo menos um bilhete simples no idioma que conhecer. Esta informação consta na variável ‘alfabetizado’ dos dados compatibilizados.

```
* alfabetizado = 0 - não
*              1 - sim
```

Com respeito à escolaridade, sempre foi pesquisada a frequência à escola, embora o conteúdo deste quesito tenha se modificado ao longo do tempo. Em todos os anos, foram excluídos cursos rápidos de especialização profissional e cursos de extensão cultural. Foi considerado como frequência à escola o fato de o indivíduo estar regularmente matriculado, ainda que afastado por motivos como doença ou férias escolares. Em 1970, excluíram-se também os indivíduos que apenas frequentavam jardim de infância. Este grupo passou a ser considerado como frequentador de escola em 1980 e 1991, anos em que, por outro lado, não se considerava atendimento a creche. Além disso, em 1980, foi considerada a “frequência” das pessoas que assistiam a aulas ministradas por rádio ou TV como forma de preparação para exame de aptidão escolar, condição desconsiderada nos anos seguintes. No Censo 2000 foi investigada simultaneamente a frequência a creche ou escola. Finalmente, no Censo 2010, a frequência a cursos pré-vestibulares foi desconsiderada.

Por conta dos fatos acima, para construir uma variável compatível em todos os Censos, deve-se considerar aqueles que frequentam creche, pré-escola e curso pré-vestibular como não frequentadores de escola (além daqueles que fazem curso supletivo por rádio e TV em 1980). A variável ‘freq_escola’ foi construída dessa forma.

```
* freq_escola = 0 - não
*              1 - sim
```

A variável 'freq_escolaB' inclui quem frequenta pré-escola na alternativa '1'. Nesse caso, não é compatível com 1970, mas vale para todos os demais Censos.

```
* freq_escolaB = 0 - não
*                1 - sim
```

Em 2000 e 2010, também foi investigada a rede de ensino frequentada pelo aluno, registrada na variável 'rede_freq' :

```
* rede_freq = 0 - rede privada
*            1 - rede pública
```

Exceto em 1991, é possível identificar se o aluno frequenta escola no mesmo município em que reside. Criou-se a variável 'mun_escola'. Vale lembrar que, para compatibilizar com 1970, é preciso considerar os casos em que 'freq_escola' é igual a 1, e não a variável 'freq_escolaB'.

```
* mun_escola = 0 - não
*            1 - sim, escola localiza-se no município de residência
```

O nível de escolaridade das pessoas foi investigado de forma distinta em diferentes anos. Em 1970 é perguntada a série mais elevada concluída com aprovação para cada indivíduo. Nos Censos 1980 e 2000, esta informação é pesquisada apenas para indivíduos que não frequentam escola, enquanto os que frequentam informam a série e o grau atualmente cursados. Assim, um indivíduo que se matricula novamente num nível de ensino inferior ao mais elevado que cursou “perde” (no sentido da mensuração efetuada no Censo) esta escolaridade mais avançada. Em 1991 segue-se o procedimento de 1980 e 2000, mas há uma instrução explícita no manual do entrevistador para a recuperação da informação do curso mais elevado em casos como esse.

Logo, nos Censos 1970 e 1991 é possível calcular o número de anos de estudo de cada pessoa, considerando o nível e a série mais altos cursados. Esta é a informação contida na variável 'anos_estudo'. Para 1980 e 2000 é possível realizar um cálculo alternativo para as pessoas que frequentam escola, considerando como série final aquela que precede imediatamente o nível atualmente cursado. A variável 'anos_estudoB' combina esse conteúdo com a informação obtida da primeira forma para os não estudantes.

Em 2010, a mesma “perda” ocorrida em 1980 e 2000 se repete; porém, se o curso frequentado é superior de graduação, o entrevistado responde se já concluiu outro curso de graduação. Por outro lado, tanto para quem frequenta curso superior quanto para quem não frequenta escola, não é investigado qual a série ou o ano atualmente cursado ou mais elevado que fora frequentado anteriormente. Dessa forma, é impossível construir uma variável “contínua” de anos de estudo para 2010. Para haver uma variável de escolaridade para 2010 compatível com outros Censos, criou-se uma variável de grupos de anos de estudo, válida para todos os anos, com exceção de 1970 (já que o cálculo considera a variável anos_estudoB).

* anos_estudoC = 0 - sem instrução ou menos de 3 anos de estudo (primário incompleto)

* 1 - de 4 a 7 (fundamental/ ginásio/ 1°. Grau/ médio primeiro ciclo incompleto)

* 2 - de 8 a 10 (médio/ 2°. Grau/ médio segundo ciclo incompleto)

* 3 - de 11 a 14 (médio/ 2°. Grau/ médio segundo ciclo completo ou superior incompleto)

* 4 - 15 ou mais (superior completo, mestrado, doutorado)

Todos os Censos possuem uma variável indicando o curso superior concluído. Ao longo dos anos, novos cursos foram criados, outros tiveram seu nome alterado e houve aqueles com mudanças substanciais em seu conteúdo, tornando a compatibilização uma tarefa além dos objetivos deste projeto. Houve uma tentativa de agrupar os cursos por área, utilizando as próprias denominações utilizadas pelo IBGE. Duas classificações foram criadas, uma utilizando classificações mais ‘antigas’ e outra mais ‘modernas’, esta última seguindo a classificação CONCLA. Além disso, a variável original foi mantida, com o nome ‘curso_concl’.

* cursos_c1 = 3 ciências humanas
 * 4 ciências biológicas
 * 5 ciências exatas
 * 6 ciências agrárias
 * 7 ciências sociais
 * 8 militar
 * 9 outros cursos

* cursos_c2 = 1 Educação
 * 2 Artes, Humanidades e Letras
 * 3 Ciências Sociais, Administração e Direito
 * 4 Ciências, Matemática e Computação
 * 5 Engenharia, Produção e Construção
 * 6 Agricultura e Veterinária
 * 7 Saúde e Bem-Estar Social
 * 8 Militar
 * 9 Outros

D.9. Situação conjugal

O Censo 1991 apresenta o questionário mais extenso, investigando duração da situação conjugal atual, e idade em que o indivíduo teve sua primeira união, além do estado conjugal na data da pesquisa. Em 1970 e 1980, somente o estado conjugal foi obtido, embora – como será discutido adiante – com uma pequena diferença conceitual. O estado conjugal também pode ser obtido para 2000 e 2010 (como veremos a seguir), de modo que para todos os anos foi construída a variável:

* estado_conj = 1 casamento civil e religioso
 * 2 só casamento civil
 * 3 só casamento religioso
 * 4 união consensual

- * 5 solteiro
- * 6 separado(a)
- * 7 desquitado(a)/separado(a) judicialmente
- * 8 divorciado(a)
- * 9 viúvo(a)

Duas qualificações devem ser observadas quanto à construção desta variável. Em primeiro lugar, os conceitos de “separado (a)” e “solteiro (a)” empregados em 1970 e 1980 diferem dos considerados mais recentemente. Enquanto naqueles anos considerou-se como “separada” apenas a pessoa que teve casamento civil e/ou religioso e não vive com cônjuge, a partir de 1991 também passa a constar nessa situação quem terminou uma união consensual. Assim, um indivíduo que teve união consensual mas não se casou e, posteriormente, deixou de viver com cônjuge é classificado, até 1980, como “solteiro”, e desde 1991 como “separado”.

Outra questão a se considerar é que o estado conjugal não foi diretamente investigado nos Censos 2000 e 2010, sendo necessária a obtenção indireta, a partir das informações disponíveis. Nesses anos, foi perguntado se o indivíduo vive ou já viveu com cônjuge, a natureza da última união e o estado civil corrente. Como o *estado civil* (definido como situação jurídica) não coincide com o *estado conjugal* (que considera se indivíduos vivem no mesmo domicílio), foi adotado o seguinte procedimento para obter esta informação nos dados de 2000 e 2010:

- indivíduos que nunca viveram com cônjuge têm *estado conjugal* “solteiro(a)”;
- indivíduos que vivem com cônjuge no momento da pesquisa, “casado(a)”;
- indivíduos que não vivem, mas viveram anteriormente, com cônjuge:
 - têm *estado conjugal* “divorciado(a)”, ou “desquitado(a)”, ou “viúvo(a)”, caso seja o nome do seu *estado civil*;
 - caso contrário, “separado(a)”.

Estas regras de atribuição de *estado conjugal* equivalem às prescritas pelo manual do entrevistador de 1991, exceto por ser impossível distinguir um indivíduo que teve união estável e tenha se separado daquele cujo cônjuge tenha morrido.

Nos dados compatibilizados também consta a variável ‘vive_conjuge’, derivada do estado conjugal, que informa se a pessoa vive em companhia de cônjuge. Para os anos de 1991, 2000 e 2010 (mas não para 1970 e 1980) também foi possível determinar quais indivíduos viveram, em algum momento, em companhia de cônjuge, informação registrada na variável ‘teve_conjuge’.

- * vive_conjuge, teve_conjuge = 0 - não
- * 1 - sim

D.10. Atividade econômica e renda

Até 1991, o foco da investigação da atividade econômica dos indivíduos foi a atividade principal nos doze meses anteriores à pesquisa. Em 2000 e 2010, quesitos análogos aos dos anos anteriores foram aplicados à atividade econômica exercida na semana anterior. Assim, a

maior parte das informações obtidas nesses anos não são diretamente comparáveis com as coletadas anteriormente.

Atividade econômica ao longo do último ano (Censos 1970 a 1991)

De 1970 a 1991, foram identificados os indivíduos que trabalharam nos doze meses anteriores à pesquisa, período de referência dos quesitos de emprego e renda. Considerou-se como trabalho a atividade econômica que visava remuneração (em dinheiro, produtos ou mercadorias) e aquela sem remuneração que fosse: (i) realizada em auxílio a morador do mesmo domicílio; (ii) colaborando com instituições de caridade, beneficentes, sociais ou de cooperativismo; ou (iii) como estagiário ou aprendiz. Nos anos de 1980 e 1991, como requerimento adicional, a atividade não remunerada somente poderia ser classificada como trabalho caso fosse exercida por pelo menos 15 horas semanais.

Para os trabalhadores, foram investigadas, além de características gerais do seu trabalho – isto é, considerando todas as ocupações exercidas –, algumas informações mais detalhadas da ocupação habitual (ou principal). Este conceito corresponde ao trabalho no qual o indivíduo passou a maior parte do período de referência ou à ocupação mais recente, quando ocorreu mudança com intenção definitiva. No caso de pessoas que exerciam simultaneamente mais de uma ocupação, adotaram-se como critérios para definir a principal: (i) aquela em que se ocupava por mais horas semanais e, em caso de igualdade, (ii) a que tivesse maior remuneração.

Foram registradas a ocupação habitual do trabalhador e a atividade (finalidade ou ramo de negócio) do estabelecimento onde ele exercia essa ocupação (ou o “principal” estabelecimento, de acordo com os mesmos critérios usados na ordenação das ocupações). Essas informações são apresentadas nas variáveis denominadas ‘ocup_hab’ e ‘ativ_hab’. Estes quesitos possuem codificação distinta ente os anos: 1980 e 1991 possuem os mesmos códigos para as mesmas ocupações e atividades, mas são diferentes dos códigos de 1970. Porém, os quesitos foram mantidos porque podem ser utilizados na elaboração de agregados compatíveis entre os anos. Uma das possibilidades é identificar os grupos de ocupações e setores de atividades, o que foi feito para os três anos, nas variáveis ‘grp_ocup_hab’ e ‘set_ativ_hab’, descritas abaixo.

```
* grp_ocup_hab = 1 administrativas
*               2 técnicas, científicas, artísticas e assemelhadas
*               3 agropecuária e da prod. extrativa vegetal e animal
*               4 produção extrativa mineral
*               5 indústrias de transformação e construção civil
*               6 comércio e atividades auxiliares
*               7 transportes e comunicações
*               8 prestação de serviços
*               9 defesa nacional e segurança pública
*               10 outras / mal definidas / não declaradas

* set_ativ_hab = 1 ativ. agropecuárias, de extração vegetal e pesca
*               2 indústria de transformação
*               3 indústria da construção civil
*               4 outras ativ. industriais (extração min. e serviços
```

```

*      industriais de utilidade pública)
*      5 comércio de mercadorias
*      6 transporte e comunicação
*      7 serv. aux. da ativ. econ. (técnico-profissionais
*      e auxiliares das atividades econômicas)
*      8 prest. de serv. (aloj. e alimentação, reparação e
*      conservação, pessoais, domiciliares e diversões)
*      9 social (comunitárias, médicas, odontológicas e
*      ensino)
*      10 adm. pública, defesa nacional e segurança pública
*      11 outras ativ. (instituições de crédito, seguros e
*      capitalização, com. e adm. de imóveis e valores
*      mobiliários, org. internacionais e representações
*      estrangeiras, ativ. não compreendidas nos demais
*      ramos e ativ. mal definidas ou não declaradas)

```

De 1970 a 1991 também foi pesquisada a posição na ocupação, isto é, a relação do trabalhador com o estabelecimento (principal) em que realiza a atividade. Geralmente, os indivíduos puderam ser divididos em “empregados”, “trabalhadores autônomos ou por conta-própria” e “empregadores”. Além disso, receberam tratamento especial, compondo categorias à parte:

- Os “trabalhadores agrícolas volantes”, pessoas que não tinham trabalho fixo, prestando serviço em estabelecimentos da agricultura ou extração vegetal, remunerados por tarefa, dia ou hora. Em 1980 foram divididos em “sem” e “com intermediário”, conforme eram pagos pelos proprietários dos estabelecimentos ou por agentes intermediários.
- Os “parceiros ou meeiros”, indivíduos que trabalhavam no setor primário e que recebiam parte da produção como pagamento pelo trabalho ou pagavam com parte da produção pelo acesso à terra, trecho de garimpo ou embarcação para pesca. Em 1980 foram subdivididos em “empregadores”, “conta-própria” e “empregados”; em 1991 os dois últimos grupos foram mantidos como categorias separadas, mas o primeiro foi agregado aos demais empregadores.
- e os “trabalhadores domésticos” (identificados pelo código da ocupação), somente em 1991, quando foram divididos ainda em “empregados”, quando trabalhavam diariamente para um só patrão, e “autônomos ou conta-própria” quando trabalhavam normalmente para mais de uma pessoa, não diariamente.

Assim, para 1980 e 1991, é possível a construção da variável:

```

* pos_ocup_hab = 0 sem remuneração
*      1 trabalhador agrícola volante
*      2 parceiro ou meeiro - empregado
*      3 parceiro ou meeiro - autônomo ou conta-própria
*      4 trabalhador doméstico - empregado
*      5 trabalhador doméstico - autônomo ou conta-própria
*      6 empregado
*      7 autônomo ou conta-própria
*      8 empregador

```

Em 1980 o quesito originalmente não distinguia “trabalhadores domésticos” dos demais empregados ou trabalhadores por conta-própria. Para obter a variável acima, identifica-se a categoria através do código da ocupação (naquele ano, tinha ‘ocup_hab’ igual a 805).

Em 1970, os “trabalhadores agrícolas volantes” não formaram uma categoria separada, sendo incluídos entre os “empregados”; além disso, os parceiros ou meeiros, desde que não fossem empregadores, formaram uma única categoria. Novamente, para identificar os trabalhadores domésticos, recorre-se ao código de ocupação (‘ocup_hab’ igual a 813). A variável abaixo vale de 1970 até 1991.

```
* pos_ocup_habB = 0 sem remuneração
*                1 parceiro ou meeiro
*                2 trabalhador doméstico - empregado
*                3 trabalhador doméstico - autônomo ou conta-própria
*                4 empregado
*                5 autônomo ou conta-própria
*                6 empregador
```

A contribuição para previdência é registrada para todas as posições na ocupação. Essa informação está contida na variável ‘previd’.

```
* previd_A = 0 - não
*           1 - sim, contribui p/ previdência
```

Desde 1980, investigou-se a remuneração e a quantidade de horas habitualmente trabalhadas na ocupação habitual (agregando todos os estabelecimentos) e em todas as ocupações. Em 1991, registrou-se o número de horas (com arredondamento para cima no caso de fração maior ou igual a 30 minutos), mas em 1980 foi adotada uma divisão por categorias, seguida no arquivo compatibilizado. Assim, temos, para a ocupação principal e para todas as ocupações, respectivamente:

```
* hrs_oc_hab, hrs_todas_oc = 1 - menos de 15 horas
*                          2 - de 15 a 29 horas
*                          3 - de 30 a 39 horas
*                          4 - de 40 a 48 horas
*                          5 - 49 horas ou mais
```

Em 1970, são pesquisadas as horas de trabalho, mas apenas para a ocupação habitual, e foram excluídos os trabalhadores do setor agropecuário ou extrativista (que na mesma pergunta, em vez de horas por semana, deveriam reportar meses trabalhados por ano). Como foram adotados intervalos distintos dos de 1980, são oferecidas duas variáveis: ‘hrs_oc_habB’, com todas as categorias de 1970 (disponível para 70 e 91, mas não para 80); e ‘hrs_oc_habC’ com agregados que podem ser obtidos nos três anos. Ambas as variáveis excluem os trabalhadores do setor agropecuário ou extrativista.

```
* hrs_oc_habB = 1 - menos de 15 horas
*              2 - de 15 a 39 horas
*              3 - de 40 a 49 horas
*              4 - 50 horas ou mais

* hrs_oc_habC = 1 - menos de 15 horas
*              2 - de 15 a 39 horas
```

*

3 - 40 horas ou mais

O rendimento, por sua vez, é calculado, de acordo com a posição na ocupação, da seguinte maneira:

- Para os “empregados”, computam-se os salários atualmente recebidos pelo exercício da ocupação (ou recebidos no último mês em que a exerceu), acrescido da média mensal de componente variável da remuneração (gorjeta, comissões, horas extras). São desconsiderados a participação nos lucros e o 13º salário;
- Para os “empregadores” e “conta-própria”, o rendimento consiste na média mensal no período de referência (ou na parte desse período em que indivíduo exerceu a ocupação) dos rendimentos auferidos, corrigidos para o mês da pesquisa.

Em 1980 e 1991, há distinção entre trabalho principal (ou habitual) e outros trabalhos (ou ocupações). As variáveis de rendimento do trabalho principal e demais trabalhos são registradas, respectivamente, em ‘rend_ocup_hab’ e ‘rend_outras_ocup’. Há também a separação da renda do trabalho de outras fontes de renda, como aluguéis e doações. Porém, as outras fontes de renda variam de ano para ano, não havendo, em dois Censos seguidos, um conjunto com as mesmas denominações. Por conta disso, optou-se por criar uma variável que indique a renda total do indivíduo, ‘rend_total’, que é a soma da renda do trabalho com a de todas as outras fontes. Em 1970, existe apenas a renda total, uma vez que não há distinção de fontes de renda, nem entre tipos de trabalho (principal ou secundário).

O Censo também pesquisa a condição de atividade dos indivíduos que não trabalham. Em 1980 e 1991 foram identificados, dentre os indivíduos de 10 anos de idade ou mais que não trabalhavam, aqueles que tomaram providência para encontrar trabalho nos dois meses anteriores à pesquisa. Estes, juntamente com os que trabalharam no período de referência, compuseram a população economicamente ativa (PEA).

Para os anos de 1980 e 1991, os indivíduos que não faziam parte da PEA foram classificados nas categorias 3 a 9 descritas na variável ‘cond_ativ’:

```
* cond_ativ = 0 trabalhou nos últimos 12 meses
*             1 procurando trabalho - já trabalhou
*             2 procurando trabalho - nunca trabalhou
*             3 aposentado ou pensionista
*             4 vive de renda
*             5 detento
*             6 estudante
*             7 doente ou inválido
*             8 afazeres domésticos
*             9 sem ocupação
```

Em 1970, há importantes distinções na forma de obtenção desses dados. Em primeiro lugar, não há clara distinção entre quem trabalhou e quem procura trabalho, embora os indivíduos que não tinham nem procuravam trabalho tenham sido classificados nas mesmas categorias que os indivíduos fora da PEA em 80 e 91. Uma classificação apenas para 1970 foi mantida para não haver perda de informação para este quesito:

```

* cond_ativB = 1 trabalhou/procurando trabalho (mas já trabalhou)
*             2 procurando trabalho - nunca trabalhou
*             3 aposentado ou pensionista
*             4 vive de renda
*             5 detento
*             6 estudante
*             7 doente ou inválido
*             8 afazeres domésticos
*             0 sem ocupação

```

Em segundo lugar, a classificação de indivíduos na categoria “procurando trabalho” dependeu apenas de declaração do indivíduo sobre o que fazia nos últimos doze meses, sem referência ao momento quando foi tomada a última providência para tanto. Além disso, a classificação das pessoas nas categorias “não-PEA” dependeu da atividade que elas consideravam principal, em vez da hierarquia adotada a partir de 1980. Para haver uma variável compatível entre 1970 e 1991, foi criada a variável ‘pea’, indicando a população economicamente ativa. Nesta variável, em 1970, considera-se valor igual a ‘1’ para quem ‘cond_ativB’ é igual a ‘1’ ou ‘2’; e para 1980 e 1991, considera-se valor igual a ‘1’ para quem ‘cond_ativ’ é igual ou menor que ‘3’.

```

* pea = 1 economicamente ativo
*      0 inativo

```

Apenas para 1970 e 1980, há um quesito acerca da atividade econômica exercida na última semana, o mesmo período de referência para toda a seção de trabalho e rendimento dos Censos 2000 e 2010. Para 1970 e 1980, tem-se a variável ‘trab_semana’:

```

* trab_semana = 1 só trabalho habitual
*              2 trabalho habitual e outros
*              3 só trabalho diferente do habitual
*              4 outros

```

Apesar da existência desta última variável, que, a princípio, poderia ser utilizada para compatibilizar a informação com 2000 e 2010 (que também investigam o trabalho na última semana antes da pesquisa), respondem a este quesito apenas aqueles que trabalharam ou procuraram trabalho nos últimos 12 meses. Nos anos mais recentes, essa informação é irrelevante para a caracterização da ocupação investigada. Assim, é possível que pessoas inativas no período de 12 meses estejam trabalhando na semana de referência, o que torna a comparação inviável. Nesse sentido, optou-se por não compatibilizar essas informações.

Atividade econômica na semana de referência (2000 e 2010)

A investigação da atividade econômica dos indivíduos foi modificada drasticamente no Censo 2000, quando se considerou o exercício de trabalho ou a atividade na semana de referência. Isto implicou avaliar a situação de curto prazo dos indivíduos, no lugar das condições prevalentes no horizonte de um ano, como vinha sendo feito. Assim, não se pode

comparar diretamente, em geral, as informações obtidas em 2000 e 2010 com as discutidas acima para os censos anteriores.

Todas as variáveis abaixo se referem somente aos Censos 2000 e 2010. Para esses anos, é possível identificar se o indivíduo exerceu atividade remunerada na semana de referência e se estava afastado de algum trabalho por motivo de saúde, férias, licença etc. Esses casos estão registrados nas variáveis abaixo:

```
* trab_rem_sem = 0 - não
*                 1 - sim, trabalhou em atividade remunerada

* afast_trab_sem = 0 - não
*                 1 - sim, estava afastado de trabalho remunerado
```

Dentre aqueles que não trabalharam nem estavam afastados, investigou-se a possibilidade do indivíduo ter ajudado algum morador do domicílio em sua atividade ou ter exercido atividade de cultivo, extração vegetal, criação de animais, etc., destinada à alimentação de moradores. Porém, há diferenças na obtenção dessas informações. Em 2000, há explícita referência ao trabalho não-remunerado como aprendiz ou estagiário em ajuda a algum morador. Além disso, não é especificado se a atividade do morador ajudado é remunerada ou não. Em 2010, há apenas uma referência genérica de ajuda em atividade remunerada de outro morador. Apesar disso, essas divergências não devem impactar de forma substancial o conceito de trabalho não-remunerado, registrado em duas variáveis:

```
* nao_remun = 0 - não
*             1 - sim, ajudou sem remuneração

* trab_proprio_cons = 0 - não
*                   1 - sim, exerceu atividades de cultivo para
alimentação dos moradores
```

Há também informação sobre o número de trabalhos que a pessoa tinha na semana de referência:

```
* mais_de_um_trab = 0 - não, tinha somente um trabalho
*                 1 - sim, tinha mais de um trabalho
```

Os códigos da ocupação principal são registrados nas variáveis 'ocup2000' e 'ocup2010'. Em 2010, existem as duas variáveis. Isso ocorre por conta de uma nova classificação utilizada pelo IBGE em 2010. Para haver comparabilidade, o IBGE manteve as duas classificações. Para 2000, as ocupações também são classificadas de acordo com os códigos de 1991. Isso foi mantido a variável 'ocup1991' para o caso de uma eventual tentativa de compatibilização, lembrando que o horizonte temporal em 2000 difere daquele observado em 1991.

A posição na ocupação no trabalho principal da semana foi investigada de forma diferente nos anos 2000 e 2010. Além da questão do trabalho não-remunerado mencionada anteriormente, os Censos diferem no quesito empregado doméstico. Enquanto em 2000 o indivíduo tem a

opção de escolher essa categoria, com ou sem carteira de trabalho, em 2010 não há referência alguma no questionário aplicado, sendo os trabalhadores domésticos inseridos na posição 'empregado' com ou sem carteira. Foi possível construir a variável abaixo utilizando variáveis derivadas, que não constam no questionário de 2010 (V6930 e V6940):

```
* pos_ocup_sem = 1 - Empregado com carteira
*                2 - Militar e Funcionário Públicos
*                3 - Empregado sem carteira
*                4 - Trabalhador doméstico com carteira
*                5 - Trabalhador doméstico sem carteira
*                6 - Conta - própria
*                7 - Empregador
*                8 - Não remunerado
*                9 - Trabalhador na produção para o próprio consumo
```

Para os empregadores, pode-se obter o número de empregados no empreendimento, após recodificação das alternativas em 2000 para adequar-se a 2010:

```
* qtos_empregados = 1 - Um a cinco empregados
*                  2 - Seis ou mais
```

A contribuição para previdência é registrada para os empregados sem carteira de trabalho assinada (exclusive servidores públicos e militares), para trabalhadores por conta-própria e para empregadores. Porém, em 2000, pergunta-se apenas sobre o trabalho principal, enquanto em 2010, há opção de responder se contribui em outros trabalhos. Essa última informação foi ignorada para construir a variável 'previd_B'.

```
* previd_B = 0 - não
*            1 - sim, contribui p/ previdência no trab. principal
```

O número de horas trabalhadas no trabalho principal da semana é registrado em 'horas_trabprin'. Em 2000, a informação a respeito do número de horas trabalhadas nos demais trabalhos foi coletada e mantida na variável 'horas_outros_sem', para alguma eventualidade, apesar dos problemas de compatibilização com 1980 e 1991.

Para aqueles que não trabalharam, de forma remunerada ou não, pergunta-se se tomou alguma providência para conseguir trabalho no mês de referência. Isso é registrado em:

```
* tomou_prov = 0 - não
*              1 - sim, tomou providência para conseguir trabalho
```

O rendimento foi investigado para o trabalho principal 'rend_ocup_prin' e os demais trabalhos 'rend_outras_ocup', sendo computada a remuneração bruta ou a retirada do empreendimento (rendimento bruto menos gastos realizados), e excluindo benefícios, como auxílio moradia e transporte. Caso o rendimento seja em produtos ou mercadorias, consta o

valor real ou estimado; se sazonal, consta a média mensal para o mês de referência.¹ Além das variáveis 'contínuas', tanto em 2000 quanto em 2010 há variáveis indicando o rendimento em salários mínimos: 'rend_prin_sm' e 'rend_outras_sm'.

Os rendimentos não oriundos do trabalho são registrados em 'rend_outras_fontes'. Da mesma forma que os Censos anteriores, os componentes desta variável não são uniformes entre 2000 e 2010, por isso a opção por incluir todos em uma mesma variável. Por exemplo, em 2000, há uma variável computando valores recebidos de pensão alimentícia, mesada e doação, ao passo que em 2010, todas essas fontes entram na variável de outros programas sociais ou transferências.

Ainda sobre rendimentos, tem-se a renda total 'rend_total' e a renda total em salários mínimos 'rend_total_sm', que nada mais são do que a soma da renda do trabalho com a de outras fontes.

Por fim, para todos os anos, é possível obter a renda familiar, que inclui a renda de todos os membros da família, com exceção dos empregados domésticos, parentes do empregado doméstico e pensionistas. Para 1991, essa variável já existe no arquivo original. Para os demais censos, é necessário construí-la. A variável compatibilizada, que sofre de todos os problemas discutidos anteriormente, é registrada em 'rend_fam'.

Todos os rendimentos são deflacionados para julho/2010, mas as variáveis com os rendimentos nominais são mantidas na base de dados. As variáveis com os rendimentos deflacionados possuem o mesmo nome que as variáveis nominais, acrescidas do sufixo 'def'.

Município em que trabalha

Exceto em 1991, investiga-se o município onde o indivíduo trabalha, mas apenas em 2010 essa pergunta refere-se explicitamente a trabalho; nos demais censos, a pergunta refere-se a trabalho ou estudo, ou seja, não é possível identificar se a resposta do indivíduo refere-se ao local de trabalho ou de estudo se o mesmo estuda e trabalha. Optou-se pela suposição de que o indivíduo respondeu o local de trabalho nesse caso de dupla jornada. Foi construída uma variável dummy para este quesito, válida somente para quem trabalha (independentemente da frequência escolar):

```
* mun_trab = 0 - não, trabalha em município diferente do qual reside
*           1 - sim, trabalha no mesmo município em que reside
```

Nota sobre Rendimentos

Em todos os Censos desde 1970 foi feita alguma medida de rendimento mensal. Essa medida exclui recursos obtidos esporadicamente (doações não regulares, heranças, prêmios de loterias). Entre 1970 e 1991, para fins de cálculo, os componentes da renda podiam receber

¹ Essas orientações estão contidas na documentação do Censo 2000, mas não aparecem explicitamente nas descrições das variáveis em 2010; por outro lado, não há indicações de que tenham sido alteradas substancialmente.

dois tratamentos diferentes, conforme considerados “fixos” ou “variáveis”. Os primeiros foram computados de acordo o valor referente ao mês da coleta do Censo; os outros, pela média dos doze meses anteriores. Em 2000 e 2010, por outro lado, todos os rendimentos foram medidos conforme seu valor no mês de referência. O total dos rendimentos do indivíduo, disponível em qualquer ano, é registrado na variável ‘rend_total’.

Enquanto em 1970 foi obtida apenas a renda total de cada indivíduo, nos anos seguintes esta foi desagregada de acordo com a origem dos rendimentos. Assim, além do rendimento do trabalho – discutido acima nesta seção –, são discriminados:

- desde 1980, o rendimento de “aposentadoria e pensão”. Este conceito, no entanto, variou ao longo do tempo. Em 1991, consideraram-se estritamente valores pagos a título de aposentadoria e pensão (mas não o abono permanência) por instituto de previdência oficial; em 2000 incluiu-se a previdência privada fechada; e em 1980 foram considerados diversos outros tipos de recebimento, como o abono permanência e o 14º salário (abono salarial) recebido do PIS/PASEP.
- em 1980 e 2000: “aluguel e arrendamento” e “doação, mesada e pensão alimentícia” (transferências recebidas de não moradores).

D.11. Fecundidade

Para mulheres de 15 anos de idade ou mais e, a partir de 1991, para aquelas de 10 anos ou mais, foram registrados os números de filhos tidos, filhos nascidos vivos, filhos nascidos mortos e filhos ainda vivos. Nos arquivos compatibilizados, estas quantidades compõem, respectivamente, as variáveis ‘filhos_tot’, ‘filhos_nasc_vivos’, ‘filhos_nasc_mortos’ e ‘filhos_vivos’.

De 1980 em diante, esses valores foram computados separadamente de acordo com o sexo dos filhos. A partir dessa informação foram construídas variáveis a seguir, onde o sufixo ‘hom’ refere-se a homens e ‘mul’ refere-se a mulheres:

```
filhos_hom,filhos_mul = número de filhos tidos  
f_nasc_vivos_hom, f_nasc_vivos_mul = filhos nascidos vivos  
f_nasc_mortos_hom, f_nasc_mortos_mul = filhos nascidos mortos  
filhos_vivos_hom,filhos_vivos_mul = filhos ainda vivos
```

Desde 1980, adicionalmente, foi possível obter a idade do último filho nascido vivo, e a partir de 1991 o sexo do mesmo. Esses dados são registrados nas variáveis ‘idade_ult_nasc_v’, valor em anos completos, e ‘sexo_ult_nasc_v’, descrita a seguir.

```
* sexo_ult_nasc_v = 0 - feminino  
*                  1 - masculino
```

Os “filhos tidos” são aqueles cuja gestação alcançou pelo menos sete meses. Estes são classificados como “nascidos vivos” (ou “nascidos mortos”) conforme a presença de algum (ou

ausência de qualquer) sinal de vida – como respiração, batimento cardíaco, pulsação do cordão umbilical etc.– depois de separados do corpo da mãe.

Por fim, vale destacar que, para compatibilizar todos os anos, é necessário restringir a idade nos Censos 1991, 2000 e 2010 para 15 anos ou mais de idade, já que nesses anos, todas as mulheres com 10 anos ou mais foram entrevistadas, o que não ocorreu em 1970 e 1980.

D.12. Peso da pessoa

A variável 'peso_pess' registra o peso amostral da pessoa. Esta variável foi apenas renomeada para ter o mesmo nome em todos os censos.

D.13. Número de moradores da família

As variáveis 'n_pes_fam', 'n_homem_fam' e 'n_mulher_fam' indicam o número total de moradores da família, o número de moradores do sexo masculino e o número de moradores do sexo feminino, respectivamente. Essas variáveis não estão presentes no arquivo de domicílios compatibilizado, uma vez que pode haver mais de uma família por domicílio.