



CENSO DEMOGRÁFICO
DE 1991
- ASPECTOS METODOLÓGICOS -

Novembro 1992

CENSO DEMOGRÁFICO DE 1991 - ASPECTOS METODOLÓGICOS

INTRODUÇÃO

O Censo é prioridade número um do IBGE, que agrega à preocupação com a qualidade dos dados, a consciência da necessidade de uma divulgação rápida e abrangente dos resultados, nas mais diversas formas (publicações, disquetes, arquivos magnéticos etc.) visando preencher a lacuna de informações provocada pelo seu adiamento.

ÂMBITO EM EXTENSÃO E PROFUNDIDADE

Por âmbito em extensão entende-se os objetos que serão alvos da investigação. No Censo Demográfico de 1991 eles se constituíram dos domicílios, das famílias e dos indivíduos.

O âmbito em profundidade compreende os temas e os quesitos que serão levantados em relação a cada um desses objetos. Ele deve obedecer a critérios tais como: atender as necessidades nacionais de informações bem como as recomendações internacionais; respeitar, sempre que possível, a comparabilidade histórica, não desprezando a introdução de novos aperfeiçoamentos. Os temas e os quesitos que fazem parte de um censo também estão sujeitos às restrições de ordem técnica e de ordem financeira.

O âmbito em profundidade do Censo Demográfico de 1991 foi definido a partir de consultas e debates com usuários das informações censitárias, representativos dos diversos segmentos sociais. A partir das sugestões foram construídos questionários a serem testados em Censo Experimental.

Os Questionários Básico-CD.1.01 e da Amostra-CD 1.02, testados em Limeira, sofreram algumas alterações decorrentes das avaliações feitas. Além dessas modificações, por solicitação direta da Presidência da República, foi incluído um quesito específico para investigação das pessoas portadoras de deficiências físicas (visual, auditiva e motora) ou mentais. Posteriormente, a Lei nº 7853, de 24 de outubro de 1989, que dispõe sobre o apoio às pessoas portadoras de deficiência, estabeleceu que "deverão ser incluídas no Censo Demográfico de 1990, e nos subseqüentes, questões concernentes à problemática da pessoa deficiente, objetivando o conhecimento atualizado do número de pessoas portadoras de deficiência no País".

Para atender à demanda de informações sobre o mercado informal, foram introduzidos dois novos quesitos no bloco referente

às características econômicas e feita uma adequação do quesito que investiga a posição na ocupação. Neste sentido, também foi realizado, por ocasião da coleta, em instrumento próprio (CD 1.08-Folha de Cadastramento para os Censos Econômicos), um arrolamento dos domicílios com exploração de atividade econômica, visando pesquisas posteriores.

Os quesitos novos se referem a "quantas pessoas trabalham no estabelecimento, negócio, instituição etc. onde exerce a ocupação principal" e ao "local de trabalho onde exerceu a ocupação", discriminando, se no domicílio, na via pública, em propriedade agropecuária, em empresa ou firma, em casa do cliente ou patrão ou em outro local.

A adequação do quesito posição na ocupação permite identificar os empregados do setor privado, do setor público, discriminando se servidor público ou de empresa estatal, e os trabalhadores domésticos, discriminando empregados e conta própria.

ASPECTOS DE AMOSTRAGEM - INOVAÇÕES NO CENSO 91

Com relação aos aspectos de amostragem, foram introduzidas duas grandes modificações no Censo Demográfico de 1991, relativamente aos Censos Demográficos anteriores. Essas modificações referem-se à fração amostral e ao método de expansão da amostra.

Em 1960, 1970 e 1980, a fração amostral era de 25% dos domicílios para todos os municípios brasileiros. Em 1991, foram adotadas duas frações amostrais distintas, dependendo do tamanho do município, medido em termos de população estimada. A fração amostral adotada foi de 20% para os municípios com população até 15.000 habitantes e 10% para os municípios com mais de 15.000 habitantes.

Em 1960 e 1970, as estimativas a partir da amostra de domicílios e pessoas foram obtidas por meio do estimador simples de pós-estratificação. Em 1980, para a obtenção das estimativas foi utilizado o Processo Iterativo de Estimação por Totais Marginais - PIETOM, conhecido na literatura internacional como *Raking Ratio Estimation Procedure*. Para o Censo Demográfico de 1991, pretende-se utilizar o estimador associado ao método *Two Step Generalized Least Squares Estimation Procedure - GLSEP*, desenvolvido por M. Bankier do Statistics Canada e que foi aplicado para a obtenção das estimativas associadas à amostra do Censo Canadense de População de 1991.

A metodologia associada ao GLSEP, que no IBGE está sendo denominado MQG2 - Mínimos Quadrados Generalizados em duas Etapas, supõe o ajuste do modelo em dois momentos a fim de buscar consistência das estimativas de algumas variáveis para as quais são conhecidos os valores de universo, em dois níveis geográficos distintos, um deles sendo o setor censitário, ou seja, a área de coleta,

e o outro a tradicional área de ponderação. No Censo de 1980 cada município constituiu uma área de ponderação, com exceção de alguns municípios grandes, para os quais a área de ponderação foi definida segundo alguma subdivisão administrativa do município. Para o Censo de 1991, ainda não foram definidas as áreas de ponderação, mas em função de método de estimação proposto já se sabe que terão que ser definidas áreas de ponderação de tamanho, medido em número de domicílios, bem menores do que um município.

ETAPAS DO CENSO DEMOGRÁFICO

A execução do Censo Demográfico pode ser desagregada em quatro grandes etapas, que devem ser precedidas por planejamento. São elas:

- coleta;
- apuração;
- análise;
- divulgação.

1. COLETA DAS INFORMAÇÕES

1.1 - Adiamento do Censo Demográfico de 1990

A legislação (Decreto nº 96.705, de 15 de setembro de 1988) fixava 1º de setembro de 1990 a data de referência do Censo Demográfico, quando teria início a coleta das informações. Para o desenvolvimento desta atividade, o IBGE deveria contar com 22 mil supervisores e 166 mil recenseadores. Por questões de natureza administrativa externas ao IBGE, somente na primeira semana de agosto foi dada autorização para seleção desse pessoal.

Isto implicaria na transferência do início da coleta para a segunda quinzena do mês de outubro, o que tecnicamente não era recomendável pelo afastamento da data de referência e por prolongar a coleta nos meses de dezembro e janeiro.

Nessas circunstâncias, fundamentada em criteriosa avaliação feita pelas áreas técnicas e operacionais do IBGE, foi transferida a data de referência para 1/9/1991.

Como conseqüências do adiamento podem ser citadas entre outras:

- postergar o conhecimento do volume e distribuição espacial da população;
- postergar o conhecimento de informações sócio-econômicas e demográficas ao nível nacional e estadual e especialmente ao nível municipal;
- introduzir alguma dificuldade na análise das informações por não respeitar a decenalidade dos intervalos censitários.

Entretanto, o adiamento do censo se apresentou como medida necessária para resguardar a qualidade e confiabilidade do levantamento.

1.2 - Período da Coleta

A coleta teve início em 1º de setembro e tinha duração prevista para um período de três meses.

No final do mês de novembro, quando começou a greve dos servidores do IBGE, era a seguinte a posição da coleta:

Total de setores.....	163.196
Setores com coleta encerrada.....	153.520
Setores com coleta em andamento.....	9.287
Setores com coleta não iniciada.....	389

A coleta prosseguiu durante o mês de dezembro, tendo em vista o compromisso do IBGE informar ao Tribunal de Contas da União a população dos municípios brasileiros até o final do ano.

Ao chegar a data prevista para a prestação dessa informação, 100 dos 4491 municípios instalados ainda não tinham sua coleta concluída. Esses municípios estavam localizados nos Estados de Pernambuco, Bahia, Rio de Janeiro e São Paulo.

Para estes municípios, a coleta se prolongou pelo mês de janeiro de 1992, tendo o IBGE divulgado os Resultados Preliminares do Censo, baseado nos elementos de controle da coleta, em 5 de fevereiro de 1992.

1.3 - Verificação da Coleta

Além das atividades de supervisão sistemática dos trabalhos dos recenseadores previstas no planejamento do Censo, para os municípios de Recife, Salvador, Rio de Janeiro e São Paulo, onde as consequências da greve se fizeram sentir mais diretamente sobre a coleta, foi realizado um trabalho especial de verificação que constou das seguintes etapas:

- comparação da população e do número de domicílios ocupados e não ocupados referentes aos setores censitários com áreas comparáveis nos Censos de 1980 e 1991;
- identificação daqueles setores cuja diferenças eram significativas. Inicialmente se trabalhou com diferenças iguais ou superiores a 10% e posteriormente se reduziu esse valor.
- verificação dos setores selecionados por equipe especial e correção das informações, quando necessária.

Este trabalho de verificação proporcionou resultados positivos, principalmente em alguns bairros da zona Sul do Rio de Janeiro, uma vez que várias falhas de cobertura puderam ser corrigidas.

2. APURAÇÃO DAS INFORMAÇÕES

2.1 - Descentralização da Apuração

A apuração dos Censos Demográficos, realizados anteriormente no Brasil, era executada de forma totalmente centralizada, desde a transcrição até a tabulação final dos dados.

A centralização dos processos de apuração gerava a necessidade de contratação (temporária) de equipes numerosas, manuseando um grande volume de documentos. Como decorrência, diversos níveis de "supervisores" eram necessários para gerenciar as várias equipes de preparo, transcrição, conferência, codificação, etc.

Este gigantismo nas equipes centralizadas de apuração acarretava problemas sérios na administração de produção e gerava diversos transtornos na época de liberação dos recursos, ao término do Censo.

Outra característica típica da produção centralizada é a de que os processos voltados a grandes volumes quantitativos exigem controles operacionais sofisticados e quase não permitem paralelismo entre eles. A consequência da implantação de con-

troles rígidos (necessários, neste contexto) tornam a linha de produção inflexível e o resultado da linearidade adotada na execução dos processos estendem os prazos de apresentação dos resultados.

Com os objetivos de encurtar os prazos de produção das informações, evitando concentração excessiva de material e pessoal, e de melhorar a qualidade dos resultados finais, o Censo Demográfico de 1991 está tendo sua apuração descentralizada nas diversas Unidades Regionais do IBGE.

O modelo de descentralização adotado foi testado no Censo Experimental de Limeira. Os resultados obtidos foram positivos e, se, por um lado, confirmaram a viabilidade da proposta, por outro lado subsidiaram o aperfeiçoamento do modelo na implementação do sistema definitivo.

O sistema de descentralização considera a complexidade e o volume do trabalho de apuração dos Questionários Básico e da Amostra. O Questionário Básico, de estrutura mais simples, que foi aplicado à totalidade das unidades domiciliares, teve as etapas de empastamento, digitação, execução das críticas básica e de quantidade realizadas em centros de apuração localizados em todas as Unidades Regionais, exceto Rondônia, Acre, Amazonas, Roraima, Amapá e Tocantins. A seguir, os dados foram transmitidos para o computador central no Rio de Janeiro, para consolidação.

O Questionário da Amostra, pela sua extensão e multiplicidade de temas abordados, é composto por maior número de quesitos, alguns com resposta aberta, requerendo tratamento diferenciado, seja pela especialização das equipes envolvidas, como pela necessidade de recursos computacionais mais sofisticados. Dessa forma, a sua apuração será descentralizada em menor número de centros, com implantação prevista nas seguintes Unidades: Pará, Ceará, Pernambuco, Bahia, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Rio Grande do Sul e Goiás. Nesses centros, além de serem atendidas as fases de digitação e execução das críticas básica e de quantidade, serão desenvolvidas as tarefas de codificação dos quesitos abertos, assistida por computador, e execução de crítica de consistência, que irão preceder a transmissão de dados para o CPD no Rio de Janeiro.

2.2 - Codificação assistida por computador

Outro grande ponto a ser destacado para o Censo Demográfico é a substituição de toda a etapa de codificação manual por processo assistido por computador, também executado a nível regional, trazendo grandes vantagens na qualidade, velocidade e, principalmente, na uniformização de procedimentos.

Esta atividade é desenvolvida a partir dos seguintes procedimentos.

Os quesitos em aberto no questionário (religião, migração, último grau concluído, ocupação e atividade) tiveram o seu conteúdo (texto) transcrito na etapa de digitação.

A partir da relação de códigos e descrição dos códigos disponíveis e utilizados anteriormente, pela codificação manual, criou-se o Banco de Códigos que contém código, descrições para este código e palavras constantes destas descrições transformadas, segundo regra fonética aplicada.

Para a pesquisa do código, foi utilizada a pesquisa parcial de texto, que consiste em comparar, no lugar de todo o texto, as palavras de que é composto. Portanto, os textos constantes das descrições foram divididos em palavras e no momento da codificação o texto digitado oriundo do questionário também é dividido em palavras e estas transformadas foneticamente para a pesquisa. A transformação fonética das palavras, permite que, mesmo com diferentes grafias, a recuperação do código seja bem sucedida.

2.3 - Correção Automática das Informações

A metodologia de crítica e correção de dados qualitativos adotada para o Censo de 1991 é de autoria de **Fellegi e Holt** implementada no Sistema DIA - *Detección e Imputación Automática de errores para datos cualitativos*.

Os princípios básicos, nos quais se baseia a metodologia, são:

- 1) os dados de cada questionário devem satisfazer todas as regras de críticas, modificando o menor número possível de variáveis, procurando, assim, manter a maior quantidade possível de informação original;
- 2) não se faz necessário especificar procedimentos de correção, uma vez que estes devem ser deduzidos automaticamente das regras de críticas impostas aos dados;
- 3) a tarefa de correção (imputação) procura resguardar, na medida do possível, as distribuições de freqüências dos dados bons, quer em termos das distribuições marginais como da distribuição conjunta, reveladas através da distribuição de freqüências dos registros bons (aqueles que passaram pelas críticas sem serem rejeitados).

O Sistema DIA (Sistema de Deteccion e Imputación Automática de errores para datos cualitativos) trata-se de um pacote

computacional desenvolvido por técnicos do Instituto Nacional de Estatística (INE - Espanha), baseado na metodologia de **Fellegi e Holt**, que tem o objetivo de facilitar a depuração de Censos e grandes pesquisas estatísticas.

Este sistema permite não só a utilização de correção probabilística, como também a de correção determinística para ser usada no tratamento de erros sistemáticos, e ainda garante uma consistência entre as regras de incompatibilidade e as de imputação determinística.

A viabilidade do uso desta nova metodologia foi verificada através de teste realizado com os dados coletados no Censo Experimental, mais precisamente do CD 1.01-Questionário Básico. Os resultados indicaram que a adoção desta nova metodologia deverá trazer ganhos significativos de qualidade e de segurança na correção automática das informações, bem como economia de tempo e de recursos no desenvolvimento desta atividade.

2.4 - Posição dos Trabalhos de Apuração

2.4.1 - Resultados Preliminares

Os resultados já divulgados do Censo, no mês de fevereiro, foram obtidos a partir de elementos de controle da coleta e se referem a população total e por sexo dos municípios.

2.4.2 - Apuração do CD 1.10-Boletim Resumo

Permite conhecer as seguintes informações:

- população por sexo para desagregações geográficas menores que o município (distritos, bairros, favelas) e número de domicílios ocupados, vagos, fechados e de uso ocasional, segundo a situação urbano/rural;
- população e domicílios por setor censitário.

Já está disponível arquivo magnético para exploração de dados, através de procedimentos informatizados para emissão de relatórios ou gravação em meio magnético (disquete ou fita), permite a seleção de itens geográficos e/ou agregações das informações.

A partir desse arquivo estão sendo geradas tabulações para ampla divulgação.

2.4.3 - Apuração do CD 1.01-Questionário Básico

2.4.3.1 - Situação e perspectiva da apuração descentralizada

Pelo andamento dos trabalhos espera-se que somente o Rio de Janeiro conclua a digitação no mês de janeiro, devendo as demais unidades estar com essa atividade terminada ainda este ano.

2.4.3.2 - Situação e perspectiva da apuração centralizada

As atividades centralizadas subseqüentes são: crítica e correção automática das informações; tabulação; e análise para liberação dos resultados.

A Crítica e correção automática encontram-se em fase de teste para implantação na produção. Também estão sendo selecionados os relatórios que serão utilizados na avaliação do processo.

O plano tabular está definido e encontra-se em fase de programação. A previsão é de que se tenha até dezembro as tabulações relativas às UFs que tenham suas informações consolidadas e liberadas pela correção automática. No momento, já se encontram consolidadas as informações das seguintes unidades: AC, AP, TO, PI, RN, AL, SE, ES, MS, GO e DF.

2.4.4 - Apuração do CD 1.02-Questionário da Amostra

Encontra-se em fase de planejamento da implantação da apuração descentralizada que tem início com a transcrição (digitação) dos questionários nos dez pólos de apuração.

Em relação ao plano tabular, o IBGE preparou um projeto, constante de 144 tabelas, que foi discutido e apreciado por um Grupo de Trabalho especialmente designado pela ABEP que fez sugestões e o reencaminhou ao IBGE para uma avaliação da Diretoria de Pesquisas e posterior encaminhamento para programação.

3. ANÁLISE DAS INFORMAÇÕES

A análise para liberação dos resultados está planejada em três fases: fundamentalmente produto das análises da Sinopse; análise da pré-crítica e da correção automática; e análise dos resultados, além de análises demográficas complementares.

A análise dos resultados da apuração do CD 1.10-Boletim Resumo consta de:

- comparação com resultados preliminares e com as estimativas de população;
- comparação com dados retrospectivos;
- análises demográficas;
- análises a nível municipal e setorial (identificação de unidades com problemas ainda recuperáveis).

4. DIVULGAÇÃO

A próxima publicação contendo resultados censitários será a Sinopse Preliminar do Censo Demográfico 1991, que contemplará um volume para cada Unidade da Federação e um volume para o total Brasil.

As publicações serão compostas por um conjunto de tabelas e de textos, incluindo uma análise das Mesorregiões para cada Unidade da Federação. As tabelas começaram a ficar disponíveis na primeira quinzena de outubro/92. Para obtenção do produto completo é demandado um tempo maior (elaboração das análises, editoração e impressão) devendo estar concluído na primeira quinzena de novembro.

5. AVALIAÇÃO DOS RESULTADOS

A ocorrência de erros é inerente a qualquer processo de medição. Por conseguinte, diversos tipos de erros são naturalmente presumíveis em um levantamento censitário, operação de grande complexidade, envolvendo o trabalho de milhares de pessoas desde seu planejamento até a divulgação dos resultados.

Assim, faz parte da operação censitária um programa de avaliação dos seus resultados.

Para o Censo Demográfico de 1991, o IBGE planejou e realizou uma Pesquisa de Avaliação da precisão da cobertura censitária, pois a omissão de domicílios e de pessoas são os erros mais graves de um censo.

A Pesquisa de Avaliação compreende um levantamento, em uma amostra de setores, com coleta independente do Censo, de forma a captar e medir as tendenciosidades ocorridas.

A apuração da Pesquisa de Avaliação está sendo desenvolvida descentralizadamente nas Divisões de Pesquisas do IBGE nos estados, se apoia na comparação (batimento) das Folhas de Coleta e Questionários do Censo e irá permitir o cálculo de taxas de omissão de domicílios e de pessoas para as diversas Unidades da Federação, exceto as da Região Norte, que serão tratadas de maneira agregada.

O cronograma de trabalho prevê que, a divulgação dos resultados do Questionário Básico seja acompanhada desses indicadores.

Também fazem parte da estratégia global de avaliação da cobertura censitária os seguintes procedimentos:

- comparação com os resultados da Operação Listagem da PME de março de 1992, para os setores compatíveis nas Regiões Metropolitanas;
- utilização da Listagem da PNAD-92, iniciada em junho, para detectar a omissão de domicílios.