
Notas técnicas

Informações gerais

No Cadastro Central de Empresas - CEMPRE, estão armazenados dados cadastrais e econômicos de todas as pessoas jurídicas formalmente constituídas no Território Nacional, independentemente da atividade exercida ou da natureza jurídica.

Os principais dados cadastrais disponíveis em nível das organizações como um todo são: razão social, natureza jurídica, classificação da atividade principal e ano de fundação; e no nível de unidade local são: nome fantasia, endereço completo, classificação da atividade e ano de fundação. Quanto aos dados econômicos, estão armazenadas informações de pessoal ocupado, de salários pagos e de receitas, obtidas em diversas fontes, que permitem também a identificação do porte das unidades.

A atualização do CEMPRE é feita anualmente, conjugando as informações do Cadastro de Empregadores da Relação Anual de Informações Sociais - RAIS, inscritos no Cadastro Nacional da Pessoa Jurídica - CNPJ, com aquelas obtidas a partir das pesquisas anuais nas áreas de Indústria, Comércio, Construção Civil e Serviços, realizadas pelo IBGE. As informações coletadas pelo IBGE prevalecem sobre as demais.

As Pesquisas Anuais de Indústria, de Comércio, de Construção e de Serviços, realizadas pelo IBGE, são pesquisas amostrais nas quais o desenho amostral distingue dois estratos. As empresas com porte acima de determinado corte, segundo o número total de pessoas ocupadas na empresa (20 no Comércio e nos Serviços, e 30 na Indústria e na Construção Civil) compõem o estrato certo, sendo, assim, investigadas censitariamente. As empresas abaixo desses cortes são selecionadas aleatoriamente, com base em critérios de amostragem probabilística.

Atualmente, o CEMPRE é composto de registros de aproximadamente 10,6 milhões de endereços de atuação de cerca de 9,6 milhões de organizações¹ formalmente constituídas, inscritas no Cadastro Nacional da Pessoa Jurídica - CNPJ, contemplando todo o conjunto de atividades econômicas. Cerca de 90% do CEMPRE é formado por entidades empresariais, sendo os demais registros distribuídos entre órgãos da administração pública e entidades sem fins lucrativos, que também possuem inscrição no CNPJ.

Para fins da divulgação das Estatísticas do Cadastro Central de Empresas referentes a um determinado ano, são selecionadas as unidades ativas no ano em questão, segundo critérios que estão descritos a seguir.

Critério de seleção das unidades ativas

O CEMPRE utiliza duas fontes básicas para sua alimentação: as pesquisas anuais do IBGE nas atividades de Indústria, Comércio, Serviços e Construção Civil, e a RAIS. Os dados oriundos das pesquisas do IBGE prevalecem sobre os da RAIS, no caso de informantes comuns às duas fontes.

O critério para selecionar as unidades ativas que fazem parte do âmbito das Estatísticas do Cadastro Central de Empresas considera as seguintes situações:

- Empresas ou outras organizações informantes de pesquisa do IBGE ou declarantes da RAIS, inclusive da RAIS negativa no ano em questão, com situação normal de funcionamento; e
- Empresas ou outras organizações que apresentarem situação cadastral indicando encerramento de atividades até 31.12, seja via RAIS, ou via pesquisas do IBGE, caso tenham declarado alguma informação econômica (pessoal ocupado ou salários pagos) ao longo do ano.

Critério para atribuição de valores econômicos de pessoal ocupado e salários

Quando uma mesma empresa é informante tanto do IBGE quanto da RAIS, os valores econômicos de pessoal ocupado e salários, relativos à empresa como um todo, declarados à pesquisa do IBGE prevalecem sobre os da RAIS. No entanto, para as unidades locais, o mesmo procedimento não era adotado até 2000, visto que a unidade básica de investigação das pesquisas do IBGE é a empresa e não a unidade local (exceto no caso da Pesquisa Industrial Anual, na qual, para algumas grandes empresas, são também obtidas informações para seus endereços de atuação). Nesse caso, vinha-se adotando apenas a RAIS como fonte básica de informações econômicas para as unidades locais.

A partir de 2001, com o objetivo de tornar essas informações compatíveis com as das empresas investigadas pelas pesquisas do IBGE, implementou-se um procedimento de ajuste nos valores das unidades locais. Tal ajuste consiste em distribuir proporcionalmente os valores de pessoal ocupado assalariado e de salários pagos das empresas, informado nas pesquisas institucionais, entre suas unidades locais, obedecendo à distribuição dessas informações na RAIS. No caso de empresa com uma

¹ Por organizações entende-se pessoas jurídicas, inscritas no Cadastro Nacional da Pessoa Jurídica - CNPJ, compreendendo entidades empresariais, órgãos da administração pública e entidades sem fins lucrativos.

única unidade local, a atribuição do valor da empresa é imediata. Com este procedimento, reduz-se a diferença até então observada entre os totais de unidades locais e de empresas, em função de estarem sendo computados a partir de fontes distintas.

Procedimentos de crítica e qualidade

Conforme já mencionado, o CEMPRE utiliza duas fontes básicas para sua alimentação: as pesquisas anuais do IBGE e a RAIS.

A apropriação dos dados por ambas as fontes não é direta, estando sujeita a diversos procedimentos de verificação de modo a garantir sua qualidade, dentre os quais se destacariam:

- Ampla verificação do código de atividade econômica das empresas, outras organizações e unidades locais, mediante verificação de palavras-chave no conteúdo da razão social, em especial, nas informações oriundas da RAIS. Para as empresas ou unidades locais informantes de pesquisas do IBGE, o código CNAE 1.0 é o mesmo atribuído por essas pesquisas, sempre prevalecendo sobre o código declarado na RAIS;
- Validação das informações econômicas de pessoal ocupado e salários, mediante críticas de dados agregados e de microdados, tomando por base informações de outras fontes, resultados de anos anteriores, bem como a verificação do *ranking* das maiores unidades de cada atividade para identificar eventuais erros de magnitude;
- Verificação das principais mudanças de atividade econômica, em relação ao ano anterior, para Unidades da Federação e município, bem como grandes variações nos valores de pessoal ocupado e de salários, em relação ao ano anterior; e
- Identificação e confirmação da ausência de grandes unidades que faziam parte do universo do ano anterior e que deixaram de integrar as estatísticas do presente ano. Atenção especial é dada aos órgãos da administração pública (ministérios, governos estaduais, prefeituras, etc.) que eventualmente ficam omissos com relação à declaração da RAIS e que, por esse critério, não integrariam o universo de referência. Para evitar que suas informações deixem de ser contempladas, afetando os resultados, pois normalmente empregam uma quantidade significativa de pessoas, seus dados são imputados com base nos valores do ano anterior, uma vez que sua existência pode ser averiguada.

Esses procedimentos refletem o amadurecimento dos trabalhos de compatibilização entre informações provenientes de registros administrativos e as produzidas pelas pesquisas do IBGE, partes constitutivas do CEMPRE.

Âmbito

Tendo em vista que o CEMPRE engloba registros de pessoas jurídicas inscritas no CNPJ, independentemente da atividade exercida ou da natureza jurídica, os dados desta publicação referem-se tanto a entidades empresariais, como a órgãos da administração pública e outras entidades sem fins lucrativos, aqui representadas pela denominação *organização*.

Nesta publicação, são tabulados dados de 5,7 milhões de empresas e outras organizações e de 6,1 milhões de unidades locais do CEMPRE consideradas ativas em 2005.

Estas informações resultam da consolidação de dados cadastrais e econômicos de fontes administrativas com os das pesquisas econômicas realizadas pelo IBGE, dando-se prioridade aos dados obtidos por essas últimas.

Os critérios para seleção das unidades consideradas nas Estatísticas do Cadastro Central de Empresas referentes ao ano de 2005 foram descritos no item Critério de seleção das unidades ativas.

Classificação de atividades

Todas as unidades estão classificadas de acordo com a principal atividade econômica desenvolvida com base na Classificação Nacional de Atividades Econômicas - CNAE 1.0², que é a classificação oficialmente utilizada pelo Sistema Estatístico Nacional, cuja estrutura de códigos é apresentada nos Anexos.

Disseminação dos resultados

Conteúdo das tabelas

Os resultados estão organizados em 12 tabelas impressas constantes desta publicação, e em 39 tabelas no CD-ROM que a acompanha. Todas as tabelas impressas constam, também, do CD-ROM, com a mesma numeração. Nos Quadros 1 e 2, a seguir, relaciona-se o conjunto de tabelas para o volume da publicação e do CD-ROM, respectivamente, associando a cada uma seu conteúdo, como um guia de leitura para os usuários.

Quadro 1 – Apresentação do conteúdo das tabelas da publicação.

Quadro 2 – Apresentação do conteúdo das tabelas do CD-ROM.

Dentre as tabelas publicadas, oito referem-se a dados de empresas e outras organizações, e quatro a dados das unidades locais. Estatísticas relativas a cada um dos municípios brasileiros estão disponibilizadas no CD-ROM.

O confronto desses dados com outras informações publicadas pelo IBGE deve levar em consideração o ano de referência das bases de dados em que as pesquisas se apóiam, a cobertura de cada uma das pesquisas envolvidas, a unidade de investigação das mesmas e os conceitos implícitos na descrição de cada variável.

Solicitações de tabulações especiais e dúvidas relacionadas a aspectos metodológicos podem ser encaminhadas para o *e-mail*: ibge@ibge.gov.br e endereçadas à Gerência do Cadastro Central de Empresas, da Diretoria de Pesquisas.

² Para conhecer as regras da Classificação Nacional de Atividades Econômicas - CNAE, bem como sua interpretação e estrutura de códigos, consultar a página de Classificações Estatísticas no endereço: <http://www.ibge.gov.br/concla>.

Quadro 1 - Apresentação do conteúdo das tabelas impressas

Conteúdo	Número das tabelas											
	1	2	3	4	5	6	7	8	9	10	11	12
Unidades de referência												
Empresas e outras organizações	X	X	X	X	X	X	X	X				
Unidades locais									X	X	X	X
Natureza Jurídica												
1 - Administração pública								X				
2 - Entidade empresariais						X	X	X				
3 - Entidade sem fins lucrativos								X				
Variáveis												
Número de empresas e outras organizações	X	X	X	X	X			X				
Número de empresas						X	X					
Número de unidades locais									X	X	X	X
Pessoal ocupado total	X	X	X		X	X	X	X	X	X	X	X
Pessoal ocupado assalariado	X	X	X		X			X	X	X	X	X
Salários e outras remunerações	X	X	X					X	X	X	X	X
Salário médio mensal	X	X			X				X			
Períodos de ano de fundação				X								
Natureza jurídica								X				
Níveis de agregação regional												
Brasil	X	X	X	X	X	X	X	X	X	X	X	
Grandes Regiões										X		
Unidades da Federação										X		
Municípios das Capitais												X
Municípios												
Classificação Nacional de Atividades Econômicas - CNAE 1.0												
Total geral	X	X	X	X				X	X	X	X	
Totais por seção	X	X	X	X		X	X	X	X	X	X	X
Totais por divisão		X	X			X	X					
Totais por grupo		X				X						
Totais por classe					X							
Tamanho por faixas de pessoal ocupado total												
Total geral	X		X	X					X		X	
Totais por seção			X								X	
Totais por divisão			X									
Indicadores especiais												
Índices de concentração econômica						X						
Diversificação espacial e de atividades							X					

Regras de arredondamento

O arredondamento foi feito aumentando-se de uma unidade à parte inteira do total da variável, quando a parte decimal era igual ou superior a 0,5. Desse modo, podem ocorrer pequenas diferenças de arredondamento entre os totais apresentados e a soma das parcelas em uma mesma tabela, bem como entre a mesma variável apresentada em tabelas distintas.

Regras de desidentificação

Considera-se que há risco de identificação do informante quando o número de unidades, para o nível de agregação tabulado, for igual ou inferior a dois. Neste caso, os dados não podem ser divulgados.

Devido à legislação que assegura o sigilo das informações estatísticas, foram adotadas regras de desidentificação para evitar a individualização dos informantes a partir dos dados divulgados. A regra básica consiste em desidentificar, no mesmo nível de subtotalização ou totalização, as colunas para as quais se tenham informações relativas a apenas uma ou duas unidades econômicas. Tal procedimento consistiu em aplicar um (x), na célula correspondente ao valor a ser omitido, nas variáveis Pessoal Ocupado Total, Pessoal Ocupado Assalariado, e Salários e Outras Remunerações, preservando-se os valores referentes ao número de unidades (empresas ou unidades locais) que não sofreram desidentificação.

Em alguns casos, pode ocorrer omissão de informação referente a um conjunto maior de unidades, visando a preservar possíveis identificações através de diferenças entre os níveis de totalização das tabelas.