NONONONONONONONO

**RELATÓRIO**

**METODOLÓGICO**

—

**PESQUISA**

**TIC EMPRESAS**

2021



**Relatório Metodológico TIC Empresas**

Comitê Gestor da Internet no Brasil (CGI.br), por meio do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), departamento do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), apresenta a metodologia da Pesquisa sobre o uso das tecnologias de informação e comunicação nas empresas brasileiras

**O**

– TIC Empresas. A pesquisa é realizada em todo o território nacional e aborda os seguintes temas:

* **Módulo A:** Informações gerais sobre os sistemas TIC;
* **Módulo B:** Uso da Internet;
* **Módulo E:** Comércio eletrônico;
* **Módulo F:** Habilidades no uso das TIC;
* **Módulo G:** *Software;*
* **Módulo H:** Novas tecnologias;
* **Módulo X:** Privacidade e proteção de dados.

# Objetivos da pesquisa

A pesquisa TIC Empresas tem como objetivo principal medir a posse e o uso das tecnologias de informação e comunicação (TIC) entre as empresas brasileiras com 10 ou mais pessoas ocupadas.

# Conceitos e definições

A pesquisa TIC Empresas é desenvolvida com a preocupação de manter a comparabilidade internacional. Para isso, faz-se uso dos padrões metodológicos propostos no manual da Conferência das Nações Unidas sobre Comércio e

Desenvolvimento (UNCTAD), elaborado pela parceria entre a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o Instituto de Estatísticas da Comissão Europeia (Eurostat) e o Partnership on Measuring ICT for Development (UNCTAD, 2009) – este último, uma coalizão formada por diversas organizações internacionais, busca a harmonização de indicadores-chave em pesquisas sobre TIC.

## MERCADO DE ATUAÇÃO

Para a definição do público-alvo da pesquisa, é utilizada a Classificação Nacional das Atividades Econômicas (CNAE 2.0) e a Tabela de Natureza Jurídica 2009.1, da Comissão Nacional de Classificação (Concla).

A Tabela de Natureza Jurídica identifica a constituição jurídico-institucional das entidades públicas e privadas no país segundo cinco grandes categorias: administração pública; entidades empresariais; entidades sem fins lucrativos; pessoas físicas e organizações internacionais; e outras instituições extraterritoriais.

A CNAE pode ser definida como uma estrutura-base sobre a qual as pessoas jurídicas no Brasil estão categorizadas de acordo com suas atividades econômicas, oficialmente adotada pelo Sistema Estatístico Nacional e pelos órgãos federais gestores de registros administrativos. A CNAE 2.0 é derivada da International Standard Industrial Classification of All Economic Activities (ISIC 4), cujo gestor é a Divisão de Estatísticas das Nações Unidas (UNSD).

A CNAE 2.0 não distingue os tipos de propriedade, natureza jurídica, tamanho do negócio, modo de operação e a legalidade da atividade. Sua estrutura hierárquica tem cinco níveis de detalhamento: seções, divisões, grupos, classes e subclasses. Para a TIC Empresas, utiliza-se o nível seção para classificação das empresas em seus mercados de atuação. As seções “Atividades imobiliárias” (Seção L), “Atividades profissionais, científicas e técnicas” (Seção M) e “Atividades administrativas e serviços complementares” (Seção N) foram agrupadas em uma só categoria (L+M+N). Já as seções “Artes, cultura, esporte e recreação” (Seção R) e “Outras atividades de serviços” (Seção S) foram agrupadas em uma categoria (R+S).

## PORTE

A pesquisa TIC Empresas considera pequenas, médias e grandes empresas aquelas com, respectivamente, 10 a 49 pessoas ocupadas, 50 a 249, e 250 pessoas ocupadas ou mais. As microempresas, aquelas com 1 a 9 pessoas ocupadas, não entram no escopo da pesquisa.

## PESSOAS OCUPADAS

Pessoas ocupadas são aquelas com ou sem vínculo empregatício, remuneradas diretamente pela empresa. Sendo que o número de pessoas ocupadas considera os assalariados, autônomos remunerados diretamente pela empresa, empregadores e sócios, pessoas da família e trabalhadores temporários. Não são considerados terceirizados e consultores.

# População-alvo

O universo abordado na pesquisa compreende todas as empresas brasileiras ativas com 10 ou mais pessoas ocupadas cadastradas no Cadastro Central de Empresas (Cempre) do Instituto Brasileiro de Geografia e Estatística (IBGE), pertencentes aos setores da CNAE 2.0 de interesse da pesquisa TIC Empresas e à Natureza Jurídica 2 – entidades empresariais, exceto as empresas públicas (Natureza Jurídica 201-1), de maneira a preservar a comparabilidade internacional. As empresas investigadas correspondem às seções:

C – Indústria de transformação; F – Construção;

G – Comércio; reparação de veículos automotores e motocicletas; H – Transporte, armazenagem e correio;

I – Alojamento e alimentação; J – Informação e comunicação; L – Atividades imobiliárias;

M – Atividades profissionais, científicas e técnicas;

N – Atividades administrativas e serviços complementares; R – Artes, cultura, esporte e recreação;

S – Outras atividades de serviços.

# Unidade de análise e referência

A unidade de investigação é a empresa, que, segundo o IBGE, é definida como a pessoa jurídica caracterizada por uma firma ou razão social que engloba o conjunto de atividades econômicas exercidas em uma ou mais unidades locais (o espaço físico, geralmente uma área contínua, onde uma ou mais atividades econômicas são desenvolvidas, correspondendo a um endereço de atuação da empresa).

Como o Cempre é composto por estabelecimentos e unidades locais, é necessário adequar a base de dados, de modo a obter um universo composto por empresas. Isso é obtido depois de adotados os seguintes procedimentos:

* As empresas são ordenadas por meio do número do Cadastro Nacional de Pessoa Jurídica (CNPJ).
* As unidades locais são agrupadas pelos oito primeiros dígitos do CNPJ, que são os que identificam a empresa. Nesse processo, são mantidas as informações de seção CNAE e de região do primeiro registro. Além disso, soma-se o número de pessoas ocupadas de todas as unidades locais.
* São excluídas as empresas com menos de 10 pessoas ocupadas no campo criado na etapa anterior.
	+ São excluídas as empresas que pertencem às seções A, B, D, E, K, O, P, Q, T e U, pois não pertencem à população-alvo da pesquisa.
	+ São excluídas as empresas que não pertencem à Natureza Jurídica 2, que engloba as entidades empresariais. Também são excluídas as empresas públicas que pertencem à Natureza Jurídica 201-1.

# Domínios de interesse para análise e divulgação

Para as unidades de análise e referência, os resultados são divulgados para domínios definidos com base nas variáveis e níveis descritos a seguir:

* **região:** corresponde à divisão regional do Brasil, segundo critérios do IBGE, nas macrorregiões Centro-Oeste, Nordeste, Norte, Sudeste e Sul;
* **porte:** corresponde à divisão por pequenas, médias e grandes empresas segundo o número de pessoas ocupadas, respectivamente, de 10 a 49 pessoas ocupadas, de 50 a 249, e 250 pessoas ocupadas ou mais. Destaca-se que, desde a edição 2017, a informação divulgada tem como base aquela disponível no cadastro e não a declarada pelo respondente no momento da entrevista, como acontecia até a edição de 2015;
* **mercados de atuação – CNAE 2.0:** corresponde à classificação das empresas nas seções mostradas como: C, F, G, H, I, J, L+M+N, R+S.

# Instrumento de coleta

## INFORMAÇÕES SOBRE OS INSTRUMENTOS DE COLETA

Para coleta das informações de interesse na pesquisa foi construído um questionário estruturado, com perguntas fechadas e abertas (quando fosse o caso). Para mais informações a respeito do questionário, ver item “Instrumento de coleta” em “Relatório de Coleta de Dados”.

# Plano amostral

O plano amostral é estratificado, e as empresas são selecionadas aleatoriamente dentro de cada estrato.

## CADASTRO E FONTES DE INFORMAÇÃO

O Cempre do IBGE fornece a consolidação e a atualização das informações de empresas e outras organizações formais, inscritas no CNPJ da Secretaria da Receita Federal, e suas respectivas unidades locais que responderam às pesquisas econômicas do IBGE e/ou declararam a Relação Anual de Informações Sociais (Rais). O IBGE disponibiliza anualmente um panorama geral das organizações formais ativas no país, com destaque para informações sobre natureza jurídica, pessoas ocupadas e atividades econômicas.

Com o objetivo de produzir um retrato do uso das TIC nas empresas brasileiras, considerando-se as diferenças entre os mercados de atuação, portes (número de pessoas ocupadas) e regiões brasileiras, a pesquisa TIC Empresas utiliza informações oriundas do Cempre, que serve como cadastro-base para o desenho da amostra e para a seleção das empresas a serem contatadas. A escolha das seções da CNAE, assim como a da estrutura de porte das empresas, segue as recomendações propostas no manual estatístico da UNCTAD (2009).

FIGURA 1

**PLANO AMOSTRAL DA PESQUISA TIC EMPRESAS**

**PLANO AMOSTRAL**

**TIC EMPRESAS**

**PORTE REGIÃO MERCADO DE ATUAÇÃO**

**Metodologias internacionais (UNCTAD, OCDE,**

**Eurostat e Partnership on Measuring ICTs for Development)**

**Concla / CNAE 2.0**

**Cempre**

**(IBGE)**


## CRITÉRIOS PARA DESENHO DA AMOSTRA

A amostra da pesquisa é desenhada utilizando-se a técnica de amostragem estratificada, que visa melhorar a precisão das estimativas e garantir a inclusão de subpopulações de interesse. A estratificação ocorre em duas etapas.

A primeira delas compreende a definição de estratos naturais a partir do cruzamento das variáveis: região geográfica (Centro-Oeste, Nordeste, Norte, Sudeste e Sul) e mercado de atuação CNAE 2.0 (C, F, G, H, I, J, L+M+N, R+S), conforme descrito na seção “Domínios de interesse para análise e divulgação”. Assim, são formados 40 estratos naturais não nulos. A partir de cada estrato natural, são definidos os estratos finais, que consideram a divisão dos estratos naturais por porte da empresa. As faixas de porte consideradas são: 10 a 19 pessoas ocupadas; 20 a 49 pessoas ocupadas; 50 a 249 pessoas ocupadas; e 250 pessoas ocupadas ou mais.

Não havendo empresas no universo em algum estrato, esse estrato é agrupado com uma faixa de porte anterior, mantendo as informações de região e mercado de atuação.

Definidas as variáveis de estratificação, os estratos possibilitam que todas as regiões, mercados de atuação e portes estejam representados na amostra, além de permitir análises para os domínios definidos por essas três variáveis individualmente. Contudo, não é possível tirar conclusões para categorias resultantes do cruzamento entre pares de variáveis.

## DIMENSIONAMENTO DA AMOSTRA

O tamanho da amostra planejada da pesquisa TIC Empresas é de 7 mil empresas.

## ALOCAÇÃO DA AMOSTRA

A amostra de empresas é obtida por amostragem aleatória simples sem reposição em cada estrato final. Dessa forma, as probabilidades de seleção são iguais dentro de cada estrato final.

A alocação da amostra de 7 mil empresas considera as distribuições marginais das variáveis “mercado de atuação” e “região”. Para as regiões, há um aumento na participação para obtenção da amostra final, dado que essa variável apresenta menos estratos a serem representados.

A distribuição do total de empresas por porte segue a distribuição: empresas de 10 a 19 pessoas ocupadas representam 30% da amostra, de 20 a 49 pessoas ocupadas

representam 25%, também 25% para 50 a 249 pessoas ocupadas e 20% para empresas com 250 pessoas ocupadas ou mais. Além disso, em estratos que têm o universo com até dez empresas, aloca-se apenas uma entrevista. Também se controla a fração amostral dentro de cada estrato, ou seja, a razão entre o tamanho da amostra e o tamanho da população – assim, em cada estrato pode-se ter no máximo 30% dessa fração. O tamanho final da amostra foi distribuído pelos estratos pré-definidos e é apresentada no “Relatório de Coleta de Dados”.

***n*** =***n*** ×

***N***

***h***

***h***

***N***

***N* é o tamanho total da população**

***Nh* é o tamanho total da população do estrato *h n* é o tamanho da amostra**

***nh* é o tamanho da amostra dentro de cada estrato *h***

FÓRMULA 1

## SELEÇÃO DA AMOSTRA

Dentro de cada estrato, as empresas são selecionadas por amostragem aleatória simples, conforme a Fórmula 1.

Logo, as probabilidades de inclusão das unidades de amostragem ***i*** para cada estrato ***h***

são dadas pela Fórmula 2.

π =

***n***

***h***

***ih***

***Nh***

FÓRMULA 2

Considera-se a taxa de resposta das empresas da onda anterior e, com isso, é selecionada aleatoriamente em cada estrato uma amostra reserva com o intuito de aproximar a amostra final do número inicialmente previsto de empresas. O uso da amostra reserva depende dos controles realizados para obtenção de entrevistas.**1**

# Coleta de dados em campo

## MÉTODO DE COLETA

As empresas são contatadas por meio da técnica de entrevista telefônica assistida por computador (do inglês, *computer-assisted telephone interviewing* [CATI]).

Em todas as empresas pesquisadas, busca-se entrevistar o responsável pela área de informática, tecnologia da informação, gerenciamento da rede de computadores ou área equivalente, o que corresponde a cargos como:

* diretor da divisão de informática e tecnologia;
* gerente de negócios (vice-presidente sênior, vice-presidente de linha de negócios, diretor);
* gerente ou comprador do departamento de tecnologia;
* influenciador tecnológico (funcionário do departamento comercial ou de operações de TI com influência sobre as decisões a respeito de questões tecnológicas);
* coordenador de projetos e sistemas;
* diretor de outros departamentos ou divisões (excluindo informática);
* gerente de desenvolvimento de sistemas;
* gerente de informática;
* gerente de projetos;
* dono da empresa ou sócio.

**1 Conforme disposto no item “Procedimentos e controle de campo” do "Relatório de Coleta de Dados".**

Nas empresas que declaram no momento da entrevista ter 250 pessoas ocupadas ou mais, a estratégia é entrevistar um segundo profissional, preferencialmente o gestor da área contábil ou financeira. Caso não seja encontrado, busca-se o responsável pela área administrativa, jurídica ou de relações com instituições governamentais, a quem cabem exclusivamente as respostas sobre comércio eletrônico e atividades realizadas na Internet.

# Processamento dos dados

## PROCEDIMENTOS DE PONDERAÇÃO

A cada empresa da amostra foi associado um peso amostral básico, obtido pela razão entre a quantidade de empresas existentes no estrato e o tamanho da amostra no estrato final correspondente, conforme a Fórmula 3.

FÓRMULA 3

FÓRMULA 4

Para corrigir os casos onde não se obtém a resposta de todos os selecionados, é realizada uma correção de não resposta. A correção de não resposta é dada pela Fórmula 4.

***w*** = =

**1 *N***

***h***

***ih***

π***ih***

***nh***

***wih* é o peso básico, inverso da probabilidade de seleção, associado a cada empresa respondente *i* no estrato *h***

***nh* é o tamanho da amostra de empresas no estrato *h Nh* é o total de empresas no estrato *h***

***w\**** =***w*** ×

***NH***

***ih ih***

***∑w***

***w\** é o peso com correção de não resposta da empresa *i* no estrato *h***

***ih***

***1 ih***

# Erros amostrais

As medidas ou estimativas dos erros amostrais dos indicadores da TIC Empresas levam em consideração em seus cálculos o plano amostral por estratos empregado na pesquisa.

Assim, a divulgação dos erros amostrais, expressos pela margem de erro, é feita a partir das variâncias estimadas. As margens de erro são calculadas para um nível de confiança de 95%. Isso indica que os resultados, baseados nessa amostra, são considerados precisos, dentro do intervalo definido pelas margens de erro. Se a pesquisa for repetida várias vezes, em 95% delas o intervalo poderá conter o verdadeiro valor populacional. Outras medidas derivadas dessa estimativa de variabilidade são comumente apresentadas, tais como erro padrão, coeficiente de variação ou intervalo de confiança.

O cálculo da margem de erro considera o produto do erro padrão (raiz quadrada da variância) pelo valor 1,96 (valor da distribuição amostral que corresponde ao nível de significância escolhido de 95%). Esses cálculos são feitos para cada variável de cada uma das tabelas, o que significa que todas as tabelas de indicadores possuem margens de erro relacionadas às suas estimativas apresentadas em cada célula.

# Disseminação dos dados

Os resultados desta pesquisa são divulgados de acordo com as seguintes variáveis de cruzamento: porte da empresa, mercado de atuação e região geográfica.

Arredondamentos fazem com que, em alguns resultados, a soma das categorias parciais difira de 100% em questões de resposta única. O somatório de frequências em questões de respostas múltiplas usualmente é diferente de 100%. Vale ressaltar que, nas tabelas de resultados, o hífen ( - ) é utilizado para representar a não resposta ao item. Por outro lado, como os resultados são apresentados sem casa decimal, as células com valor zero significam que houve resposta ao item, mas ele é explicitamente maior do que zero e menor do que um.

Os resultados desta pesquisa são publicados em formato *online* e disponibilizados no *website* ([https://www.cetic.br](https://www.cetic.br/)) e no portal de visualização de dados do Cetic.br|NIC.br ([https://data.cetic.br](https://data.cetic.br/)). As tabelas de proporções, totais e margens de erros calculadas para cada indicador estão disponíveis para *download* em português, inglês e espanhol. Mais informações sobre a documentação, os metadados e as bases de microdados da pesquisa estão disponíveis na página de microdados (https://[cetic.br/microdados/](http://cetic.br/microdados/)).

# Referências

Conferência das Nações Unidas sobre Comércio e Desenvolvimento. (2009). *Manual for the production of statistics on the information economy 2009*. [http://www.unctad.org/en/](http://www.unctad.org/en/docs/sdteecb20072rev1_en.pdf) [docs/sdteecb20072rev1\_en.pdf](http://www.unctad.org/en/docs/sdteecb20072rev1_en.pdf)