**Avaliação de métodos de imputação
na variável *Receita* das empresas da
Pesquisa Anual de Comércio - PAC-IBGE**

João Carlos Silva Rodrigues

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Profa. Dra. Lúcia Pereira Barroso

Durante parte do desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, fevereiro de 2019

**Avaliação de métodos de imputação
na variável *Receita* das empresas da
Pesquisa Anual de Comércio - PAC-IBGE**

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 07/06/2019. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Profa. Dra. Lúcia Pereira Barroso - IME-USP
- Prof. Dr. Rinaldo Artes - Insper
- Profa. Dra. Denise Britz do Nascimento Silva - ENCE

“To call in the statician after the experiment is done may be no more than asking to perform a postmortem examination: he may be able to say what the experiment died of”
Fisher (1938)

Agradecimentos

Primeiramente, gostaria de agradecer à minha orientadora, Prof.^a Lúcia Pereira Barroso, pela paciência, disponibilidade e todo o apoio dado durante o processo de elaboração da presente dissertação. Os meus mais sinceros agradecimentos!

Também agradeço aos professores Rinaldo Artes e Denise Britz do Nascimento Silva pela participação na banca e por toda a contribuição oferecida.

Aos meus colegas de turma e aos tantos amigos que fiz na USP ao longo deste período, parceiros de angústias e alegrias; e aos amigos que fiz na EY para a vida, que me foram essenciais no período de adaptação à cidade de São Paulo - muito obrigado!

Obrigado, também, à atual gerente da equipe de Métodos da COSEC do IBGE, Adriana Bandeira Moraes, e a sua antecessora, Maria Deolinda Borges Cabral, pelo apoio, compreensão e solidariedade para com a elaboração desta dissertação. Não há como deixar de agradecer, também, a todos os meus colegas de trabalho, que muito colaboraram, cada um a sua maneira, no processo de aprendizado nessa minha trajetória. Importante, igualmente, citar os três coordenadores aos quais fui subordinado no IBGE ao longo desta jornada: Flávio Magheli, Vânia Prata e Alessandro Piniheiro, pela compreensão ao atendimento de necessidades específicas que envolveram a construção deste trabalho dissertativo.

Gostaria de agradecer também aos meus pais por sempre estarem ao meu lado, nos melhores e piores momentos da minha vida, independentemente das minhas escolhas, sempre dando o suporte necessário, fundamentalmente no aspecto emocional, para seguir sempre adiante e quebrar barreiras cada vez mais difíceis.

Um agradecimento especial ao meu companheiro, que é peça fundamental nessa conquista - me apoiando, me incentivando, brigando nos momentos necessários e me deixando quieto em outros, bem como se mostrando sempre parceiro ao meu lado independentemente das minhas decisões. Também por me dar o suporte que eu mais precisava, DIARIAMENTE, para seguir e conquistar os meus objetivos. Obrigado pelo carinho, pelo amor e pelo companheirismo de sempre! Te amo!

E, finalmente, a todos que de alguma forma colaboraram, não apenas de forma direta, mas também indiretamente em toda essa fase - muito OBRIGADO!

Resumo

RODRIGUES, J. C. S. **Avaliação de métodos de imputação na variável *Receita* das empresas da Pesquisa Anual de Comércio - PAC-IBGE**. 2019. 146 f. Dissertação - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

O presente trabalho utiliza as informações da Pesquisa Anual do Comércio - PAC, uma das quatro pesquisas econômicas estruturais do IBGE, para avaliar o Modelo de Imputação atual da pesquisa comparando-o com outros modelos disponíveis na literatura. Foi feito um recorte da base da PAC-IBGE dos anos de 2014 e 2015 e foram testados vinte modelos de imputação.

Na PAC, tem sido observado um aumento do impacto das não-respostas nas estimativas de seus totais. Isto deriva da alta assimetria das variáveis econômicas em conjunto com o pequeno número de empresas de alguns estratos, somados ainda ao aumento populacional de algumas atividades econômicas - e, por consequência, dos pesos amostrais - e ainda do elevado número de mortes (fechamento) de empresas pequenas. Tais problemas apresentados geram a necessidade de se estudar alternativas de tratamento para essas empresas não-respondentes.

Os modelos foram analisados selecionando algumas empresas aleatoriamente e assumindo que elas não tivessem respondido à pesquisa. Posteriormente, essas empresas foram submetidas aos modelos de imputação selecionados e os resultados foram avaliados utilizando Erro Quadrático Médio (EQM) e Variação Percentual (VP) dos totais estimados contra o real. Foi escolhida a variável de *RECEITA* para ser usada nos testes.

Os modelos utilizados podem ser agrupados em quatro grupos: de médias de respondentes; através de uma regressão com uso de variáveis auxiliares de cadastro; média dos respondentes mais próximos através de uma função distância; e através de uma regressão dos respondentes mais próximos com uso de uma função distância.

Ao final das análises, verificou-se que apesar de alguns modelos também terem tido bons desempenhos, não foi observado um fator relevante que indique a troca do modelo atual de imputação utilizado na PAC-IBGE.

Palavras-chave: Imputação de Pesquisas Econômicas, Pesquisa Econômica Estrutural por Amostra, Pesquisa Anual do Comércio do IBGE.

Abstract

RODRIGUES, J. C. S. **An evaluation of Imputation Methods on the *Revenue* variable from the Annual Survey of Commerce's (PAC-IBGE) companies 2019.** 146 f. Dissertation - Institute of Mathematics and Statistics, University of São Paulo, São Paulo, 2019.

The present work uses the information from the Annual Survey of Commerce - PAC, one of the four structural surveys of IBGE, to evaluate its current imputation model against other available models in the literature. The dataset used was obtained from PAC in the years of 2014 and 2015 and twenty imputation models were tested.

At PAC, there has been an increase in the impact of non-responses on its totals estimative. This is due to the high asymmetry of the economic variables together with the small number of companies of some strata, added to the population increase of some economic activities - and, consequently, of their sample weights - and also with the high number of deaths (closure) of small businesses. Such problems present the need to study alternatives treatments for these non-responding companies.

The analysis of models were made by selecting some companies randomly and assuming that they had not responded the survey. Subsequently, these companies were submitted to the selected imputation models and the results were evaluated using Mean Square Error (MSE) and the Percent Variation (PV) between the estimated totals against the real ones. The *Revenue* variable was the one chosen to be used in the tests.

The models used can be grouped into four groups: average of the respondents; through a regression function using auxiliary variables of cadastre; average of the closest respondents through a distance function; and through a regression function of the closest respondents using a distance function.

At the end of the analyzes, it was verified that although some imputation models presented good results, there is no relevant factor indicating the change of the current one.

Keywords: Imputation of Economic Surveys, Structural Economic Sample Surveys, Annual Survey of Commerce of IBGE.

Sumário

Lista de Abreviaturas	xi
Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Considerações gerais	1
1.2 Motivação	1
1.3 Objetivos	3
1.4 Organização do trabalho	4
2 O Instituto Brasileiro de Geografia e Estatística (IBGE) e a Pesquisa Anual do Comércio (PAC)	5
2.1 Origem e missão do IBGE	5
2.2 Principais objetivos da PAC	6
2.3 Âmbito e unidade de investigação da PAC	7
2.4 Conceituação das variáveis investigadas	9
2.5 Questionário da PAC	10
2.6 Plano amostral e cadastro básico de empresas da PAC	10
2.7 Tamanho da amostra	13
2.8 Tratamento da amostra	14
2.9 Cálculo das estimativas	16
2.9.1 Estimador simples	17
2.9.2 Estimador regressão	17
2.10 Processo de crítica e imputação	18
2.11 Disseminação dos resultados	19
3 Imputação e outras formas de lidar com a não-resposta	23
3.1 Conceitos gerais	23
3.2 Imputação	26
3.2.1 Notação e considerações gerais	27
3.2.2 Principais métodos de imputação	29
3.2.3 Método de imputação utilizado na PAC	32
3.2.4 Imputação em alguns institutos de pesquisa	38

4	Proposta de trabalho e base de dados utilizada	41
4.1	Modelos de imputação utilizados	41
4.1.1	Modelo atual de imputação do IBGE da PAC	42
4.1.2	Modelos de imputação de médias de grupos	42
4.1.3	Modelos de imputação por função de regressão	43
4.1.4	Modelos de imputação múltipla de uma função distância e de regressão	43
4.1.5	Modelos de imputação múltipla de uma função distância e da mediana de grupos	43
4.1.6	Vantagens e desvantagens dos modelos utilizados	44
4.2	Critério de avaliação dos modelos aplicados	44
4.2.1	Análise por VP	45
4.2.2	Análise por EQM	45
4.3	Descrição da base de dados utilizada	45
5	Resultados obtidos	61
5.1	EQM das imputações por CNAE x ESTF	62
5.1.1	ANO 2015 / UF 25 - PB / ESTF A1	62
5.1.2	ANO 2015 / UF 25 - PB / ESTF A2	65
5.1.3	ANO 2015 / UF 25 - PB / ESTF A3	68
5.1.4	ANO 2015 / UF 25 - PB / ESTF C	71
5.1.5	ANO 2015 / UF 25 - PB / ESTF G	74
5.1.6	ANO 2015 / UF 32 - ES / ESTF A1	77
5.1.7	ANO 2015 / UF 32 - ES / ESTF A2	80
5.1.8	ANO 2015 / UF 32 - ES / ESTF A3	83
5.1.9	ANO 2015 / UF 32 - ES / ESTF C	86
5.1.10	ANO 2015 / UF 32 - ES / ESTF G	89
5.2	Diferença percentual dos totais estimados x total real	92
5.2.1	Modelos de imputação de médias de grupos	92
5.2.2	Modelos de imputação por função de regressão	93
5.2.3	Modelos utilizando cálculo da distância do IBGE entre as empresas	96
5.2.4	Modelos utilizando distância euclidiana entre as empresas	97
6	Conclusões	99
6.1	Análise dos resultados	99
6.2	Considerações finais	108
6.3	Sugestões para trabalhos futuros	108
6.3.1	Tratamento para não-respostas e mortes de empresas	109
6.3.2	Calibração	109
6.3.3	Plano amostral	109
6.3.4	Avaliar impacto das imputações na variância do estimador	109
A	Questionário Simplificado da PAC	111
B	Questionário Completo da PAC	125

C Estatísticas	143
Referências Bibliográficas	145

Lista de Abreviaturas

CAGED	Cadastro Geral de Empregados e Desempregados
CBS	Cadastro Básico de Seleção
CEEC	Coordenação de Metodologia das Estatísticas de Empresas, Cadastros e Classificações
CEMPRE	Cadastro Central de Empresas
CDDI	Centro de Documentação e Disseminação de Informações
CNAE	Classificação Nacional das Atividades Econômicas
CNPJ	Cadastro Nacional de Pessoa Jurídica
CONCLA	Comissão Nacional de Classificação
CV	Coeficiente de Variação
DGC	Diretoria de Geociências
DI	Diretoria de Informática
DPE	Diretoria de Pesquisa
ENCE	Escola Nacional de Ciências Estatísticas
ESTF	Estrato Final
ESTN	Estrato Natural
FAC	Folha de Atualização Cadastral
FGTS	Fundo de Garantia por Tempo de Serviço
IBGE	Instituto Brasileiro de Geografia e Estatística
ONU	Organização das Nações Unidas
PAC	Pesquisa Anual do Comércio
PAS	Pesquisa Anual do Serviço
PAIC	Pesquisa Anual da Indústria da Construção
PIA	Pesquisa Anual da Indústria
RAIS	Relação Anual de Informações Sociais
UF	Unidade da Federação
PIS	Programa de Integração Social
PO	Pessoal Ocupado
SIMCAD	Sistema de Manutenção Cadastral
SITCAD	Situação Cadastral
SITEMP	Situação da Empresa

Lista de Figuras

2.1	Construção do CBS	13
2.2	Tratamento aplicado às empresas no momento da expansão dos pesos amostrais associados	16
3.1	Fluxo de imputação pelo modelo do IBGE	33
4.1	Grupo de modelos de imputação escolhidos para estudo	42
5.1	EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF A1	63
5.2	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10, DISTEUCLID_REG_N20, MÉDIA_ESTN e REG_ESTN): Ano 2015 / UF 25 - PB / ESTF A1	64
5.3	EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF A2	66
5.4	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, IBGE, DISTEUCLIDREG_medN5, MÉDIA_ESTN, REG_ESTF e REG_EST): Ano 2015 / UF 25 - PB / ESTF A2	67
5.5	EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF A3	69
5.6	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, DISTEUCLIDREG_medN5, DISTREG_N10 e N20, MÉDIA_ESTN e ESTRATO e REG_ESTF, ESTN e L4): Ano 2015 / UF 25 - PB / ESTF A3	70
5.7	EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF C	72
5.8	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20 e CNAE 468): Ano 2015 / UF 25 - PB / ESTF C	73
5.9	EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF G	75
5.10	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20): Ano 2015 / UF 25 - PB / ESTF G	76
5.11	EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF A1	78
5.12	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, DISTEUCLIDREG_medN5, DISTREG_N10 e N20, MÉDIA_ESTN e REG_ESTN): Ano 2015 / UF 32 - ES / ESTF A1	79
5.13	EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF A2	81
5.14	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, MÉDIA_ESTN e REG_ESTN e UF e sem a CNAE 462): Ano 2015 / UF 32 - ES / ESTF A2	82
5.15	EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF A3	84

5.16	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, DISTREG_N20, MÉDIA_ESTN, REG_ESTF, ESTN e UF e sem a CNAE 462): Ano 2015 / UF 32 - ES / ESTF A3	85
5.17	EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF C	87
5.18	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, MÉDIA ESTRATO e REG_ESTF): Ano 2015 / UF 32 - ES / ESTF C	88
5.19	EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF G	90
5.20	EQM dos modelos de imputação (exceto DISTEUCLID_REG_N10 e N20, DISTREG_N10 e N20, MÉDIA_ESTN, REG_ESTN, L4 e UF): Ano 2015 / UF 32 - ES / ESTF G	91
6.1	Resultados do EQM em 2014	102
6.2	Resultados do EQM em 2015	103
6.3	Resultados do EQM em 2014 consolidados por UF	104
6.4	Resultados do EQM em 2015 consolidados por UF	105
6.5	Resultados do VP em 2014 e 2015 consolidados por UF	107

Lista de Tabelas

2.1	Organização hierárquica da CNAE 2.0	8
2.2	Exemplo da hierarquia das CNAE 4511 e 4512	8
2.3	Associação das variáveis de interesse ao questionário da PAC	10
2.4	Lista de SITCAD e FAC	15
2.5	CNAE 45 - Comércio de veículos, peças e motocicletas	19
2.6	CNAE 46 - Comércio por atacado	20
2.7	CNAE 47 - Comércio varejista	21
3.1	Agrupamento para o cálculo da mediana da <i>RECEITA</i>	36
4.1	Número de Empresas por SITEMP para cada UF no ano de 2014	46
4.2	Número de Empresas por SITEMP para cada UF no ano de 2015	48
4.3	Número de Estratos Naturais por UF em 2014	50
4.4	Número de Estratos Naturais por UF em 2015	51
4.5	Número de ESTF* e ESTF em 2014 e 2015 e os respectivos percentuais de ESTF*	52
4.6	Percentual de Estratos Naturais com pelo menos quatro empresas ativas em 2014	53
4.7	Percentual de Estratos Naturais com pelo menos quatro empresas ativas em 2015	54
4.8	Número Total de Estratos Finais nos estados selecionados em 2014	55
4.9	Número Total de Estratos Finais para as UF selecionadas em 2015	56
4.10	Número de Estratos Finais com pelo menos quatro Empresas Ativas para as UFs selecionadas em 2014	57
4.11	Número de Estratos Finais com pelo menos quatro Empresas Ativas para as UFs selecionadas em 2015	58
4.12	Percentual de Estratos Finais com pelo menos quatro empresas ativas em 2014	59
4.13	Percentual de Estratos Finais com pelo menos quatro empresas ativas em 2015	59
4.14	Número de empresas na base utilizada para testar os modelos de imputação por ano e UF	60
5.1	Totais de <i>RECEITA</i> estimado versus real e VP do modelo atual do IBGE	92
5.2	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_ESTRATO	92
5.3	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_ESTN	92
5.4	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_UF	93
5.5	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_L4ESTF	93
5.6	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_ESTREG	93
5.7	Totais de <i>RECEITA</i> estimado versus real e VP do modelo MÉDIA_ESTBR	93

5.8	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_ESTF	94
5.9	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_ESTN	94
5.10	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_UF	94
5.11	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_L4	95
5.12	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_ESTF_REGIONAL	95
5.13	Totais de <i>RECEITA</i> estimado versus real e VP do modelo REG_ESTF_BRASIL .	95
5.14	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTREG_N10	96
5.15	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTREG_N20	96
5.16	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTREG_medN5 . . .	96
5.17	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTREG_medN10 . .	97
5.18	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTEUCLID_REG_N10	97
5.19	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTEUCLID_REG_N20	97
5.20	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTEUCLIDREG_medN5	98
5.21	Totais de <i>RECEITA</i> estimado versus real e VP do modelo DISTEUCLIDREG_medN10	98
C.1	Número de empresas na base utilizada para testar os modelos de imputação por UF e ESTF em 2014	143
C.2	Número de empresas na base utilizada para testar os modelos de imputação por UF e ESTF em 2015	143
C.3	Frequência de Empresas por tipo de ação no questionário na PAC em 2014	143
C.4	Frequência de Empresas por tipo de ação no questionário na PAC em 2015	144
C.5	Código das Unidades Federativas e respectivas Regiões	144
C.6	Agrupamentos de classes da CNAE 2.0 nos Estados de Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul trabalhados a três dígitos .	144

Capítulo 1

Introdução

1.1 Considerações gerais

A Pesquisa Anual do Comércio (PAC) é uma pesquisa de abrangência nacional realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e será a base de estudo para esta dissertação. A PAC tem por objetivo dois papéis essenciais:

- Descrever as características estruturais das atividades relacionadas às empresas do segmento do comércio no Brasil, assim como suas mudanças ao longo do tempo; e
- Constituir um núcleo de informações dessas atividades para outras pesquisas econômicas, também de interesse nacional.

Assim, um dos principais objetivos da PAC é calcular estimativas para os totais populacionais das variáveis econômicas do setor de comércio atacadista e varejista do país, tais como Número de Funcionários, Total de Salário e Total de Receita, dentre outros.

Para atender aos seus objetivos, a PAC faz uso de um plano amostral que combina uma amostragem estratificada simples para empresas de porte pequeno a médio, e censo para as empresas consideradas grandes (estrato certo). Com isso, os estratos da PAC são formados considerando o porte da empresa, a unidade da federação onde se localiza a sede da empresa e a sua principal atividade econômica.

1.2 Motivação

Na criação da PAC, o seu plano amostral foi feito levando em consideração três características principais das empresas: sua principal atividade econômica, a localização da sede e o número total

de Pessoas Ocupadas (*PO*) - pois quando o mesmo foi elaborado, o número de pessoas ocupadas era um bom indicador para definir o tamanho (porte) de uma empresa e, a partir dessa informação, formavam-se os estratos assumindo homogeneidade quanto aos seus indicadores financeiros e contábeis.

Entretanto, com o passar dos anos e com o desenvolvimento das atividades econômicas do comércio, tais premissas não são mais necessariamente efetivas, uma vez que pequenas empresas podem apresentar grandes faturamentos e vice-versa, em função da própria atividade (e não em decorrência de algum fator externo – político, econômico, climático etc). Atenta-se para o fato de que o próprio conceito de empresa pequena pode ser redefinido dependendo da atividade econômica da qual estejamos tratando.

Limitações de orçamento e de recursos humanos do IBGE quanto à coleta das informações e limitações de acordos políticos para obtenção de novos registros administrativos para acesso a informações adicionais que contribuam na elaboração de um plano amostral de maior qualidade, associadas ao número elevado de não-respostas, recusas e/ou mortes de empresas, em alguns casos, podem causar grandes impactos nas estimativas da pesquisa. Esses fenômenos de não-respostas, recusas e/ou mortes de empresas são comuns em empresas de pequeno porte, pois elas iniciam e cessam suas atividades - nascem e morrem - ou mudam de endereço com mais facilidade do que as demais.

Os problemas citados anteriormente podem ser graves, pois dependendo da atividade econômica há estratos cuja população de empresas é bastante elevado. Com isso, empresas que, sozinhas, já teriam um grande peso amostral, representam um número ainda maior de empresas em razão da reponderação dos pesos amostrais dos não-respondentes.

Tais problemas identificados ainda se agravam, pois a pesquisa trabalha com variáveis econômicas, que, em geral, apresentam grande variância dentro dos estratos e, conseqüentemente, dos estimadores.

Sendo assim, devido ao número crescente de não-respostas e de morte na pesquisa, combinado com o aumento também crescente do tamanho do estrato certo (e conseqüente diminuição do estrato amostrado), uma preocupação recorrente (e cada vez maior) é a de qual seria o melhor tratamento para essas não-respostas, ou seja, qual seria o melhor tratamento estatístico (ou de cadastro) para cada uma delas.

A escassez de estudos referentes a pesquisas econômicas aumenta ainda mais a necessidade em se analisar esses impactos e verificar possíveis tratamentos e alternativas eficazes aplicáveis à PAC e outras pesquisas econômicas.

1.3 Objetivos

Considerando Lohr [2010], aumentar a amostra sem investigar a não-resposta não melhora em nada o seu viés, mas apenas acresce o número de não-respondentes referentes à classe a qual a empresa pertence. Para a referida autora, aumentar o tamanho da amostra indefinidamente tende apenas a piorar o viés da não-resposta e, portanto, a melhor forma de impedi-la seria evitar o seu acontecimento na origem. Para tanto, seria necessário fazer alterações ou na elaboração e/ou na forma de coleta dos questionários ou ainda na construção do cadastro utilizado para seleção da amostra. Contudo, devido às dificuldades mencionadas anteriormente, tanto de razões orçamentárias como humanas, e somadas às restrições de cadastro – provenientes de outro órgão público, o recém extinto Ministério do Trabalho –, esse tratamento seria muito custoso à pesquisa.

Lohr [2010] também discute outras três propostas de tratamento à não-resposta, além da prevenção: usar subamostras de não-respondentes para fazer inferência sobre os mesmos; usar um modelo para prever os valores dos não-respondentes; e ignorar a não-resposta, sendo este último não recomendado pela referida autora.

Dito isso, verifica-se a importância de se fazer um estudo sobre essas empresas, assim como investigar novos e melhores métodos de imputação para os dados das mesmas, de modo a melhorar as estimativas dos totais e diminuir a variância das estimativas, sem comprometer a distribuição original dos dados.

Este último ponto vai ao encontro do princípio número 5 do Código de Princípios Fundamentais das Estatísticas Oficiais do IBGE que versa sobre a eficiência do uso das fontes de dados das pesquisas, levando em consideração a qualidade, oportunidade, custos e ônus ao informante, conforme IBGE [2019].

Apesar de todos os problemas descritos anteriormente, na presente dissertação optou-se por abordar somente àqueles referentes exclusivamente aos modelos de predição dos valores das empresas não-respondentes, mais especificamente, aos Métodos de Imputação da variável *RECEITA* das Empresas de Comércio que atendam ao âmbito de pesquisa da PAC-IBGE.

Dessa forma, propõe-se estudar diferentes métodos de imputação sempre comparando-os ao Método de Imputação utilizado atualmente pelo IBGE a fim de verificar a existência de métodos mais eficazes do que o atual.

1.4 Organização do trabalho

O presente trabalho utilizou os dados referentes às unidades federativas da Paraíba, do Espírito Santo, do Paraná e do Rio Grande do Sul, que representam as quatro unidades da federação com maior número de estratos com pelo menos quatro empresas respondentes nos anos de 2014 e 2015, aos quais aplicaram-se os vinte modelos de imputação adotados para comparação com o modelo de imputação implementado pelo IBGE.

Para comparar os modelos selecionados foram escolhidos dois critérios de avaliação: Erro Quadrático Médio (EQM) e Variação Percentual (VP) do total estimado *versus* o total real. Para a melhor apresentação da análise foram utilizados gráficos e tabelas os quais expõem os resultados obtidos.

No Capítulo 2, será feito um breve resumo da PAC-IBGE e da sua importância para introduzir e conceituar o âmbito da pesquisa, as etapas e a estrutura da mesma, bem como as variáveis utilizadas no desenvolvimento do presente trabalho.

No Capítulo 3, serão abordados os principais métodos de imputação discutidos na literatura, tomando como base Albieri [1989], como também será descrito o modelo de imputação utilizado na PAC-IBGE quando necessário.

Posteriormente, no Capítulo 4, serão apresentados os modelos escolhidos para serem estudados, bem como as formas adotadas para a avaliação dos mesmos. Também serão descritos os critérios para construção da base de dados final utilizada para a avaliação dos modelos selecionados.

Finalmente, no Capítulo 5, serão apresentados os resultados encontrados para cada análise feita; e na conclusão do trabalho, no Capítulo 6, serão feitas as análises finais dos resultados expostos no capítulo anterior.

Capítulo 2

O Instituto Brasileiro de Geografia e Estatística (IBGE) e a Pesquisa Anual do Comércio (PAC)

2.1 Origem e missão do IBGE

Segundo a página oficial do [IBGE \[2017a\]](#) na Rede Internet, a instituição foi criada durante o Estado Novo, em 1936, em decorrência da necessidade do poder público federal em criar um órgão que pudesse articular e coordenar pesquisas estatísticas sob o ponto de vista de serviços especializados no Brasil. Dessa forma, o órgão passa a ser o principal fornecedor de dados e informações do país e segue até a presente data cumprindo esse papel.

Ainda de acordo com a página oficial do IBGE, sua missão consiste em identificar e analisar o território brasileiro, contar a sua população, mostrar como a sua economia evolui através do trabalho e da produção das pessoas e revelar como as pessoas vivem. Ou ainda, segundo o seu próprio estatuto:

“retratar o Brasil, com informações necessárias ao conhecimento da sua realidade e ao exercício da cidadania, por meio da produção, análise, pesquisa e disseminação de informações de natureza estatística - demográfica e socioeconômica, e geocientífica - geográfica, cartográfica, geodésica e ambiental [BRASIL \[2003\]](#).”

O IBGE é dividido atualmente em seis unidades:

- Diretoria Executiva (DE);

- Diretoria de Pesquisa (DPE);
- Diretoria de Geociências (DGC);
- Diretoria de Informática (DI);
- Centro de Documentação e Disseminação de Informações (CDDI); e
- Escola Nacional de Ciências Estatísticas (ENCE).

Cabe à DPE, diretoria na qual a PAC está inserida, a responsabilidade de produzir e sistematizar estudos, pesquisas e trabalhos de natureza estatística concernente à situação demográfica, econômica, social, ambiental e administrativa do país.

2.2 Principais objetivos da PAC

A PAC é uma pesquisa vinculada à Coordenação de Pesquisas Estruturais por Empresas cujo objetivo é acompanhar e retratar a estrutura produtiva dos diversos segmentos do Comércio através de diferentes indicadores econômicos.

A PAC foi instituída em 1988 com o intuito de suprir a falta de informações estatísticas acerca das empresas comerciais entre os anos intercensitários das atividades econômicas que ocorriam de cinco em cinco anos.

A partir de 1994, deu-se início a um processo de modernização das pesquisas econômicas que começou com a instalação do novo sistema de classificação das atividades econômicas descritos através da Classificação Nacional das Atividades Econômicas (CNAE), passando por uma mudança estrutural no modelo da pesquisa: de censos quinquenais para bases amostrais realizadas anualmente, mantendo o tratamento censitário para as maiores empresas – empresas com vinte ou mais Pessoas Ocupadas (*PO*) – do comércio.

Com isso, através da PAC é possível:

- descrever as características estruturais básicas do segmento empresarial varejista e atacadista do país, bem como retratar as suas transformações ao longo do tempo;
- analisar os diversos segmentos pesquisados extraindo estimativas de valor adicionado, emprego, salários e formação de capital;
- retratar o desempenho do comércio através de indicadores sobre a atividade das empresas, tais como faturamento médio por estabelecimento, margem de comercialização, giro de estoque, dentre outros;

- detalhar o comércio de supermercados, hipermercados e lojas de departamentos através de vendas por m^2 e receita por pessoa ocupada;
- dimensionar a distribuição da atividade regional da atividade comercial do país segundo classes de atividades econômicas para as Unidades da Federação (UF);
- apresentar a distribuição espacial das empresas comerciais para cada Unidade da Federação por atividade econômica; e
- fornecer informações sobre as empresas do comércio para outros departamentos do IBGE que assim o necessitem, como o Departamento de Contas Nacionais.

Segundo o Relatório Metodológico da PAC (IBGE [2000]), o objetivo da pesquisa é a “obtenção de estimativas dos totais populacionais referentes às variáveis investigadas, por Unidade da Federação, segundo os níveis de classificação de atividade definidos para as empresas”.

2.3 Âmbito e unidade de investigação da PAC

De acordo com o Relatório Metodológico da PAC (IBGE [2000]), entende-se por empresas comerciais aquelas cuja Receita Bruta seja proveniente de atividade comercial, sendo tais atividades descritas especificamente na seção G da CNAE 2.0 que aborda as atividades de Comércio, Reparação de veículos, objetos pessoais e domésticos, apresentada em IBGE [2007].

A CNAE foi criada com o objetivo de organizar as informações do setor por segmentos homogêneos, como por exemplo, em relação ao tipo de processo da produção, o mercado de insumos e o mercado final. Sua criação foi uma demanda tanto interna quanto externa cujo motivo foi o de padronizar as diversas fontes que se utilizam dessas informações numa linguagem única, como feito pela ONU, permitindo assim a comparação das estatísticas do mercado das empresas comerciais brasileiras com as de outros países e permitindo também fazer análises em diferentes níveis de agregação e por diferentes tipos de registros administrativos.

A Comissão Nacional de Classificação (CONCLA) é a comissão responsável pela criação e atualização da CNAE. A CONCLA divide as atividades econômicas de forma hierárquica em cinco níveis: seções, divisões, grupos, classes e subclasses, sendo a primeira a mais agregada e a última a mais desagregada com o maior detalhamento da atividade. Na Tabela 2.1, retirada de IBGE [2007], é descrita a estrutura hierárquica da CNAE e a respectiva agregação para cada um.

As empresas que respondem à pesquisa o fazem no nível de quatro dígitos da CNAE, equivalente ao nível de classe. Elas informam, então, os quatro primeiros dígitos da CNAE da sua principal

Tabela 2.1: *Organização hierárquica da CNAE 2.0*

Nome	Nível	Número de grupamentos	Identificação
Seção	Primeiro	21	Código alfabético de 1 dígito
Divisão	Segundo	87	Código alfabético de 2 dígitos
Grupo	Terceiro	285	Código alfabético de 3 dígitos
Classe	Quarto	673	Código alfabético de 4 dígitos + DV
Subclasse	Quinto	1301	Código alfabético de 7 dígitos (incluindo o DV)

atividade econômica que é aquela responsável pela maior parte de sua receita.

Na PAC, a unidade de investigação é a própria empresa, pois possui obrigações financeiras e contábeis devido a sua natureza jurídica, permitindo assim ser um bom instrumento para análise tanto econômica quanto estatística. Quando a empresa possui unidades em mais de uma Unidade da Federação, ela é analisada por unidade local, ou seja, por estabelecimento físico.

Sendo assim, o âmbito da PAC são todas as empresas que, simultaneamente:

- possuam registro no Cadastro Nacional de Pessoa Jurídica (CNPJ), no Ministério da Fazenda;
- estejam sediadas em território nacional;
- especificamente para os estados da Região Norte, estejam sediadas nas capitais; e
- estejam classificadas na seção G da CNAE 2.0 que é a seção onde se encontram as atividades econômicas relacionadas ao comércio.

Daqui em diante, o termo CNAE será sempre referente à CNAE 2.0 e tratar-se-á do nível de classe (quatro dígitos da CNAE), conforme apresentado na Tabela 2.1 e exemplificado na Tabela 2.2, em que se encontra a hierarquia das CNAE 4511 e 4512 a três dígitos (nível grupo), dois dígitos (nível divisão) e um dígito (nível seção). No momento em que forem abordados outros níveis da CNAE, tal alteração será mencionada. A seção G da CNAE 2.0 está organizada em três divisões (dois dígitos), 21 grupos (três dígitos) e 94 classes (quatro dígitos).

Tabela 2.2: *Exemplo da hierarquia das CNAE 4511 e 4512*

Nível	Identificação	Quantidade de dígitos da CNAE	Nome do agrupamento da atividade econômica
Seção:	G	1 dígito	Comércio; reparação de veículos automotores e motocicletas
Divisão:	45	2 dígitos	Comércio e reparação de veículos automotores e motocicletas
Grupo:	451	3 dígitos	Comércio de veículos automotores
Classe:	4511-1	4 dígitos	Comércio a varejo e por atacado de veículos automotores
	4512-9	4 dígitos	Representantes comerciais e agentes do comércio de veículos automotores

2.4 Conceituação das variáveis investigadas

No relatório metodológico da PAC (IBGE [2000]), pode-se encontrar a conceituação de todas as variáveis utilizadas na pesquisa. A seguir, serão apresentadas apenas as variáveis abordadas no presente trabalho.

- **Pessoal Ocupado (PO)** - número de pessoas efetivamente ocupadas em 31 de dezembro do ano de referência, independentemente de terem ou não vínculo empregatício, desde que tenham sido remuneradas diretamente pela empresa;
- **Pessoal Assalariado (PO^A)** - número de pessoas que recebem por meio de salário da empresa efetivamente ocupadas em 31 de dezembro do ano de referência;
- **Pessoal Não-Assalariado (PO^{NA})** - número de pessoas que não recebem por meio de salário efetivamente ocupadas em 31 de dezembro do ano de referência, como diretores e presidentes;
- **Salário, Retiradas e Outras remunerações (SAL)** - valores pagos a título de salários fixos, retiradas pró-labore, honorários da diretoria, comissões sobre vendas, horas extras, participações nos lucros, ajudas de custo, 13º salários, abono financeiro de 1/3 de férias etc, não deduzidos dos correspondentes valores de INSS. Não estão incluídos pagamentos relacionados às diárias de viagens e indenizações derivadas de dispensa;
- **Salário ($SALARIO$)** - valores pagos a título somente de salários fixos e outras remunerações pagas aos empregados da empresa;
- **Receita Bruta ($RECEITA$)** - receitas brutas provenientes do exercício das atividades da empresa (tanto a principal, que a caracteriza ser do ramo de Comércio, como as secundárias) sem deduções dos impostos devidos, das vendas canceladas e dos abatimentos;
- **Revenda ou Comissão sobre venda (REC^1)** - receitas provenientes da revenda de mercadorias ou da comissão sobre vendas de representação comercial e agentes de comércio, o que for maior; e
- **Revenda ou Comissão sobre venda (REC^2)** - receitas provenientes da revenda de mercadorias e da comissão sobre vendas de representação comercial e agentes de comércio.

2.5 Questionário da PAC

Para coletar as informações necessárias à pesquisa, a PAC utiliza dois questionários: um simplificado, que é aplicado às empresas com $PO < 20$; e outro mais completo, com um nível maior de detalhamento das informações investigadas para as empresas com $PO \geq 20$ ao qual é feito censo, denominado como *estrato certo*.

É feita essa distinção uma vez que as empresas do estrato certo possuem suas informações contábeis e financeiras de forma mais ampla, detalhada e organizada.

Na Tabela 2.3, são apresentadas as relações das variáveis de interesse com os questionários da PAC, de forma que $sABC$ corresponde à pergunta de número ABC do questionário simplificado e $cDEF$ à pergunta de número DEF do questionário completo da PAC.

Tabela 2.3: Associação das variáveis de interesse ao questionário da PAC

Variável de interesse	Cálculo através do questionário da PAC
PO	$= PO^A + PO^{NA} = s016 + c020$
PO^A	$= s013 + c016 + c017$
PO^{NA}	$= s014 + s015 + c018 + c019$
SAL	$= SALARIO + s027 + c063$
$SALARIO$	$= s026 + c061 + c062$
$RECEITA$	$= s017 + s017a + s018$ $+ c021 + c021A + c022 + c023 + c024 + c025$
REC^1	$= \max\{s017 + c021; s017A + c021A\}$
REC^2	$= s017 + s017A + c021 + c021A$

Os questionário da PAC, tanto o *Simplificado* quanto o *Completo*, podem ser encontrados nos Anexos A e B, respectivamente.

2.6 Plano amostral e cadastro básico de empresas da PAC

O plano amostral utilizado na pesquisa é inicialmente discutido em Góes [1996], o qual indica a divisão das empresas em dois estratos para a seleção, sendo censo para as grandes empresas que são consideradas as mais representativas de suas atividades econômicas e as demais empresas por amostragem probabilística, de forma que a parte censitária é a maior possível. Esse corte é determinado pelo número de PO das empresas. Definiu-se que as empresas que não fizeram parte do estrato censitário serão selecionadas através de Amostragem Aleatória Estratificada Simples.

Os estratos da PAC são formados por dois tipos: naturais e finais – sendo Estratos Naturais (ESTN) a combinação de UF e CNAE; e os Estratos Finais (ESTF) o porte da empresa determinado

por faixas de *PO*.

Exemplo, o ESTN **354711** representa as empresas do estado de São Paulo de *Comércio a varejo e por atacado de veículos automotores*, sendo os dois primeiros códigos referentes à UF (35 é o código para São Paulo - a relação dos códigos encontra-se na Tabela C.5 apresentados no Anexo C) e os demais dígitos o código da CNAE. Enquanto o estrato **354711-A1** representa as empresas do estado de São Paulo de *Comércio a varejo e por atacado de veículos automotores* com menos de 5 pessoas ocupadas, isto é, do estrato final A1.

Segundo o Relatório metodológico da PAC (IBGE [2017b]), no Sul e Sudeste (exceto Espírito Santo) a maior parte dos estratos naturais são compostos pela UF e pela CNAE a quatro dígitos (nível de classe) - exemplo 354711 e 354712 (empresas do estado de São Paulo de *Representantes comerciais e agentes do comércio de veículos automotores*); enquanto nas demais regiões e Espírito Santo, os estratos naturais são compostos pela UF e pela CNAE a três dígitos (nível de grupo) - 29471 (empresas da Bahia de *Comércio e reparação de veículos automotores e motocicletas*). A lista das atividades econômicas que são descritas a três dígitos no Sul e Sudeste (exceto ES) encontra-se na Tabela C.6 do Anexo C.

Já os estratos finais são obtidos pelo tamanho da empresa e são eles: Certo, Gerencial e Amostrado. A alocação das empresas em cada um desses estratos é dada pelo porte da empresa, que é medido pelo número de *PO* e o número de estados onde atuam, de acordo com o cadastro básico da pesquisa.

Assim, os estratos finais são divididos em Estrato Certo para as Empresas com mais de 20 Pessoas Ocupadas, no qual se realiza o censo das Empresas; Estrato Gerencial que contempla as empresas com *PO* menor de 20 e que atuam em dois ou mais estados; e Estrato Amostrado para as empresas com *PO* menor que 20 e situadas em apenas uma unidade da federação, sendo este último dividido ainda em:

- A1: empresas com menos de cinco de *PO*;
- A2: empresas com *PO* entre 5 e 9; e
- A3: empresas com *PO* entre 10 e 19.

A parte censitária é composta pelos Estratos Certo e Gerencial, enquanto a parte amostral é composta pelo Estrato Amostrado. Aqui será chamado apenas de *Estrato* a combinação de ESTN com ESTF, isto é, o agrupamento final de seleção das empresas.

O cadastro utilizado para seleção da amostra é baseado nas empresas cadastradas no Cadastro Básico de Seleção (CBS) que é uma derivação do Cadastro Central de Empresas (CEMPRE), atualizado pela Coordenação de Metodologia das Estatísticas de Empresas, Cadastros e Classificações (CEEC) e é alimentado, principalmente, pelo Registro Administrativo da Seguridade Geral, conhecido por Relação Anual de Informações Sociais (RAIS), na época de responsabilidade do Ministério do Trabalho.

A RAIS possibilita a obtenção do número de empresas para cada CNAE em cada UF, além de informar o número de Pessoas Ocupadas (PO_{RAIS}) e o valor gasto com Salários para cada empresa ($SALARIO_{RAIS}$), bem como se a empresa declarou ter exercido atividade (empresa ativa) ou não no ano de referência.

Contudo, a base disponível para o IBGE da RAIS é obtida com uma defasagem temporal de um ano em relação ao ano base da pesquisa. Assim, para trabalhar com uma base cadastral mais atualizada, utilizam-se as informações da RAIS combinadas às do Cadastro Geral de Empregados e Desempregados (CAGED) que possui as informações de entradas e saídas de funcionários de todas as empresas nacionais.

Caso a empresa tenha respondido à pesquisa no ano anterior, essa informação também é levada em consideração na montagem do CEMPRE, sendo esta prioritária em sua formulação.

Há uma pesquisa paralela à PAC somente para atualização do CEMPRE chamada Sistema de Manutenção Cadastral (SIMCAD), em que é feita uma amostra de empresas não selecionadas na PAC para atualizar informações básicas de cadastro das empresas.

Dessa forma, utilizando-se da RAIS, do CAGED, do SIMCAD e das informações obtidas nas pesquisas anteriores (para o caso de empresas já selecionadas na pesquisa) mantém-se atualizado o CEMPRE, que serve como cadastro base para a PAC.

Na Figura 2.1, segue uma representação gráfica da atualização do CEMPRE e criação do CBS.

Figura 2.1: Construção do CBS

A partir do CBS, são selecionadas as empresas que responderão as quatro pesquisas estruturais por empresas do IBGE: Pesquisa Anual do Comércio, Pesquisa Anual do Serviço (PAS), Pesquisa Anual da Indústria (PIA) e Pesquisa Anual da Indústria da Construção (PAIC).

2.7 Tamanho da amostra

O tamanho da amostra da PAC para os estratos amostrados é calculado de forma a assegurar que o Coeficiente de Variação (CV) do Total de PO seja menor que 10% para cada ESTN. Assim, seja $n_{A;k}$ o tamanho da amostra para o estrato amostrado A do ESTN k , então obtém-se $n_{A;k}$ da seguinte forma:

$$n_{A;k} = \frac{N_{A;k}^2 \sum_{i=1}^3 W_{A_i;k} S_{A_i;k}^2}{CV^2 (PO_{RAIS;C;k} + PO_{RAIS;A;k})^2 + N_{A;k} \sum_{i=1}^3 W_{A_i;k} S_{A_i;k}^2}, \quad (2.1)$$

em que:

$PO_{RAIS;C;k}$ é o total populacional da RAIS de PO do Estrato Certo C do ESTN k ;

$PO_{RAIS;A;k}$ é o total populacional da RAIS de PO do Estrato Amostrado A do ESTN k ;

$S_{A_i;k}^2$ é a variância populacional da RAIS do PO do Estrato Amostrado A_i do ESTN k ;

$N_{A_i;k}$ é o número de Empresas do Estrato Amostrado A do ESTN k ;

$W_{A_i;k} = \frac{A_i;k}{N_{A_i;k}}$ é a proporção populacional de empresas do do estrato amostrado A_i em relação total do estrato A ;

CV é o coeficiente de variação pré-fixado para o estimador do total de PO em cada ESTN.

Note que $PO_{RAIS;C;k} + PO_{RAIS;A;k} = PO_{RAIS;k}$ é o total de PO da RAIS do ESTN k .

E seja $n_{A_i;k}$ o tamanho da amostra do ESTF A_i para o ESTN k . Então temos que:

$$n_{A_i;k} = \max \left\{ \left\lceil \frac{n_{A_i;k} N_{A_i;k} S_{A_i;k}}{\sum_{j=1}^3 N_{A_j;k} S_{A_j;k}} \right\rceil ; 5 \right\}, \quad (2.2)$$

em que:

$\lceil . \rceil$ é a função teto de . ;

$N_{A_i;k}$ é o tamanho populacional do ESTF A_i para o ESTN k ; e

$S_{A_i;k}$ é o desvio-padrão do ESTF A_i para o ESTN k .

2.8 Tratamento da amostra

Uma parte do plano de crítica da PAC inclui acompanhar, analisar e tratar as empresas que venham a não responder a pesquisa. Essas empresas são então codificadas de acordo com o seu respectivo tipo de não-resposta (motivo pelo qual não foi possível coletar suas informações a respeito da pesquisa) e dependendo do tipo de não-resposta, são aplicados tratamentos diferentes aos dados.

Para cada situação associada à coleta dos questionários, aplica-se uma Situação Cadastral (SITCAD). Quando esta SITCAD estiver relacionada a um questionário não coletado, então aplica-se também uma Folha de Atualização Cadastral (FAC).

Na Tabela 2.4, de acordo com a Nota Técnica da PAC (IBGE [2017b]), segue a lista de Situações Cadastrais das empresas.

Existem quatro tipos diferentes de tratamentos às empresas da pesquisa. São eles:

- expansão com os valores informados/criticados - consiste em aplicar os respectivos pesos das empresas às suas respostas. Esses tratamento é aplicado às empresas que respondem à pesquisa. Essas empresas são chamadas de *Empresas Ativas*. São as empresas com SITCAD 01, 03 e 04;

Tabela 2.4: *Lista de SITCAD e FAC*

Código	Descrição
<i>Aplicadas aos questionários</i>	
01	Em operação
03	Paralisada com informação de receita
04	Extinta com informação de receita
<i>Aplicadas às empresas sem questionários</i>	
02	Em implantação
05	Paralisada sem informação
06	Extinta sem informação
07	Extinta até dezembro do ano decorrente de fusão ou incorporação
08	Atividade fora do âmbito da pesquisa
09	Mudança para endereço ignorado
10	Endereço inexistente ou incompleto
11	Impossibilitada de prestar informações
15	Empresa fora do âmbito geográfico da Região Norte
17	Empresa constituída juridicamente, porém nunca funcionou
20	Recusa a prestar informação

- expansão com valores zerados - consiste em aplicar os respectivos pesos das empresas como ZERO, isto é equivalente a considerar todas as variáveis do questionário com o valor de ZERO e aplicar a expansão anterior. Isso acontece às empresas que forem identificadas como fechadas no ano de referência da pesquisa. A essas empresas é dado o nome de *Empresas Mortas*. São também conhecidas como *FAC de Morte*. São as empresas de SITCAD 02, 05, 06 e 07;
- retirada da amostra e redistribuição dos pesos - aplicada às empresas que não são encontradas, que se recusem a responder à pesquisa (mesmo que obrigadas), que tenham seu questionário julgado inconsistente ou que não possam responder por motivo de força maior. Nesse caso, os pesos dessas empresas são redistribuídos para as demais empresas dos grupos anteriores que sejam do mesmo Estrato. Essas empresas também são chamadas de *FAC de Reponderação*. São as empresas de SITCAD 09, 10, 11, 17 e 20;
- retirada do universo e da amostra - esse tratamento é aplicado ao caso das empresas que estão fora do âmbito da pesquisa. Ela é então retirada do universo da pesquisa e o seu peso original menos um é redistribuído para as empresas dos dois primeiros itens. São as empresas de SITCAD 08 e 15.

As empresas que forem identificadas como fechadas serão tratadas como FAC de Morte. Tal tratamento é decorrente do fato do IBGE receber a RAIS defasada em um ano em relação ao ano da pesquisa em questão. Por exemplo, na PAC 2015, as empresas foram selecionadas a partir da RAIS do ano de 2014, que fora atualizada para 2015 através do CAGED desse ano, sendo a coleta

das informações realizada em 2016. Em outras palavras, esse tratamento estaria estimando o número de mortes das empresas nesse lapso temporal.

Logo, essas distâncias temporais contribuem para que algumas empresas, na data da coleta, não estejam mais em funcionamento ou tenham mudado de endereço (e não atualizaram a RAIS). Essas ocorrências explicam tanto o número significativo de FAC 05 quanto FAC 09, conforme notação da Tabela 2.4. Assim, é fácil perceber a preocupação do IBGE em encontrar as formas mais adequadas possíveis para tratar tais dados ausentes, uma vez que tais ocorrências crescem a cada ano nas empresas, levando seus respectivos pesos amostrais consigo. Esse fenômeno é mais frequente nos estratos finais A1, portanto existe uma preocupação maior com ele.

Portanto, resumidamente, existem dois tratamentos aplicados às empresas no momento da expansão dos respectivos pesos amostrais das empresas, conforme apresentado na Figura 2.2.

Figura 2.2: Tratamento aplicado às empresas no momento da expansão dos pesos amostrais associados

2.9 Cálculo das estimativas

Na PAC, são utilizados dois tipos de estimadores para se obter os totais de interesse quando a amostra é expandida: um estimador regressão e um simples.

Usa-se o estimador regressão para garantir que os totais estimados de *PO* e *SALARIO* sejam iguais aos respectivos totais do Registro Administrativo utilizado para selecionar as empresas da amostra (calibração). Para tal, o estimador regressão usa as variáveis *PO* e *SALARIO* do CEMPRE como variáveis auxiliares.

O estimador simples é utilizado em apenas duas situações: quando o número de empresas respondentes no estrato for menor que cinco ou quando os pesos resultantes da calibração apresentaram

valores negativos, segundo o Relatório Metodológico da PAC (IBGE [2000]).

2.9.1 Estimador simples

Seja Y o total de uma variável de interesse, então o estimador simples para Y para um determinado domínio D é obtido da seguinte forma:

$$\hat{Y}^D = \sum_{h=1}^H \sum_{i=1}^{n_h} w_{hi} y_{hi} \mathbb{I}_D(u_{hi}), \quad (2.3)$$

em que:

H é o número de estratos em que a população foi dividida;

n_h é o número de empresas respondentes na amostra pertencentes ao estrato h ;

u_{hi} é a unidade investigada i do estrato h ;

$w_{hi} = \frac{N_h}{n_h}$ é o peso atribuído à u_{hi} , sendo N_h é o tamanho populacional do estrato h ;

$\mathbb{I}_D(u_{hi})$ a função indicadora do domínio D da unidade u_{hi} ;

y_{hi} é o valor da variável da pesquisa y para a unidade u_{hi} .

Um estimador da variância do estimador simples para o total da variável Y no domínio D dado pela expressão (2.3) é:

$$v(\hat{Y}^D) = \sum_{h=1}^H N_h^2 \frac{1 - f_h}{n_h} s_{hD}^2, \quad (2.4)$$

em que:

$f_h = \frac{n_h}{N_h}$ é a fração amostral final de h ;

$s_{hD}^2 = \frac{\sum_{i=1}^{n_h} (\alpha_{hi} - \bar{\alpha}_h)^2}{n_h - 1}$ é o estimador da variância de α_{hi} no estrato h ;

$\alpha_{hi} = y_{hi} \mathbb{I}_D(u_{hi}), i = 1, \dots, N_h$

$\bar{\alpha}_h = \frac{\sum_{i=1}^{n_h} \alpha_{hi}}{n_h}$ é o estimador de média de α_{hi} no estrato h .

Portanto, o estimador do CV do estimador simples dado em (2.3) para a variável Y no domínio D pode ser obtido por:

$$CV(\hat{Y}^D) = 100 * \frac{\sqrt{v(\hat{Y}^D)}}{\hat{Y}^D}. \quad (2.5)$$

2.9.2 Estimador regressão

O estimador regressão para o Total da variável Y no domínio D é dado por:

$$\hat{Y}_{Reg;h}^D = \sum_{i=1}^{n_h} w_{Reg;h} y_{hi} \mathbb{I}_D(u_{hi}), \quad (2.6)$$

em que:

$w_{Reg;h} = w_{hi}g_{hi}$ é o peso final da unidade u_{hi} ;

$g_{hi} = 1 + (\mathbf{C} - \hat{\mathbf{C}}) \left\{ \sum_{h=1}^2 \sum_{i=1}^{n_h} w_{hi} \mathbf{c}_{hi} \mathbf{c}_{hi}' \right\}^{-1} \mathbf{c}_{hi}$ é o fator de calibração associado à unidade u_{hi} ;

$\mathbf{C} = (C_1, \dots, C_J)'$ é um vetor de dimensão $J \times 1$ com os Totais Populacionais das explicativas variáveis c_j , tal que $j = 1, \dots, J$;

$\hat{\mathbf{C}} = (\hat{C}_1, \dots, \hat{C}_J)'$ é um vetor de dimensão $J \times 1$ com os Estimadores dos Totais Populacionais das variáveis explicativas $c_j, j = 1, \dots, J$;

$\mathbf{c}_{hi}$ é o vetor de dimensão $J \times 1$ de valores das variáveis explicativas para a unidade u_{hi} .

A variância desse estimador pode ser estimada por:

$$v(Y_{Reg;h}^{\hat{D}}) = \sum_{h=1}^H N_h^2 \frac{1 - f_h}{n_h} s_{Reg;h_D}^2, \quad (2.7)$$

em que:

$s_{Reg;h_D}^2 = \frac{\sum_{i=1}^{n_h} (m_{hi} - \bar{m}_h)^2}{n_h - 1}$ é o estimador da variância de m_{hi} no estrato h ;

$m_{hi} = g_{hi} d_{hi} \mathbb{I}_D(u_{hi}), i = 1, \dots, N_h$

$\bar{m}_h = \frac{\sum_{i=1}^{n_h} m_{hi}}{n_h}$ é o estimador de média de m_{hi} no estrato h .

$\hat{d}_{hi} = \alpha_{hi} - \mathbf{c}_{hi}' \hat{\mathbf{B}}_h$ é o resíduo estimado para a unidade u_{hi} ;

$\hat{\mathbf{B}}_h = (\hat{\mathbf{B}}_1, \dots, \hat{\mathbf{B}}_J) = \left(\sum_{i=1}^{n_h} w_{hi} \mathbf{c}_{hi} \mathbf{c}_{hi}' \right)^{-1} \left(\sum_{i=1}^{n_h} w_{hi} \mathbf{c}_{hi} \mathbf{y}_{hi} \right)$ é um vetor de dimensão $J \times 1$ composto pelos estimadores dos coeficientes de regressão apresentados em [Särndal e Lundström \[2005\]](#).

E temos, para o estimador regressão, o coeficiente de variação podendo ser estimado por:

$$CV(\hat{Y}_{Reg;h}^D) = 100 * \frac{\sqrt{v(\hat{Y}_{Reg;h}^D)}}{\hat{Y}_{Reg;h}^D}. \quad (2.8)$$

2.10 Processo de crítica e imputação

A crítica e imputação dos dados de uma pesquisa é um dos seus processos mais importantes, pois é nele que se assegura a qualidade dos dados que foram coletados. Segundo [Snijkers et al. \[2013\]](#) é no processo de crítica que se analisam as relações das variáveis para cada questionário coletado, incluindo a coerência das estimativas finais quando analisadas entre si, em conjunto com outras fontes de dados e com relação à própria série histórica da pesquisa. Ainda segundo [Snijkers et al. \[2013\]](#), o processo de crítica e imputação de uma pesquisa divide-se em duas etapas: *crítica do microdado* e *macrocrítica*.

O processo de crítica de microdado consiste em olhar a unidade de investigação da pesquisa,

que no caso da PAC é a empresa como um todo, ou seja, avaliar o preenchimento e a coerência de cada questionário. Já o processo de macrocrítica consiste em analisar os valores estimados de forma agregada após a expansão da pesquisa e verificar possíveis inconsistências de resultados e, quando houver, revisitar a etapa de crítica do microdado do domínio em questão. Portanto, o processo de crítica é cíclico, sendo necessário revisitar diversas vezes cada etapa para certificar a qualidade dos dados.

Ao final de cada rodada do processo de crítica é natural que haja valores faltantes ou inconsistentes, as *não-respostas*, e no capítulo a seguir serão apresentados os principais tratamentos para esses casos.

2.11 Disseminação dos resultados

Segundo o IBGE [2018], os resultados da PAC são divulgados em 23 tabelas divulgados apenas no portal do IBGE na internet seguindo o seguinte formato:

- Resultados do total das empresas comerciais no Brasil para o ano de referência e tabelas de confronto entre o ano de referência e o ano anterior para todas as empresas comerciais segundo divisões, grupos e classes de atividade. Além disso, são apresentados também dados gerais das empresas comerciais, segundo as Grandes Regiões, Unidades da Federação de atuação das empresas, por divisões e grupos de atividades; e
- Resultados das empresas com 20 ou mais pessoas ocupadas, segundo divisões, grupos e classes de atividade.

O plano tabular da PAC para efeito de divulgação dos resultados segue o agrupamento descrito nas Tabelas 2.5, 2.6 e 2.7 para as divisões de CNAE 45, 46 e 47, respectivamente. Cada linha apresentada nessas tabelas são denominadas de *linha de divulgação* e são compostas ou por uma única classe de CNAE ou por grupo de classes de CNAE, conforme segue nas tabelas supracitadas.

Tabela 2.5: CNAE 45 - Comércio de veículos, peças e motocicletas

Denominação	Códigos CNAE 2.0
Veículos automotores	45.11-1 e 45.12-9
Peças para veículos	45.30-7
Motocicletas, peças e acessórios	45.41-2 e 45.42-1

Tabela 2.6: *CNAE 46 - Comércio por atacado*

Denominação	Códigos CNAE 2.0
Representantes e agentes do comércio (exceto de veículos e motocicletas)	46.11-7, 46.12-5, 46.13-3, 46.14-1, 46.15-0, 46.16-8, 46.17-6, 46.18-4, 46.19-2
Comércio de matérias-primas agrícolas e animais vivos	46.21-4, 46.22-2, 46.23-1
Comércio especializado em produtos alimentícios, bebidas e fumo	46.31-1, 46.32-0, 46.33-8, 46.34-6, 46.35-4, 46.36-2, 46.37-1, 46.39-7
Tecidos, artefatos de tecidos e de armarinho	46.41-9
Artigos do vestuário e acessórios	46.42-7
Calçados e artigos de viagem	46.43-5
Produtos farmacêuticos, perfumaria, cosméticos, médicos, ortopédicos, odontológicos e veterinários	46.44-3, 46.45-1, 46.46-0
Artigos de escritório e de papelaria; papel, papelão e seus artefatos; livros, jornais e outras publicações	46.47-8
Outros equipamentos e artigos de uso pessoal e doméstico não especificados anteriormente	46.49-4
Combustíveis e lubrificantes	46.81-8 e 46.82-6
Madeira, ferragens, ferramentas, material elétrico e material de construção	46.71-1, 46.72-9, 46.73-7, 46.74-5, 46.79-6
Produtos químicos, adubos e fertilizantes	46.83-4 e 46.84-2
Produtos siderúrgicos e metalúrgicos	46.85-1
Papel e papelão em bruto e de embalagens	46.86-9
Resíduos e sucatas	46.87-7
Outros produtos	46.89-3
Comércio de equipamentos e produtos de tecnologia de informação e comunicação	46.51-6 e 46.52-4
Comércio de máquinas, aparelhos e equipamentos (exceto de tecnologia de informação e comunicação)	46.61-3, 46.62-1, 46.63-0, 46.64-8, 46.65-6, 46.69-9
Comércio não especializado	46.91-5, 46.92-3, 46.93-1

Tabela 2.7: CNAE 47 - Comércio varejista

Denominação	Códigos CNAE 2.0
Hipermercados e supermercados	47.11-3
Outros tipos de comércio não especializado com predominância de produtos alimentícios	47.12-1
Comércio não especializado sem predominância de produtos alimentícios	47.13-0
Comércio de produtos alimentícios, bebidas e fumo	47.21-1, 47.22-9, 47.23-7, 47.24-5, 47.29-6
Tecidos e artigos de armarinho	47.55-5
Artigos do vestuário e complementos	47.81-4
Calçados, artigos de couro e viagem	47.82-2
Combustíveis e lubrificantes	47.31-8 e 47.32-6
Produtos farmacêuticos, perfumaria e cosmético e artigos médicos, ortopédicos e de óptica	47.71-7, 47.72-5, 47.73-3, 47.74-1
Eletrodomésticos, equipamentos de áudio e vídeo, instrumentos musicais e acessórios	47.53-9 e 47.56-3
Móveis, artigos de iluminação, peças e acessórios e outros artigos de uso doméstico	47.54-7, 47.57-1, 47.59-8
Material de construção	47.41-5, 47.42-3, 47.43-1, 47.44-0
Equipamentos de informática e comunicação	47.51-2 e 47.52-1
Artigos culturais, recreativos e esportivos	47.61-0, 47.62-8, 47.63-6
Gás liquefeito de petróleo (GLP)	47.84-9
Outros produtos novos	47.83-1 e 47.89-0
Comércio de artigos usados	47.85-7

Capítulo 3

Imputação e outras formas de lidar com a não-resposta

3.1 Conceitos gerais

Segundo [Albieri \[1989\]](#), existem dois tipos de dados ausentes: o primeiro é aquele no qual não se tem nenhuma informação do questionário, chamado de “questionário ausente” ou “não-resposta das unidades”; e o segundo é referente à falta da informação de um item do questionário, contudo as demais encontram-se respondidas, chamado de *item perdido* ou *não-resposta ao item*.

A maioria dos casos de *questionários ausentes* são decorrentes de *Recusa*, *Impossibilidade de resposta da empresa por força maior* e a *Não localização da empresa*. De acordo com [Kalton e Kasprzyk \[1982\]](#), além da *não-resposta total* e da *não-resposta ao item*, um outro tipo de dado ausente em pesquisas por amostras pode também ocorrer pela *falta de cobertura* de toda a população. No caso das Pesquisas Estruturais por Empresas do IBGE, sabe-se que existe *falta de cobertura* das empresas nascidas no ano de referência das pesquisas, decorrente do recebimento tardio da RAIS, contudo o tema não será aprofundado.

O presente trabalho considerará somente a variável *RECEITA* para imputação e para essa questão assumiremos o caso de *questionários ausentes*, ou seja, a empresa não respondeu a nenhum item do questionário. Entretanto, os valores de *PO* e *SALARIO* da RAIS serão utilizados como variáveis explicativas.

Para evitar o acontecimento da não-resposta, [Lohr \[2010\]](#) sugere investigar detalhadamente os motivos que movem uma empresa a omitir sua resposta para tentar reverter esse quadro, ou pelo menos, mitigar o seu acontecimento, pois não há estimação que seja melhor do que o valor real do

informante.

Lohr [2010] discute diversas formas de lidar com a *não-resposta*, elegendo como a melhor forma de evitar a *não-resposta*: mitigar o seu acontecimento. A sua implementação consiste em investigar os motivos que ocasionam as não-respostas e, a partir dessa informação, criar um plano de ação para combater, ou pelo menos reduzir, suas ocorrências. Para que seja possível fazer o levantamento dessas informações, aconselha-se selecionar uma sub-amostra de não-respondentes e investigar o motivo para tal. Para isso, melhor seria retornar à etapa de coleta e entrar novamente em contato com os informantes, ou ainda, tentar acrescentar essa informação no momento da primeira coleta. Entretanto, como as bases utilizadas para nossas estimativas já passaram pelas etapas de coleta e crítica, esse procedimento não poderá ser abordado.

Além da prevenção, Lohr [2010] aborda três outras formas de lidar com as *não-respostas*:

1. Ignorar a não-resposta;
2. Usar sub-amostras de não-respondentes para fazer inferência sobre os mesmos; e
3. Usar um modelo para prever os valores dos não-respondentes, ajuste dos pesos amostrais em cada classe, métodos de imputação e modelos paramétricos.

A primeira proposta indicada fala sobre ignorar a não-resposta. O problema quanto a isso é o viés implícito à não-resposta. Acredita-se que empresas que se recusam a responder têm algum motivo para tal e, sendo assim, ao desconsiderarmos esses respondentes poderemos aumentar significativamente o vício das estimativas.

Há ainda pesquisas que excluem questionários incompletos e analisam apenas questionários completos. Isso aumenta ainda mais o viés da não-resposta, pois, seguindo a lógica exposta anteriormente, estaríamos sub-representando a população na amostra, sendo, portanto, não recomendado a adoção dessa opção.

A segunda proposta sugere que as respostas sejam divididas em dois estratos: respondentes iniciais (aqueles que responderam a pesquisa na primeira oportunidade) e não-respondentes iniciais (aqueles que só responderam a pesquisa após uma nova visita à localidade). Lohr [2010] chama esse método de *two-phase sampling* ou *double sampling*. Entretanto, como a pesquisa não diferencia respondentes iniciais de outros respondentes, não temos como aplicar tais procedimentos nos moldes atuais da pesquisa.

Dito isso, o presente trabalho se concentrará no terceiro item supracitado e, portanto, serão abordados somente modelos que predizem os valores dos não-respondentes. Mais especificamente,

nos atentaremos aos **métodos de imputação**.

Segundo Little e Rubin [2002] *Apud* Lohr [2010], existem três tipos de não-respostas: *Missing Completely at Random* - *MCAR*; *Missing at Random (Given Covariates)* - *MAR*; e *Not Missing at Random* - *NMAR*.

Para poder descrever cada tipo, primeiro serão definidas algumas variáveis.

Seja: Z_i a variável indicadora da unidade i estar na amostra; e R_i a variável indicadora da unidade i responder dado que está na amostra, tal que:

$$P(Z_i = 1) = \pi_i \text{ e } P(R_i = 1) = \Phi_i. \quad (3.1)$$

Então, $P(\text{unidade}_i \text{ ser selecionada e responder}) = \pi_i \Phi_i$. Assim, o peso final de um respondente é $(\pi_i \Phi_i)^{-1}$.

Seja também Y_i a resposta de interesse e X_i o vetor de variáveis dependentes conhecido sobre a unidade i da amostra, temos que:

- *MCAR* – se Φ_i não depende de X_i , Y_i ou do desenho da amostra, ou de outra forma, se as probabilidades Φ_i são iguais para todas as unidades i e os eventos $\{R_i = 1\}$ são condicionalmente independentes entre si e do processo de seleção da amostra dado n_R , então os dados são *MCAR*;
- *MAR* – se Φ_i depende de X_i , mas não de Y_i então os dados são *MAR*. Nesse caso, os dados dependem apenas das variáveis observadas e poderiam ser facilmente modelados com base nas variáveis conhecidas; e
- *NMAR* – se Φ_i depender dos valores não respondidos das variáveis e não puder ser completamente explicado pelos valores observados, então os dados são *NMAR*. Nesse caso, modelos podem ser utilizados para prever os valores ausentes, pois eles podem depender também dos valores conhecidos, contudo esse modelo nunca será totalmente suficiente.

Segundo Lohr [2010], as probabilidades de respostas Φ_i são usuais para descrever o tipo de não-resposta, entretanto como são desconhecidas, não se pode saber com certeza qual o tipo presente. Para distinguir MAR de MCAR, pode-se criar um modelo a fim de prever as probabilidades observadas de respostas para cada sub-grupo, se os coeficientes de uma regressão logística predizendo as não-respostas forem significativamente diferente de 0 (zero), então os dados ausentes provavelmente não são MCAR. Já para distinguir MAR de NMAR é mais complicado, porém é mais razoável esperar mais respostas do tipo NMAR do que MAR.

De acordo com Lohr [2010], os métodos de reponderação (de ajustes da ponderação) assumem que as respostas desconhecidas podem ser estimadas através dos valores conhecidos para todas as unidades, logo assumem que são do tipo MAR.

Särndal e Lundström [2005] dizem que os dois principais métodos indicados na literatura para os casos de tratamento de não-resposta são a *Reponderação dos pesos* e a *Imputação*, sendo que Imputação ainda é um método muito controverso dentro da Estatística.

Segundo Silva [2017], quando os padrões de não-resposta são descritos como modelos do tipo MCAR ou MAR, admite-se o uso de imputação, entretanto quando os modelos são do tipo NMAR, é muito difícil, ou até mesmo impossível, utilizar algum método de imputação nos dados.

3.2 Imputação

Kalton e Kasprzyk [1982] descrevem diversos métodos de imputação classificados por eles a partir de duas características básicas: com uso ou não uso de variáveis auxiliares; e pelo tipo de erro atribuído ao modelo. A seguir destacamos os dez métodos por eles apresentados, porém utilizando a nomenclatura de Albieri [1989], que originalmente também utilizou os mesmos métodos traduzindo-os para o português.

Vale ressaltar que as escolhas das variáveis auxiliares do modelo são de grande importância. Segundo Kalton e Kasprzyk [1982], tal escolha deve ser feita pela capacidade das variáveis auxiliares em conseguir prever a variável de interesse.

Os métodos podem ser classificados em dois tipos:

- Determinístico – quando os resultados da imputação são únicos para o mesmo banco de dados, ou seja, mesmo repetindo o procedimento de imputação várias vezes, ainda assim seriam obtidos os mesmos resultados; e
- Estocástico – quando são acrescentados alguns tipos de erro aos resultados de forma que, caso repetíssemos o mesmo procedimento, seriam obtidos resultados diferentes.

Os métodos de imputação descritos em Kalton e Kasprzyk [1982] e Albieri [1989] são os seguintes:

1. Imputação dedutiva;
2. Imputação da média geral;
3. Imputação aleatória geral;

4. Imputação da média dentro de classes;
5. Imputação aleatória dentro de classes;
6. Imputação da média de regressão;
7. Imputação da média de regressão com efeito aleatório;
8. Imputação *Hot Deck* sequencial;
9. Imputação por associação flexível;
10. Imputação por função distância.

Antes de detalhar cada método citado acima, vamos apresentar a notação que será adotada para a apresentação desses métodos.

3.2.1 Notação e considerações gerais

Seja $\mathcal{P}$ uma população finita de N unidades compostas de duas subpopulações: respondentes ($\mathcal{R}$) e não-respondentes ($\mathcal{M}$), tal que $\mathcal{R} \cup \mathcal{M} = \mathcal{P}$.

Sejam R e M os números de unidades nas subpopulações de respondentes ($\mathcal{R}$) e não-respondentes ($\mathcal{M}$), respectivamente. Então, temos que $R + M = N$. Suponhamos que R e M sejam grandes.

Seja A_n uma amostra aleatória simples, sem reposição, de tamanho n retirada de $\mathcal{P}$ e um conjunto de variáveis é observado: Y , X e $\mathbf{Z} = \{Z_1, Z_2, \dots, Z_f\}$.

Os dados para X e $\mathbf{Z}$ são observados para todas as unidades da amostra, enquanto os dados da variável Y são obtidos para r unidades, $1 < r < n$. Assim, temos $m = n - r$ unidades com valores ausentes para Y .

Dessa forma, podemos decompor a amostra A_n da seguinte forma: $A_n = A_m \cup A_r$, tal que A_r é a amostra que responderam todas as variáveis de Y , X e $\mathbf{Z}$ e A_m a amostra que respondeu a X e $\mathbf{Z}$, mas não a Y .

Com base nas suposições acima, podemos definir:

$$\begin{aligned}\bar{R} &= \frac{R}{N} \\ \bar{M} &= \frac{M}{N} \\ \bar{R} + \bar{M} &= 1 \\ \bar{Y} &= \sum_{i=1}^N \frac{y_i}{N}\end{aligned}$$

$$\begin{aligned}
\bar{Y}_R &= \sum_{i=1}^R \frac{y_{ri}}{R} \\
\bar{Y}_M &= \sum_{i=1}^M \frac{y_{mi}}{M} \\
\bar{X} &= \sum_{i=1}^N \frac{x_i}{N} \\
\bar{X}_R &= \sum_{i=1}^R \frac{x_{ri}}{R} \\
\bar{X}_M &= \sum_{i=1}^M \frac{x_{mi}}{M} \\
S_Y^2 &= \sum_{i=1}^N \frac{(y_i - \bar{y})^2}{N-1} \\
S_{RY}^2 &= \sum_{i=1}^R \frac{(y_{ri} - \bar{y}_r)^2}{R-1} \\
S_{MY}^2 &= \sum_{i=1}^M \frac{(y_{mi} - \bar{y}_m)^2}{M-1} \\
S_{XY} &= \sum_{i=1}^N \frac{(y_i - \bar{y})(x_i - \bar{x})}{N-1} \\
S_{RXY} &= \sum_{i=1}^R \frac{(y_{ri} - \bar{y}_r)(x_{ri} - \bar{x}_r)}{R-1} \\
S_{MXY} &= \sum_{i=1}^M \frac{(y_{mi} - \bar{y}_m)(x_{mi} - \bar{x}_m)}{M-1}, \\
\bar{r} &= \frac{r}{n} \\
\bar{m} &= \frac{m}{n}
\end{aligned}$$

tal que:

y_i e x_i - valores das variáveis Y e X correspondentes à i -ésima unidade da população, respectivamente;

y_{ri} e x_{ri} - valores das variáveis Y e X correspondentes à i -ésima unidade da subpopulação de respondentes, respectivamente;

y_{mi} e x_{mi} - valores das variáveis Y e X correspondentes à i -ésima unidade da subpopulação de não-respondentes, respectivamente.

A seguir, seguem algumas das relações observadas:

$$\begin{aligned}
\bar{Y} &= \bar{R}\bar{Y}_r + \bar{M}\bar{Y}_m \\
S_y^2 &= \bar{R}S_{ry}^2 + \bar{M}S_{my}^2 + \bar{R}\bar{M}(\bar{Y}_r - \bar{Y}_m)^2 \\
S_{xy}^2 &= \bar{R}S_{rxy}^2 + \bar{M}S_{mxy}^2 + \bar{R}\bar{M}(\bar{Y}_r - \bar{Y}_m)(\bar{X}_r - \bar{X}_m) \\
\bar{y} &= \bar{r}\bar{y}_r + \bar{m}\bar{y}_m \\
s_y^2 &= \bar{r}s_{ry}^2 + \bar{m}s_{my}^2 + \bar{r}\bar{m}(\bar{y}_r - \bar{y}_m)^2 \\
s_{xy}^2 &= \bar{r}s_{rxy}^2 + \bar{m}s_{mxy}^2 + \bar{r}\bar{m}(\bar{y}_r - \bar{y}_m)(\bar{x}_r - \bar{x}_m),
\end{aligned}$$

em que as definições de $\bar{y}_r$, $\bar{y}_m$, s_{ry}^2 , s_{my}^2 , s_{rxy}^2 , s_{mxy}^2 são análogas às dos parâmetros denotados pelas mesmas letras maiúsculas na população, só que para a amostra.

Muitos modelos de imputação usam um conjunto de variáveis auxiliares para ajudar a construir um modelo para estimar as respostas ausentes. Isso é muito frequente, por exemplo, quando lidamos com os casos de *não-resposta ao item* em que costuma haver uma quantidade substancial de informações acerca das outras variáveis investigadas. E, em casos como esse, podemos utilizar toda a base teórica exposta anteriormente.

Dessa forma, podemos escrever o valor imputado como $\hat{y}_{mi} = f(\mathbf{z}) = f(z_{1i}, z_{2i}, \dots, z_{fi})$, tal que $f(\mathbf{z})$ é uma função de variáveis auxiliares.

A função $f(\mathbf{z})$ pode ser, frequentemente, expressa como uma função do tipo linear $\beta_0 + \sum_{j=1}^k \beta_j z_{ji}$ cujos parâmetros poderiam ser estimados em A_m . Assim, temos:

$$y_{ri} = \beta_{r0} + \sum_{j=1}^k \beta_{rj} z_{rji} + e_{ri} \text{ em } A_r$$

e

$$y_{mi} = \beta_{r0} + \sum_{j=1}^k \beta_{rj} z_{mji} + e_{mi} \text{ em } A_m .$$

[Albieri \[1989\]](#) reforça que o fator principal para a escolha das variáveis auxiliares deve ser a capacidade delas em predizer o valor da variável ausente. Ainda segundo [Albieri \[1989\]](#), os métodos de imputação se classificam em duas formas: pelo uso ou não de variáveis auxiliares; e pelas suposições atribuídas ao erro: modelo determinístico ou estocástico.

3.2.2 Principais métodos de imputação

A seguir, a partir de [Albieri \[1989\]](#), será apresentada uma análise sucinta dos principais métodos de imputação utilizados na literatura.

Método por Imputação Dedutiva - MID

Este método depende de alguma informação redundante dos dados de forma que a resposta ausente seja explicada exatamente por uma informação auxiliar, ou seja, $y_{mi} = f(\mathbf{z}_i)$. Um exemplo disso ocorre quando um registro é obtido a partir da soma de uma série de parcelas e um desses registros está ausente. Nesse caso, seria possível obtê-lo pela subtração do total das demais parcelas. Um outro exemplo acontece quando, em uma pesquisa painel, uma variável permanece praticamente inalterada ao longo do tempo e num dado momento encontra-se ausente, sendo possível imputá-la do momento anterior ou posterior ao ausente.

Método por Imputação da Média Geral - MIMG

Este método atribui a média geral dos respondentes ($\bar{y}_r$) para todas as respostas ausentes. É a forma determinística da função linear sem variáveis auxiliares, ou seja, $y_{mi} = \beta_{r0} = \bar{y}_r$.

Método por Imputação Aleatória Geral - MIAG

Este método atribui o valor y de um respondente selecionado aleatoriamente da amostra de respondentes para cada resposta ausente. Este método seria a forma estocástica da função linear sem variáveis auxiliares, $y_{mi} = \bar{y}_r + e_{mi} = y_{ri'}$, para algum respondente i' .

Método por Imputação da Média Dentro de Classes - MIMDC

Este método considera a existência de classes de imputação formadas a partir de variáveis auxiliares e dentro de cada classe a média dos respondentes da variável y é atribuída às respostas ausentes da mesma classe, ou seja, $y_{mci} = \bar{y}_{rh}$ para o i -ésimo não-respondente na classe c . Em geral, o modelo pode ser representado por $y_{mi} = \beta_{r0} + \sum_{j=1}^k \beta_{rc} z_{cmi}$, tal que z_{cmi} é a função indicadora do i -ésimo não-respondente estar na classe c . Este método é um método determinístico.

Método por Imputação Aleatória Dentro de Classes- MIADC

Este método é semelhante ao MIAG, sendo que no presente método, ele é aplicado dentro de classes. Tal método é o estocástico do MIMDC, sendo $y_{mhi} = \bar{y}_{rc} + e_{mci} = y_{rci'}$, para alguma empresa i' pertencente à classe c .

Alternativamente, podemos usar a expressão $y_{mji} = \beta_{r0} + \sum_{j=1}^k \beta_{rj} z_{jmi} + e_{mji}$, onde e_{mji} é selecionado aleatoriamente dentro da classe de respondentes a qual o não-respondente i pertence. Este método também é conhecido como *Hot Deck* Aleatório.

Método por Imputação da Média de Regressão - MIMR

O método em questão usa os dados dos respondentes para regredir y em função das variáveis auxiliares. Dessa forma, os valores ausentes de y são imputados através da equação $y_{mi} = \beta_{r0} + \sum_{j=1}^k \beta_{rj} z_{jmi}$. As variáveis auxiliares podem ser tanto quantitativas quanto qualitativas e transformações nas variáveis e uso de interações podem ser aplicadas se necessário. Este método é um método determinístico.

Método por Imputação da Média de Regressão com Efeito Aleatório - MIMRE

Este método é a versão estocástica do MIMR, isto é, soma-se um resíduo aos valores preditos, ou seja, $y_{mi} = \beta_{r0} + \sum_{j=1}^k \beta_{rj} z_{jmi} + e_{mi}$.

Método por Imputação *Hot Deck* Sequencial - MIHDS

Esse método considera a amostra dividida em classes. Para cada classe, é atribuído um único valor para a variável y com o intuito de dar início ao processo que pode ser o valor de um respondente de cada classe ou o valor médio da classe, obtido numa rodada anterior. A partir daí, os registros da pesquisa são tratados sequencialmente. Se o registro seguinte tiver alguma resposta para a variável y , este passa a ser o novo valor a ser utilizado na imputação, caso contrário será atribuído o valor y previamente selecionado, e assim sucessivamente.

Se a disposição dos registros na base for aleatória, este método se assemelha ao MIADC à exceção do início do processo. Se a ordem do processo criar autocorrelação positiva entre os registros, este método propicia um grau de associação adicional.

Método por Imputação por Associação Flexível - MIAF

Esse é o nome utilizado para o método *Hot Deck* Modificado que consiste em agrupar respondentes e não-respondentes em grandes classes de imputação construídas a partir de um grande conjunto de variáveis auxiliares. A partir disso, os não-respondentes são associados aos respondentes de forma hierárquica e quando um não-respondente não puder ser associado a um respondente na classe inicial, essas classes são, então, agrupadas e a imputação é refeita num nível maior, menos específico, de forma que a associação sempre ocorra. Esse método evita o uso múltiplo de doadores em classes onde o número de respondentes excede o de não-respondentes.

Método por Imputação por Função Distância - MIFD

O método atribui o valor y do respondente *mais próximo* ao não-respondente em questão, também chamado de *vizinho mais próximo*. A escolha desse *vizinho mais próximo* é feita através de uma função de distância usando algumas variáveis auxiliares, podendo ser usadas variáveis quantitativas ou qualitativas.

As variáveis auxiliares podem ser usadas para cálculo da distância entre os registros ou para criação de classes de imputação. As variáveis auxiliares podem ou não sofrer transformações. Quando for o caso do uso de somente uma variável auxiliar, pode-se adotar como distância a diferença em valor absoluto entre o valor do doador e do respondente, sendo o candidato à imputação aquele com distância mínima. Quando várias variáveis auxiliares são utilizadas, tanto a questão das transformações quanto as do cálculo da distância são mais sensíveis. Uma possível solução é transformar todas as variáveis auxiliares em seus postos.

Assim, a função distância seria dada por $DPT_{id} = \sup_v \{W_v | PT_{vi} - PT_{vd} | \}$, tal que PT_{vi} é o posto do não-respondente i e PT_{vd} é o posto do doador em potencial d para a variável v , e W_v é o peso que representa a importância da variável v na função distância.

Outras possíveis alternativas podem ser baseadas na distância de Mahalanobis ou na função distância por $DML_{id} = |z_i - z_d|$ que foi apresentada por [Särndal e Lundström \[2005\]](#).

Método de Imputação Múltipla - MIM

[Little e Rubin \[2002\]](#) também defendem um método de imputação que consiste na combinação de duas ou mais técnicas distintas de imputação, sendo possível, então, combinar dois ou mais modelos descritos anteriormente da maneira que for mais conveniente.

3.2.3 Método de imputação utilizado na PAC

Atualmente, na PAC, o questionário de uma empresa pode ser imputado de duas formas distintas: ou através de seu *histórico* ou através de uma empresa *doadora*.

É feita imputação por histórico quando a empresa tiver respondido à pesquisa nos dois anos anteriores ou posteriores ao ano que se deseja imputar, desde que não pertença ao estrato A1 (empresas de pequeno porte), pois todas as empresas desse estrato são indicadas para imputação por doadora devido ao alto número de mortalidade das empresas de pequeno porte. Caso a empresa não atenda aos critérios de imputação por histórico é feita a imputação pela empresa doadora, conforme diagrama que segue na Figura 3.1.

Figura 3.1: Fluxo de imputação pelo modelo do IBGE

Independentemente do tipo de imputação utilizada, as informações referentes a *PO* e *SALARIO* são obtidas diretamente da RAIS através dos valores informados pelas empresas ao Ministério do Trabalho: PO_{RAIS} e $SALARIO_{RAIS}$; sendo necessária assim, somente a imputação das informações adicionais. Portanto, só podem ser imputadas empresas que tenham respondido à RAIS, tendo respondido haver exercido atividade e informado seus valores de *PO* e *SALARIO* no ano a se imputar.

No caso de imputação pelo histórico, utiliza-se a estrutura da empresa no ano considerado como histórico, isto é, a relação entre as variáveis da empresa como base para calcular um fator de variação de cada variável dentro dos blocos de *PO* e de *SALARIO* com os respectivos valores de *PO* e *SALARIO* do ano em que se tem informação (ano do histórico) e multiplica-se essa variação pelo valor informado de *PO* e de *SALARIO* do ano de referência da imputação obtido pela RAIS, isto é, PO_{RAIS} e $SALARIO_{RAIS}$, respectivo de cada bloco. Esse cálculo é replicado em todo o bloco de *PO* e *SALARIO* do questionário imputado.

Estima-se a $RECEITA^1$ por um fator de variação entre os valores informados de *RECEITA* e $SALARIO_{RAIS}$ do ano histórico, multiplicado pelo valor de $SALARIO_{RAIS}$ do ano de imputação.

¹As definições das variáveis *PO*, PO_t^A , PO^{NA} , *SAL*, *SALARIO*, *RECEITA*, REC^1 e REC^2 encontram-se nas Seções 2.4 e 2.5

Nos demais blocos, essa variação é feita em relação ao valor de REC^1 informado no ano do histórico. Com isso, multiplica-se essa variação à $RECEITA$ estimada para estimar as demais variáveis econômicas do questionário não pertencentes aos blocos citados acima.

Já no caso de imputação através de uma empresa doadora (também conhecido como *vizinho mais próximo*), primeiramente estima-se o valor da $RECEITA$ através da mediana dos valores conhecidos da variável $RECEITA$ das empresas que estejam no mesmo estrato final da empresa imputada no ano de imputação: $RECEITA_{MED}$. Caso não seja possível calcular a mediana desse modo, formam-se outras agregações até que seja possível encontrar a mediana.

Posteriormente, são utilizadas as variáveis PO_{RAIS} , $SALARIO_{RAIS}$ e $RECEITA_{MED}$ para analisar qual será a empresa doadora, sendo escolhida aquela com menor distância considerando as variáveis citadas no ano de imputação, como também respeitando a mesma agregação utilizada para estimar a $RECEITA$ no processo anterior.

Em seguida, é aplicado um procedimento similar ao realizado nas empresas imputadas por histórico aproveitando a estrutura da empresa doadora como suporte. Utiliza-se, então, um fator de variação de cada variável dentro dos blocos de PO e de $SALARIO$ com os respectivos valores de PO e $SALARIO$ do ano de imputação e multiplica-se essa variação com o valor informado de PO_{RAIS} e $SALARIO_{RAIS}$ do ano de referência da imputação. Replica-se esse cálculo em todo o bloco de PO e $SALARIO$ do questionário imputado.

Nos demais blocos, essa variação é feita em relação ao valor de REC^1 informada pela doadora. Posteriormente, aplica-se essa variação à $RECEITA_{MED}$ para estimar as demais variáveis econômicas do questionário não pertencentes aos blocos citados acima.

Os métodos de imputação por histórico e através de uma empresa doadora serão melhor detalhados adiante, mas para isso serão apresentadas algumas novas definições: as variáveis PO_t , PO_t^A , PO_t^{NA} , SAL_t , $SALARIO_t$, $RECEITA_t$, REC_t^1 , REC_t^2 , $CONSINT_t$, $PO_{RAIS;t}$ e $SALARIO_{RAIS;t}$ representam, respectivamente, as variáveis PO , PO^A , PO^{NA} , SAL , $SALARIO$, $RECEITA$, REC^1 , REC^2 , $CONSINT$, PO_{RAIS} e $SALARIO_{RAIS}$ no ano t .

Imputação por Histórico

Sejam $receita_{it}$ e $salario_{it}$ os respectivos valores das variáveis $RECEITA_t$ e $SALARIO_t$ para a empresa i , então quando a imputação é feita por histórico, pode-se estimar $receita_{it}$ pela seguinte relação:

$$\widehat{receita}_{it} = \frac{receita_{it'}}{salario_{it'}} * salario_{it},$$

em que:

t é o ano de referência da empresa que está sendo imputada; e

t' é o ano do histórico usado para estimar a *receita* da empresa i no ano t , para o qual existem informações de *RECEITA* e *SALARIO*.

As demais variáveis do questionário são estimadas replicando a estrutura da empresa no ano do histórico.

Seja *POQ* uma variável qualquer pertencente ao bloco de número 5 de perguntas dos questionários simplificado (Anexo A) e completo (Anexo B) da PAC referentes ao quadro funcional da empresa (pessoal ocupado). Seja também *SALQ* uma variável qualquer do bloco de número 8 do questionário simplificado e de número 14 do questionário completo referentes aos gastos com pessoal da empresa e seja ainda *VARQ* uma variável qualquer dos questionários simplificado ou completo pertencente à parte II dos questionários (Informações da empresa) com exceção dos blocos referentes a pessoal ocupado e gastos com pessoal, então são estimadas as demais variáveis dos questionários da seguinte forma:

- Para o bloco das variáveis de *PO*:

$$\hat{po}_{it} = \frac{po_{it'}}{po_{it'}} * po_{RAIS;it},$$

- Para o bloco das variáveis de *SALARIO*:

$$\hat{sal}_{it} = \frac{sal_{it'}}{salario_{it'}} * salario_{RAIS;it},$$

- Para as demais variáveis, temos:

$$\hat{var}_{it} = \frac{d_{it'}}{receita_{it'}} * \widehat{receita}_{it},$$

em que:

t é o ano de referência da empresa que está sendo imputada;

t' é o ano do histórico usado para estimar po_{it} , sal_{it} e var_{it} no ano t ;

$po_{RAIS;it}$ e $salario_{RAIS;it}$ são os respectivos valores de $PO_{RAIS;t}$ e $SALARIO_{RAIS;t}$ da empresa i ;

e

po_{it} , sal_{it} e d_{it} são, respectivamente, os valores das variáveis *POQ*, *SALQ* e *VARQ* da empresa i no ano t .

Imputação por Empresa Doadora

No caso de empresas imputadas por empresa doadora, estima-se a $receita_{it}$ com base na mediana dos valores dos respondentes do estrato do qual a empresa foi selecionada (UF, CNAE, e ESTF - que será chamado de *GRUPO1*). Caso o número de respondentes do *GRUPO1* seja menor do que quatro, então procura-se o valor da mediana dos respondentes a partir de outros agrupamentos, sempre começando pelo *GRUPO2*. Caso o *GRUPO2* também não tenha quatro empresas respondentes, então seguimos para o *GRUPO3*, depois para o *GRUPO4* e por último o *GRUPO5*, caso os agrupamentos continuem com menos de quatro empresas respondentes.

A descrição de todos os agrupamentos, do *GRUPO1* ao *GRUPO5* são apresentados na Tabela 3.1:

Tabela 3.1: Agrupamento para o cálculo da mediana da *RECEITA*

Grupo	Agrupamento
1	UF x CNAE x ESTF
2	REGIÃO GEOGRÁFICA x CNAE x ESTF
3	CNAE x ESTF
4	CNAE x ESTF-A ¹
5	CNAE

Ou seja, caso o estrato não tenha quatro respondentes, amplia-se a abrangência regional e caso esse corte ainda não seja suficiente para encontrar uma mediana, desconsidera-se a sua posição geográfica e se mesmo assim, ainda não for suficiente, desconsidera-se o estrato final original da empresa.

Com isso, podemos dizer que:

$$\widehat{receita}_{it} = med\{RECEITA_t\}^G, \quad (3.2)$$

em que $med\{RECEITA_t\}^G$ é a mediana da variável $RECEITA_t$ sob o o agrupamento G , sendo G um dos agrupamentos *GRUPO1*, *GRUPO2*, *GRUPO3*, *GRUPO4* ou *GRUPO5*, conforme ordenação descrita anteriormente.

Em seguida, a partir dos valores de $po_{RAIS;it}$, $sal_{RAIS;it}$ e $\widehat{receita}_{it}$ da empresa a ser imputada, seleciona-se uma empresa respondente do mesmo ano da empresa a ser imputada pelo critério da menor distância de uma base pré-criticada a fim de modelar a estrutura da empresa, estimando a relação dessas variáveis com o restante do questionário.

¹ESTF-A = (A1 + A2 + A3)

O cálculo da distância entre as empresas é feito da seguinte forma:

$$distancia_{il} = dist_{il}^{PO} + dist_{il}^{SAL} + dist_{il}^{REC},$$

tal que:

$distancia_{il}$ é o valor simbólico da distância entre as empresas i e l no ano t ;

$dist_{il}^{PO}$ é a parcela do cálculo da distância referente à variável PO entre as empresas i e l no ano t ;

$dist_{il}^{SAL}$ é a parcela do cálculo da distância referente à variável SAL entre as empresas i e l no ano t ;

$dist_{il}^{REC}$ é a parcela do cálculo da distância referente à variável $RECEITA$ entre as empresas i e l no ano t ;

As parcelas $dist_{il}^{PO}$, $dist_{il}^{SAL}$ e $dist_{il}^{REC}$ são obtidas através do seguinte cálculo:

$$\begin{aligned} dist_{il}^{PO} &= \left| \frac{po_{RAIS;it} - po_{lt}}{po_{RAIS;it}} \right|; \\ dist_{il}^{SAL} &= \left| \frac{sal_{RAIS;it} - sal_{lt}}{sal_{RAIS;it}} \right|; \\ dist_{il}^{REC} &= \left| \frac{\widehat{receita_{it}} - receita_{lt}}{\widehat{receita_{it}}} \right|; \end{aligned}$$

tal que po_{lt} , sal_{lt} e $receita_{lt}$ são os valores das variáveis PO_t , SAL_t e $RECEITA_t$ para uma empresa l , $l \neq i$.

A empresa escolhida para ser usada como doadora é aquela que tem o menor valor $distancia_{ik}$ e por isso essa empresa também é conhecida como *vizinho mais próximo*, pois acredita-se ser ela a que mais se assemelha à empresa a ser imputada.

Sejam $\widetilde{po}_{it}$, $\widetilde{sal}_{it}$, $\widetilde{var}_{it}$, $\widetilde{po}_{it}$ e $\widetilde{salario}_{it}$, respectivamente, os valores referentes às variáveis POQ , $SALQ$ e $VARQ$ da empresa selecionada como doadora da empresa imputada i no ano t , então são estimadas as demais variáveis do questionário usando como referência a estrutura dessa empresa como se segue:

- Para o bloco das variáveis de PO :

$$\widehat{po}_{it} = \frac{\widetilde{po}_{it}}{\widetilde{po}_{it}} * po_{RAIS;it}$$

- Para o bloco das variáveis de *SALARIO*:

$$\hat{salq}_{it} = \frac{\widetilde{salq}_{it}}{\widetilde{salar}_{it}} * salario_{RAIS;it}$$

- Para as demais variáveis, temos:

$$\hat{d}_{it} = \frac{\widetilde{varq}_{it}}{\widetilde{receita}_{it}} * \widehat{receita}_{it},$$

em que:

t é o ano de referência da empresa que está sendo imputada; e

t' é o ano do histórico usado para estimar poq_{it} , $salq_{it}$ e d_{it} no ano t .

$po_{RAIS;it}$ e $salar_{RAIS;it}$ são os respectivos valores de $PO_{RAIS;t}$ e $SALAR_{RAIS;t}$ da empresa i ;

e

$\widetilde{poq}_{it}$, $\widetilde{salq}_{it}$ e $\widetilde{varq}_{it}$ são, respectivamente, os valores referentes às variáveis *POQ*, *SALQ* e *VARQ* da empresa i no ano t .

3.2.4 Imputação em alguns institutos de pesquisa

Os institutos de pesquisa do Reino Unido e dos Estados Unidos da América, no que concerne às suas pesquisas econômicas, indicam a importância do uso da imputação, bem como quais modelos são utilizados e quando são aplicados em seus respectivos relatórios metodológicos.

Conforme o relatório metodológico da Annual Business Survey (ABS) do instituto de pesquisa do Reino Unido, o [Office for National Statistics \[2018\]](#), o mesmo usa modelos baseados, principalmente, no histórico da empresa em etapas anteriores ou posteriores da pesquisa para criar grupos de padrões de respostas que são usados como parâmetros para os modelos de imputação. Quando a empresa não houver participado da pesquisa em momentos anteriores ou posteriores, utiliza-se, então, um registro administrativo como *proxy* dessa resposta. São feitas imputações somente nas grandes empresas. Para as pequenas, usa-se a reponderação.

Já o relatório metodológico das Annual Economic Surveys do instituto de pesquisa dos Estados Unidos da América, o [Census Bureau \[2018\]](#), apresenta seus diversos modelos de imputação, porém não sendo claros quanto aos momentos em que cada modelo é aplicado, apenas citando-os. Dentre as diversas pesquisas econômicas existentes no instituto e seus respectivos modelos apresentados no relatório, os modelos com base no histórico da empresa - aplicados de forma semelhante à ABS - e os modelos com base em funções de regressão são os mais utilizados.

No IBGE, geralmente, utilizam-se os modelos baseados no histórico da própria empresa para imputação de uma empresa de grande porte, pois a maior parte dessas empresas já respondeu à pesquisa em algum momento, enquanto as empresas de pequeno porte são imputadas por uma doadora, conforme apresentado na Seção [3.2.3](#).

Em assim sendo, no capítulo a seguir serão indicados os modelos selecionados para estudo e apresentaremos a base de dados utilizada para tal.

Capítulo 4

Proposta de trabalho e base de dados utilizada

Neste capítulo serão apresentados os modelos de imputação que serão utilizados com o intuito de verificar seus respectivos desempenhos na base de dados selecionada. Também serão descritos os critérios adotados para definição da base de dados em questão e os critérios para comparar a eficiência de cada modelo proposto.

4.1 Modelos de imputação utilizados

Os modelos escolhidos para dar sequência ao estudo podem ser divididos basicamente em quatro grupos de modelos, como apresentado na Figura 4.1, além do próprio modelo utilizado pelo IBGE na PAC. Os modelos de imputação selecionados para estudo, então, são:

Figura 4.1: Grupo de modelos de imputação escolhidos para estudo

4.1.1 Modelo atual de imputação do IBGE da PAC

O método de imputação utilizado na PAC foi descrito na Seção 3.2.3 de forma detalhada.

4.1.2 Modelos de imputação de médias de grupos

Foram feitas seis versões do método MIMDC descrito na Seção 3.2.2, considerando o respectivo peso amostral da empresa na estimativa, e são elas:

1. Média do estrato (MÉDIA_ESTRATO);
2. Média do ESTN (MÉDIA_ESTN);
3. Média geral da UF (MÉDIA_UF);
4. Média do ESTF da linha de divulgação (MÉDIA_L4ESTF);
5. Média do estrato da Região (MÉDIA_ESTREG);

6. Média do estrato Brasil (MÉDIA_ESTBR);

4.1.3 Modelos de imputação por função de regressão

As mesmas variações do método MIMDC foram atribuídas ao método MIMR descrito na Seção 3.2.2, sendo utilizadas as informações de *PO* e *SALARIO* como variáveis auxiliares do modelo, conforme segue:

1. Regressão do estrato (REG_ESTF);
2. Regressão do ESTN (REG_ESTN);
3. Regressão da UF (REG_UF);
4. Regressão do ESTF da linha de divulgação (REG_L4);
5. Regressão do estrato da Região (REG_ESTF_REGIONAL);
6. Regressão do estrato Brasil (REG_ESTF_BRASIL);

4.1.4 Modelos de imputação múltipla de uma função distância e de regressão

Aqui foram utilizadas algumas variações dos métodos MIM combinando os métodos MIFD juntamente com o MIMR. As alterações deram-se basicamente no número de *vizinhos mais próximos* selecionados e na função distância utilizada.

1. Regressão das 10 empresas mais próximas - distância utilizada no método PAC (DISTREG_N10);
2. Regressão das 20 empresas mais próximas - distância utilizada no método PAC (DISTREG_N20);
3. Regressão das 10 empresas mais próximas - distância euclidiana (DISTEUCLID_REG_N10);
4. Regressão das 20 empresas mais próximas - distância euclidiana (DISTEUCLID_REG_N20);

4.1.5 Modelos de imputação múltipla de uma função distância e da mediana de grupos

Os métodos a seguir também são decorrentes de algumas variações dos métodos MIM combinando os métodos MIFD juntamente com uma versão mais robusta do método MIMDC. Seguindo o padrão anterior, as alterações deram-se basicamente no número de *vizinhos mais próximos* selecionados e na função distância utilizada.

1. Mediana das 5 empresas mais próximas - distância utilizada no método PAC (DISTREG_medN5);
2. Mediana das 10 empresas mais próximas - distância utilizada no método PAC (DISTREG_medN10);
3. Mediana das 5 empresas mais próximas - distância euclidiana (DSTEUCIDREG_medN5);
4. Mediana das 10 empresas mais próximas - distância euclidiana (DSTEUCIDREG_medN10);

4.1.6 Vantagens e desvantagens dos modelos utilizados

Os modelos de imputação de média, apresentados na Seção 4.1.2, são os mais simples dentre os modelos adotados e com maior facilidade de implementação. Contudo, esses modelos subestimam a variância dos estimadores, pois tratam-se de modelos determinísticos.

Os modelos de imputação que utilizam uma função de regressão, apresentados na Seção 4.1.3, não afetam tanto a variância dos estimadores por não criarem tantos picos na distribuição e ainda acrescentam variáveis auxiliares a fim de melhorar a precisão da estimação.

Já os modelos que se utilizam de uma função distância possuem as mesmas características dos modelos anteriores só que restritos às empresas com menores distâncias de acordo com o critério utilizado, de tal forma que os modelos descritos na Seção 4.1.5 se assemelham aos da Seção 4.1.2 e os modelos apresentados na Seção 4.1.4 com os da Seção 4.1.3.

As versões estocásticas dos métodos acima preservam a variância.

4.2 Critério de avaliação dos modelos aplicados

A fim de medirmos o desempenho de cada método apresentado serão selecionadas algumas empresas aleatoriamente dos domínios estudados e as omitiremos da base de dados utilizada para fazer os cálculos das imputações. Posteriormente, essas empresas serão recolocadas nas bases de dados com seus respectivos valores de imputação adquiridos por cada modelo.

Com isso, serão feitas duas análises:

- Variação Percentual (VP) dos totais estimados contra os totais reais observados na pesquisa;
- e
- Erro Quadrático Médio (EQM) de cada estimador.

Para tal, decidiu-se trabalhar com o seguinte cenário: foi selecionada aleatoriamente uma empresa de cada estrato com pelo menos quatro empresas respondentes para os anos de 2014 e 2015 de alguns estados selecionados. A escolha desses estados será descrita na Seção 4.3.

A seguir, serão descritos os cálculos realizados em cada uma das avaliações propostas.

4.2.1 Análise por VP

Seja VP_b^D a variação percentual entre o total observado na pesquisa e o total predito da variável Y sob o domínio D utilizando o método de imputação b descrito na Seção 4.1. Então tem-se que:

$$VP_b^D = \left| \left(\frac{Y_{imp;b}^{\hat{D}}}{Y_{real}^D} - 1 \right) * 100 \right|, \quad (4.1)$$

tal que Y_{real}^D é o total observado da variável Y na pesquisa sob o domínio D e $Y_{imp;b}^{\hat{D}}$ é o total estimado da variável Y pelo modelo de imputação b sob o domínio D .

4.2.2 Análise por EQM

Seja y_i o verdadeiro valor da variável de interesse Y para a empresa i e $\hat{y}_i$ o valor predito de y_i ; então temos que o $EQM(\hat{Y})$ empírico pode ser calculado da seguinte forma:

$$EQM(\hat{Y})^D = \sum_{i=1}^{n_D} \frac{(y_i - \hat{y}_i)^2}{n_D} \quad (4.2)$$

tal que $EQM(\hat{Y})^D$ é o valor do $EQM(\hat{Y})$ para o domínio D .

4.3 Descrição da base de dados utilizada

A base de dados a ser utilizada no presente trabalho para avaliarmos os diferentes métodos de imputação apresentados será composta por quatro domínios distintos, cada um formado por uma UF. Sendo assim, decidiu-se trabalhar com os quatro estados com maior número de *estratos com pelo menos quatro empresas ativas* nos anos de 2014 e 2015.

Para isso, começaremos apresentando algumas estatísticas da base de dados da PAC 2015 relativa aos anos de 2015 e 2014.

Nas Tabelas 4.1 e 4.2 constam o número de empresas que participaram da PAC nos anos de 2014 e 2015 pela Situação da Empresa (SITE MP), isto é, pela natureza de resposta das empresas para cada UF.

Tabela 4.1: *Número de Empresas por SITEMP para cada UF no ano de 2014*

UF		SITEMP				Total
		Ativa	Morte	Outro	Recusa	
11 - RO	Frequência	416	40	44	21	521
	% UF	79,85	7,68	8,45	4,03	
12 - AC	Frequência	323	32	39	11	405
	% UF	79,75	7,90	9,63	2,72	
13 - AM	Frequência	951	75	67	3	1096
	% UF	86,77	6,84	6,11	0,27	
14 - RR	Frequência	330	50	22	20	422
	% UF	78,20	11,85	5,21	4,74	
15 - PA	Frequência	681	50	97	101	929
	% UF	73,30	5,38	10,44	10,87	
16 - AP	Frequência	279	38	70	49	436
	% UF	63,99	8,72	16,06	11,24	
17 - TO	Frequência	275	58	75	10	418
	% UF	65,79	13,88	17,94	2,39	
21 - MA	Frequência	926	87	73	126	1212
	% UF	76,40	7,18	6,02	10,40	
22 - PI	Frequência	775	42	73	16	906
	% UF	85,54	4,64	8,06	1,77	
23 - CE	Frequência	1.836	105	97	94	2.132
	% UF	86,12	4,92	4,55	4,41	
24 - RN	Frequência	984	59	46	38	1.127
	% UF	87,31	5,24	4,08	3,37	
25 - PB	Frequência	959	46	67	2	1.074
	% UF	89,29	4,28	6,24	0,19	
26 - PE	Frequência	2.309	99	166	156	2.730
	% UF	84,58	3,63	6,08	5,71	
27 - AL	Frequência	658	82	110	29	879
	% UF	74,86	9,33	12,51	3,30	
28 - SE	Frequência	592	55	75	64	786
	% UF	75,32	7,00	9,54	8,14	
29 - BA	Frequência	3.019	117	348	45	3.529
	% UF	85,55	3,32	9,86	1,28	

(continua)

		(continuação)				
UF		SITEMP				Total
		Ativa	Morte	Outro	Recusa	
31 - MG	Frequência	7.648	248	352	6	8.254
	% UF	92,66	3,00	4,26	0,07	
32 - ES	Frequência	1.815	83	72	67	2.037
	% UF	89,10	4,07	3,53	3,29	
33 - RJ	Frequência	7.146	316	440	79	7.981
	% UF	89,54	3,96	5,51	0,99	
35 - SP	Frequência	19.537	671	1.559	232	21.999
	% UF	88,81	3,05	7,09	1,05	
41 - PR	Frequência	5.717	261	287	137	6.402
	% UF	89,30	4,08	4,48	2,14	
42 - SC	Frequência	3.856	198	268	280	4.602
	% UF	83,79	4,30	5,82	6,08	
43 - RS	Frequência	4.888	224	223	109	5.444
	% UF	89,79	4,11	4,10	2,00	
50 - MS	Frequência	1.188	73	76	30	1.367
	% UF	86,91	5,34	5,56	2,19	
51 - MT	Frequência	1.555	110	67	105	1.837
	% UF	84,65	5,99	3,65	5,72	
52 - GO	Frequência	2.144	119	166	45	2.474
	% UF	86,66	4,81	6,71	1,82	
53 - DF	Frequência	1.417	131	115	101	1.764
	% UF	80,33	7,43	6,52	5,73	
Total	Frequência	72.224	3.469	5.094	1.976	82.763
	% Total	87,27	4,19	6,15	2,39	

Tabela 4.2: *Número de Empresas por SITEMP para cada UF no ano de 2015*

UF		SITEMP				Total
		Ativa	Morte	Outro	Recusa	
11 - RO	Frequência	379	45	45	29	498
	% UF	76,10	9,04	9,04	5,82	
12 - AC	Frequência	360	30	33	5	428
	% UF	84,11	7,01	7,71	1,17	
13 - AM	Frequência	995	43	73	2	1.113
	% UF	89,40	3,86	6,56	0,18	
14 - RR	Frequência	357	43	19	6	425
	% UF	84,00	10,12	4,47	1,41	
15 - PA	Frequência	764	40	99	33	936
	% UF	81,62	4,27	10,58	3,53	
16 - AP	Frequência	303	33	80	37	453
	% UF	66,89	7,28	17,66	8,17	
17 - TO	Frequência	332	49	51	7	439
	% UF	75,63	11,16	11,62	1,59	
21 - MA	Frequência	1.012	80	80	120	1.292
	% UF	78,33	6,19	6,19	9,29	
22 - PI	Frequência	779	42	92	24	937
	% UF	83,14	4,48	9,82	2,56	
23 - CE	Frequência	1.972	101	104	43	2.220
	% UF	88,83	4,55	4,68	1,94	
24 - RN	Frequência	1.000	62	43	44	1.149
	% UF	87,03	5,40	3,74	3,83	
25 - PB	Frequência	1.015	42	50	3	1.110
	% UF	91,44	3,78	4,50	0,27	
26 - PE	Frequência	2.280	92	154	162	2.688
	% UF	84,82	3,42	5,73	6,03	
27 - AL	Frequência	744	68	82	15	909
	% UF	81,85	7,48	9,02	1,65	
28 - SE	Frequência	705	52	49	36	842
	% UF	83,73	6,18	5,82	4,28	
29 - BA	Frequência	3.244	94	258	34	3.630
	% UF	89,37	2,59	7,11	0,94	

(continua)

(continuação)						
UF		SITE MP				Total
		Ativa	Morte	Outro	Recusa	
31 - MG	Frequência	7.897	221	293	2	8.413
	% UF	93,87	2,63	3,48	0,02	
32 - ES	Frequência	1.758	68	59	91	1.976
	% UF	88,97	3,44	2,99	4,61	
33 - RJ	Frequência	7.269	263	412	93	8.037
	% UF	90,44	3,27	5,13	1,16	
35 - SP	Frequência	19.830	532	1.236	404	22.002
	% UF	90,13	2,42	5,62	1,84	
41 - PR	Frequência	5.858	272	310	119	6.559
	% UF	89,31	4,15	4,73	1,81	
42 - SC	Frequência	4.040	226	218	197	4.681
	% UF	86,31	4,83	4,66	4,21	
43 - RS	Frequência	5.011	202	179	90	5.482
	% UF	91,41	3,68	3,27	1,64	
50 - MS	Frequência	1.253	58	67	33	1.411
	% UF	88,80	4,11	4,75	2,34	
51 - MT	Frequência	1.647	86	74	91	1.898
	% UF	86,78	4,53	3,90	4,79	
52 - GO	Frequência	2.155	104	122	65	2.446
	% UF	88,10	4,25	4,99	2,66	
53 - DF	Frequência	1.526	99	105	80	1.810
	% UF	84,31	5,47	5,80	4,42	
Total	Frequência	74.485	3.047	4.387	1.865	83.784
	% Total	88,90	3,64	5,24	2,23	

Pode-se verificar que a taxa de empresas que responderam à pesquisa, isto é, empresas *ativas* ou *mortas* é maior do que 78% em praticamente todos os estados com exceção do Amapá, conforme apresentado nas Tabelas 4.1 e 4.2. Contudo, para a escolha de nossas bases, estamos somente interessados nas empresas que constam como *ativas* para que seja possível comparar as respostas reais com as estimadas por cada método.

Nas Tabelas 4.3 e 4.4 são apresentados os números de estratos naturais em cada UF nos anos de 2014 e 2015, respectivamente, e é possível verificar que o número de ESTN não mudou de forma significativa de um ano para o outro, como já era esperado numa pesquisa estrutural.

Tabela 4.3: *Número de Estratos Naturais por UF em 2014*

UF	Número de ESTN	% ESTN
11 - RO	20	2,13
12 - AC	20	2,13
13 - AM	20	2,13
14 - RR	20	2,13
15 - PA	20	2,13
16 - AP	20	2,13
17 - TO	20	2,13
21 - MA	20	2,13
22 - PI	20	2,13
23 - CE	20	2,13
24 - RN	20	2,13
25 - PB	20	2,13
26 - PE	20	2,13
27 - AL	20	2,13
28 - SE	20	2,13
29 - BA	20	2,13
31 - MG	87	9,28
32 - ES	20	2,13
33 - RJ	88	9,39
35 - SP	89	9,50
41 - PR	85	9,07
42 - SC	81	8,64
43 - RS	87	9,28
50 - MS	20	2,13
51 - MT	20	2,13
52 - GO	20	2,13
53 - DF	20	2,13
Total	937	100,0

Tabela 4.4: *Número de Estratos Naturais por UF em 2015*

UF	Número de ESTN	% ESTN
11 - RO	20	2,10
12 - AC	20	2,10
13 - AM	20	2,10
14 - RR	20	2,10
15 - PA	20	2,10
16 - AP	20	2,10
17 - TO	20	2,10
21 - MA	20	2,10
22 - PI	20	2,10
23 - CE	20	2,10
24 - RN	20	2,10
25 - PB	20	2,10
26 - PE	20	2,10
27 - AL	20	2,10
28 - SE	20	2,10
29 - BA	20	2,10
31 - MG	89	9,33
32 - ES	20	2,10
33 - RJ	90	9,43
35 - SP	90	9,43
41 - PR	90	9,43
42 - SC	88	9,22
43 - RS	87	9,12
50 - MS	20	2,10
51 - MT	20	2,10
52 - GO	20	2,10
53 - DF	20	2,10
Total	954	100,0

A PAC possuía 937 estratos naturais em todo o Brasil no ano de 2014, enquanto em 2015, a PAC contava com 954. Pode-se notar também nas Tabelas 4.3 e 4.4 que os estados com ESTN iguais a vinte são aqueles em que o agrupamento de CNAE dá-se a 3 dígitos, enquanto as demais são a 4 dígitos.

Entretanto, como a menor agregação considerada nesse estudo são os *estratos*, é importante que seja escolhida uma base de dados que tenha o maior número possível de respondentes para que quando esses sejam omitidos ainda assim haja um número razoável de respostas nos estratos para aplicação dos modelos em questão. Portanto, optou-se por utilizar como domínio os 4 estados com a maior taxa de empresas *ativas* em 2014 e 2015.

Tabela 4.5: *Número de ESTF* e ESTF em 2014 e 2015 e os respectivos percentuais de ESTF**

UF	ESTF*		ESTF		%	
	2014	2015	2014	2015	2014	2015
11 - RO	45	42	91	93	49,5%	45,2%
12 - AC	38	41	82	86	46,3%	47,7%
13 - AM	56	55	103	104	54,4%	52,9%
14 - RR	39	42	80	86	48,8%	48,8%
15 - PA	45	51	93	101	48,4%	50,5%
16 - AP	23	27	82	88	28,0%	30,7%
17 - TO	34	41	87	88	39,1%	46,6%
21 - MA	44	52	99	111	44,4%	46,8%
22 - PI	63	64	95	107	66,3%	59,8%
23 - CE	67	63	101	107	66,3%	58,9%
24 - RN	60	57	93	95	64,5%	60,0%
25 - PB	67	67	100	105	67,0%	63,8%
26 - PE	65	62	101	111	64,4%	55,9%
27 - AL	42	47	93	105	45,2%	44,8%
28 - SE	43	55	95	104	45,3%	52,9%
29 - BA	64	73	103	111	62,1%	65,8%
31 - MG	260	274	376	390	69,1%	70,3%
32 - ES	72	70	100	108	72,0%	64,8%
33 - RJ	226	236	372	383	60,8%	61,6%
35 - SP	272	297	390	401	69,7%	74,1%
41 - PR	275	276	379	387	72,6%	71,3%
42 - SC	239	241	367	386	65,1%	62,4%
43 - RS	255	276	374	378	68,2%	73,0%
50 - MS	59	68	99	110	59,6%	61,8%
51 - MT	58	60	99	102	58,6%	58,8%
52 - GO	74	75	107	113	69,2%	66,4%
53 - DF	55	64	100	108	55,0%	59,3%
Total	2640	2776	4261	4468		

Dessa forma, foram calculados os números de *estratos* com pelo menos quatro empresas ativas nos anos de 2014 e 2015 (ESTF*), em seguida, a partir dos números de estratos finais totais para os anos de 2014 e 2015, foram calculados os percentuais de estratos finais com pelo menos quatro empresas ativas para ambos os anos, conforme apresentado na Tabela 4.5.

Os percentuais de estratos finais com pelo menos quatro empresas ativas em 2014 são apresentados na Tabela 4.6, enquanto na Tabela 4.7 são apresentados os percentuais de estratos finais com pelo menos quatro empresas ativas de 2015, com o devido ranqueamento dos estados em cada ano.

Tabela 4.6: *Percentual de Estratos Naturais com pelo menos quatro empresas ativas em 2014*

UF	% de Estratos com pelo menos quatro empresas ativas	Classificação
11 - RO	49,5%	19
12 - AC	46,3%	22
13 - AM	54,4%	18
14 - RR	48,8%	20
15 - PA	48,4%	21
16 - AP	28,0%	27
17 - TO	39,1%	26
21 - MA	44,4%	25
22 - PI	66,3%	9
23 - CE	66,3%	8
24 - RN	64,5%	11
25 - PB	67,0%	7
26 - PE	64,4%	12
27 - AL	45,2%	24
28 - SE	45,3%	23
29 - BA	62,1%	13
31 - MG	69,1%	5
32 - ES	72,0%	2
33 - RJ	60,8%	14
35 - SP	69,7%	3
41 - PR	72,6%	1
42 - SC	65,1%	10
43 - RS	68,2%	6
50 - MS	59,6%	15
51 - MT	58,6%	16
52 - GO	69,2%	4
53 - DF	55,0%	17

Tabela 4.7: *Percentual de Estratos Naturais com pelo menos quatro empresas ativas em 2015*

UF	% de Estratos com pelo menos quatro empresas ativas	Classificação
11 - RO	45,2%	25
12 - AC	47,7%	22
13 - AM	52,9%	18
14 - RR	48,8%	21
15 - PA	50,5%	20
16 - AP	30,7%	27
17 - TO	46,6%	24
21 - MA	46,8%	23
22 - PI	59,8%	13
23 - CE	58,9%	15
24 - RN	60,0%	12
25 - PB	63,8%	8
26 - PE	55,9%	17
27 - AL	44,8%	26
28 - SE	52,9%	18
29 - BA	65,8%	6
31 - MG	70,3%	4
32 - ES	64,8%	7
33 - RJ	61,6%	11
35 - SP	74,1%	1
41 - PR	71,3%	3
42 - SC	62,4%	9
43 - RS	73,0%	2
50 - MS	61,8%	10
51 - MT	58,8%	16
52 - GO	66,4%	5
53 - DF	59,3%	14

Como desejava-se selecionar os 8 estados com maior percentual de estratos finais com pelo menos quatro empresas ativas nos anos de 2014 e 2015 para em seguida fazer uma análise mais detalhada dos estratos finais, repetindo a análise anterior dentro desses estratos finais, foram, então, selecionados os sete estados com maior percentual em 2014 e os sete estados com maior percentual em 2015, sendo seis deles coincidentes em ambos os anos.

Portanto, em ordem de maior para o menor, tem-se que os estados do Paraná, Espírito Santo, São Paulo, Goiás, Minas Gerais, Rio Grande do Sul e Paraíba são os estados com maior percentual de estratos finais com pelo menos quatro empresas ativa em 2014, conforme apresentado na Tabela 4.6, enquanto em 2015, foram os estados de São Paulo, Rio Grande do Sul, Paraná, Minas Gerais, Goiás, Bahia e Espírito Santo, apresentados na Tabela 4.7.

Assim, foram selecionados os estados da Paraíba, Bahia, Minas Gerais, Espírito Santo, São Paulo, Paraná, Rio Grande do Sul e Goiás para uma segunda análise.

Para seguir com essa análise mais detalhada são apresentados os números de estratos finais totais para os anos de 2014 e 2015 nas Tabelas 4.8 e 4.9, respectivamente.

Tabela 4.8: *Número Total de Estratos Finais nos estados selecionados em 2014*

UF		ESTF						Total
		A1	A2	A3	C	G	NS	
25 - PB	Frequência	20	20	20	20	16	4	100
	% UF	20,00	20,00	20,00	20,00	16,00	4,00	
29 - BA	Frequência	20	20	20	20	20	3	103
	% UF	19,42	19,42	19,42	19,42	19,42	2,91	
31 - MG	Frequência	75	75	74	82	61	9	376
	% UF	19,95	19,95	19,68	21,81	16,22	2,39	
32 - ES	Frequência	20	20	20	20	19	1	100
	% UF	20,00	20,00	20,00	20,00	19,00	1,00	
35 - SP	Frequência	75	75	74	87	69	10	390
	% UF	19,23	19,23	18,97	22,31	17,69	2,56	
41 - PR	Frequência	75	74	73	80	70	7	379
	% UF	19,79	19,53	19,26	21,11	18,47	1,85	
43 - RS	Frequência	75	74	74	82	60	9	374
	% UF	20,05	19,79	19,79	21,93	16,04	2,41	
52 - GO	Frequência	20	20	20	20	20	7	107
	% UF	18,69	18,69	18,69	18,69	18,69	6,54	
Total	Frequência	380	378	375	411	335	50	1.929
	% Total	19,70	19,60	19,44	21,31	17,37	2,59	

Tabela 4.9: *Número Total de Estratos Finais para as UF selecionadas em 2015*

UF		ESTF						Total
		A1	A2	A3	C	G	NS	
25 - PB	Frequência	20	20	20	20	15	10	105
	% UF	19,05	19,05	19,05	19,05	14,29	9,52	
29 - BA	Frequência	20	20	20	20	19	12	111
	% UF	18,02	18,02	18,02	18,02	17,12	10,81	
31 - MG	Frequência	75	75	74	87	63	16	390
	% UF	19,23	19,23	18,97	22,31	16,15	4,10	
32 - ES	Frequência	20	20	20	20	18	10	108
	% UF	18,52	18,52	18,52	18,52	16,67	9,26	
35 - SP	Frequência	75	75	74	89	70	18	401
	% UF	18,70	18,70	18,45	22,19	17,46	4,49	
41 - PR	Frequência	75	75	73	85	66	13	387
	% UF	19,38	19,38	18,86	21,96	17,05	3,36	
43 - RS	Frequência	75	74	73	82	58	16	378
	% UF	19,84	19,58	19,31	21,69	15,34	4,23	
52 - GO	Frequência	20	20	20	20	19	14	113
	% UF	17,70	17,70	17,70	17,70	16,81	12,39	
Total	Frequência	380	379	374	423	328	109	1.993
	% Total	19,07	19,02	18,77	21,22	16,46	5,47	

A seguir, nas Tabelas 4.10 e 4.11 são apresentados os números de estratos finais com pelo menos quatro empresas ativas para os anos de 2014 e 2015 para os estados selecionados, respectivamente.

Tabela 4.10: *Número de Estratos Finais com pelo menos quatro Empresas Ativas para as UFs selecionadas em 2014*

UF		ESTF					Total
		A1	A2	A3	C	G NS	
25 - PB	Frequência	10	18	17	17	5	67
	% UF	14,93	26,87	25,37	25,37	7,46	
29 - BA	Frequência	3	11	19	20	11	64
	% UF	4,69	17,19	29,69	31,25	17,19	
31 - MG	Frequência	31	61	72	71	25	260
	% UF	11,92	23,46	27,69	27,31	9,62	
32 - ES	Frequência	7	16	17	20	12	72
	% UF	9,72	22,22	23,61	27,78	16,67	
35 - SP	Frequência	14	60	60	84	54	272
	% UF	5,15	22,06	22,06	30,88	19,85	
41 - PR	Frequência	37	63	64	68	43	275
	% UF	13,45	22,91	23,27	24,73	15,64	
43 - RS	Frequência	40	66	60	69	20	255
	% UF	15,69	25,88	23,53	27,06	7,84	
52 - GO	Frequência	8	19	18	20	9	74
	% UF	10,81	25,68	24,32	27,03	12,16	
Total	Frequência	150	314	327	369	179	1.339
	% Total	11,20	23,45	24,42	27,56	13,37	

Tabela 4.11: *Número de Estratos Finais com pelo menos quatro Empresas Ativas para as UFs selecionadas em 2015*

UF		ESTF						Total
		A1	A2	A3	C	G	NS	
25 - PB	Frequência	11	16	16	17	5	2	67
	% UF	16,42	23,88	23,88	25,37	7,46	2,99	
29 - BA	Frequência	7	13	17	20	12	4	73
	% UF	9,59	17,81	23,29	27,40	16,44	5,48	
31 - MG	Frequência	33	63	69	79	23	7	274
	% UF	12,04	22,99	25,18	28,83	8,39	2,55	
32 - ES	Frequência	7	18	17	19	9	0	70
	% UF	10,00	25,71	24,29	27,14	12,86	0,00	
35 - SP	Frequência	29	59	60	87	55	7	297
	% UF	9,76	19,87	20,20	29,29	18,52	2,36	
41 - PR	Frequência	36	60	61	73	41	5	276
	% UF	13,04	21,74	22,10	26,45	14,86	1,81	
43 - RS	Frequência	42	66	62	73	26	7	276
	% UF	15,22	23,91	22,46	26,45	9,42	2,54	
52 - GO	Frequência	9	15	19	19	9	4	75
	% UF	12,00	20,00	25,33	25,33	12,00	5,33	
Total	Frequência	174	310	321	387	180	36	1.408
	% Total	12,36	22,02	22,80	27,49	12,78	2,56	

Foram, então, calculadas as taxas de estratos finais com pelo menos quatro empresas ativas por ESTF para as unidades da federação selecionadas, baseando-se na relação das Tabelas 4.8 e 4.10 e das Tabelas 4.9 e 4.11. Esses percentuais são apresentados nas Tabelas 4.12 e 4.13, onde estão destacados em negrito os valores maiores ou iguais a 0,80 para facilitar a análise.

Tabela 4.12: *Percentual de Estratos Finais com pelo menos quatro empresas ativas em 2014*

UF	ESTF						Total
	A1	A2	A3	C	G	NS	
25 - PB	0,50	0,90	0,85	0,85	0,31	0,00	0,67
29 - BA	0,15	0,55	0,95	1,00	0,55	0,00	0,62
31 - MG	0,41	0,81	0,97	0,87	0,41	0,00	0,69
32 - ES	0,35	0,80	0,85	1,00	0,63	0,00	0,72
35 - SP	0,19	0,80	0,81	0,97	0,78	0,00	0,70
41 - PR	0,49	0,85	0,88	0,85	0,61	0,00	0,73
43 - RS	0,53	0,89	0,81	0,84	0,33	0,00	0,68
52 - GO	0,40	0,95	0,90	1,00	0,45	0,00	0,69

Tabela 4.13: *Percentual de Estratos Finais com pelo menos quatro empresas ativas em 2015*

UF	ESTF						Total
	A1	A2	A3	C	G	NS	
25 - PB	0,55	0,80	0,80	0,85	0,33	0,20	0,64
29 - BA	0,35	0,65	0,85	1,00	0,63	0,33	0,66
31 - MG	0,44	0,84	0,93	0,91	0,37	0,44	0,70
32 - ES	0,35	0,90	0,85	0,95	0,50	0,00	0,65
35 - SP	0,39	0,79	0,81	0,98	0,79	0,39	0,74
41 - PR	0,48	0,80	0,84	0,86	0,62	0,38	0,71
43 - RS	0,56	0,89	0,85	0,89	0,45	0,44	0,73
52 - GO	0,45	0,75	0,95	0,95	0,47	0,29	0,66

Como o estrato final A1 é o que tem a pior taxa de resposta dentre os estratos finais, optou-se por trabalhar com as unidades da federação da Paraíba e Rio Grande do Sul por serem aquelas com maior percentual de resposta em 2014 e 2015 para esse estrato.

O estado do Paraná também foi acrescentado à análise por possuir mais de 70% no percentual de estratos finais com pelo menos quatro empresas ativas em ambos os anos analisados.

O uso do estado de São Paulo foi descartado por apresentar percentual muito baixo de estratos finais A1 com pelo menos quatro empresas ativas em 2014, apesar de os outros estratos finais terem apresentado altos percentuais.

A escolha da última UF ficou entre os estados de Minas Gerais e Espírito Santo por terem mais de 70% no percentual total de estratos finais com pelo menos quatro empresas ativas em pelo menos um dos anos. Logo, optou-se em trabalhar com o Espírito Santo, pois o estado de Minas Gerais tem a mesma agregação de códigos CNAE para a seleção da amostra do que os estados do Paraná e Rio Grande do Sul, ou seja, incluindo o estado do Espírito Santo nas análises temos duas unidades da federação para cada agregação de códigos CNAE.

Com isso, foram escolhidos os estados da **Paraíba, Espírito Santo, Paraná e Rio Grande do Sul** como base para o trabalho em questão e, a partir deles, foi selecionada uma empresa de

forma aleatória de cada estrato final desses estados para aplicação dos modelos selecionados.

Dessa forma, foram testadas 1269 empresas no total espalhadas entre as quatro unidades da federação escolhidas, conforme apresentado na Tabela 4.14.

Tabela 4.14: *Número de empresas na base utilizada para testar os modelos de imputação por ano e UF*

UF	Ano		Total
	2014	2015	
25 - PB	66	61	127
32 - ES	67	69	136
41 - PR	264	246	510
43 - RS	246	250	496
Total	643	626	1269

Nas Tabelas C.1 e C.2 do Anexo C apresentam-se o número de empresas imputadas por ESTF para os anos de 2014 e 2015, respectivamente.

No capítulo a seguir são apresentados os resultados obtidos para cada método descrito no início do capítulo e aplicado na base de dados descrita na sequência.

Capítulo 5

Resultados obtidos

Neste capítulo são apresentados os resultados dos erros quadráticos médios (EQM) para os modelos de imputação utilizados e as variações percentuais (VP) entre o total real da *RECEITA* (valor informado pela empresa) e o total estimado (com o valor real substituído pela imputação em questão).

Os erros quadráticos médios são expostos inicialmente de forma mais desagregada, abertos por ANO x UF x ESTF x CNAE, e posteriormente de forma menos desagregada, apenas por ANO x UF x CNAE. Já as diferenças dos totais são apresentadas somente agregadas por ANO x UF x CNAE.

A seguir são reportados os erros quadráticos médios dos modelos utilizados em cada um dos agrupamentos descritos acima. Para cada modelo é apresentada uma tabela (em forma de figura) com seus maiores valores destacados em amarelo, que representam os piores erros quadráticos médios, e junto à tabela é apresentado o seu respectivo gráfico contendo o desempenho de cada modelo.

O total de cada coluna representa o EQM total de cada modelo, considerando todas as empresas do domínio em questão, de forma que foram agregadas por CNAE (linhas) na tabela. Caso algum modelo apresentasse um resultado muito maior do que os demais, ele foi retirado da tabela e a análise foi refeita. Já o total das linhas representa o EQM total de cada CNAE considerando todos os modelos e, caso algum desses totais fosse muito maior do que os demais, essa CNAE era descartada da tabela para que a análise fosse refeita. Os resultados da análise completa são exibidos na primeira tabela e os referentes à análise refeita após as exclusões na segunda tabela.

Para auxiliar nas análises, foi destacado em verde o modelo que apresentou o MENOR EQM dentre todos os modelos da tabela. Aqueles modelos cuja raiz quadrada do EQM variasse menos do que 30% do menor EQM da tabela também teve o seu nome destacado em verde.

5.1 EQM das imputações por CNAE x ESTF

A seguir, nas Figuras 5.1 a 5.20 foram apresentados os EQM dos Estados da Paraíba e do Espírito Santo para cada ESTF no ano de 2015. Os resultados dos demais agrupamentos serão apresentados na Seção 6.1 de forma resumida.

5.1.1 ANO 2015 / UF 25 - PB / ESTF A1

Na Figura 5.1 é possível verificar que os modelos de imputação da regressão dos 10 e 20 vizinhos mais próximos obtidos pela distância euclidiana não apresentam bons resultados e atrapalham a análise dos demais modelos no gráfico, por isso foram retirados da análise. Os modelos de média e de regressão do estrato natural também não tem um desempenho muito bom e, com isso, também foram retirados da análise.

Portanto, a partir da Figura 5.2, podemos notar que o modelo do IBGE tem o menor EQM, sendo seguido dos modelos dos 10 vizinhos mais próximos obtidos pelo cálculo de distância do IBGE e das médias dos estratos regional e do Brasil.

Figura 5.2: EQM dos modelos de imputação (exceto DISTEUCID_REG_N10, DISTEUCID_REG_N20, MÉDIA_ESTN e REG_ESTN): Ano 2015 / UF 25 - PB / ESTF A1

5.1.2 ANO 2015 / UF 25 - PB / ESTF A2

A partir da Figura 5.3, podemos perceber que os modelos dos 10 e 20 vizinhos mais próximos utilizando a distância euclidiana apresentam EQM muito altos e, com isso, optou-se por retirar esses modelos da análise.

Em seguida, também foi verificado que os modelos IBGE, DISTEUCLIDREG_medN5, MÉDIA_ESTN, REG_ESTF e REG_EST. Ao retirarmos esses modelos, conforme Figura 5.4, vemos que os modelos DISTREG_medN10, DISTREG_medN5 e REG_ESTF_REGIONAL têm os menores EQM.

CNAE	IBGE	DISTECLUCID REG_N10	DISTECLUCID REG_N20	DISTECLUCID REG_mediN10	DISTECLUCID REG_mediN5	DISTREG_mediN10	DISTREG_mediN5	DISTREG_N10	DISTREG_N20	MÉDIA ESTRH	MÉDIA ESTN
451	3.905.342.725.636	4.351.667.136.741	16.737.325.782.338	21.456.841.415.409	22.848.122.760.841	19.021.003.109.556	19.236.074.951.025	13.808.437.785.475	7.241.219.803.866	11.773.957.175.963	22.459.656.332.325
453	122.495.100.049	898.059.994.684	40.477.875.978.300	332.121.113.700	353.665.711.204	5.913.840.702	10.189.993.025	11.991.660	331.692.756.493	2.284.039.743	16.227.564.467
454	52.853.550.201	68.878.517.425	3.229.074.625	43.711.101.184	43.322.675.881	4.481.365.249	60.404.567.529	17.408.070.669	30.214.917.529	881.669.067.717	542.385.173.624
462	1.324.887.201	29.533.876.454.820.100	2.348.250.088.689.410	40.126.700.172	42.334.297.009	485.768.968.784	2.706.258.595.041	1.138.544.379.857	1.293.334.818.333	35.812.680.953.124	342.253.763.832
464	5.656.794.251.225.570	1.121.820.506.667.780.000	2.756.417.937.740.790	18.358.894.033.984	179.870.934.090.409	39.430.839.184	567.697.450.849	2.249.997.656.818	1.326.598.940.828	15.146.875.017.711	2.525.265.616.570
466	1.175.381.222.500	5.323.974.903.662	409.955.951.226	960.964.562.944	465.220.849.041	726.572.974.056	5.858.689.343.729	18.768.658.183.357	5.138.967.343.729	763.194.184.223	4.610.655.081.308
467	35.952.016	1.741.743.812.237	353.550.554.313	960.964.562.944	2.107.575.159.001	972.940.864	5.650.979.929	357.378.048.274	5.176.574.262.445	4.610.655.081.308	123.748.986.236.520
468	91.328.466.436	4.039.028.974.618.040	1.386.004.444.094.040	835.825.635.225	143.147.965.801	1.660.440.252	554.482.772.496	15.114.510.753.112	37.541.761.799.816	20.595.619.176.934	19.885.571.067.623
469	113.055.320.169	23.747.640.673.350	117.469.645.585.399	107.695.548.900	2.222.997.504.784	65.042.596.190	29.995.815.249	293.282.898.760	6.860.332.999.985	3.445.807.387.278	569.075.779.284.027
471	340.210.892.176	2.986.939.439.093	4.489.036.755.688	1.331.203.673.284	1.438.265.322.729	755.132.323.272	737.264.366.881	574.788.300.706	33.807.708.715	3.613.218.205	541.247
472	15.319.507.984	4.702.119.184	13.923.205.398	627.702.916	4.178.717.449	5.937.781.249	920.636.964	5.555.805.860.236	4.177.598.330.075	60.316.048.572	52.049.598.215
473	7.404.129.428.809	244.513.649.210	38.986.543.030.582	31.545.426.090.492	143.147.965.801	4.043.815.281	347.847.525.796	777.175.569.214	259.814.833.109	2.615.894.216.872	109.065.229.180
474	227.008.414.116	64.040.022.255.001	5.536.286.905.129	4.113.816.316.770	4.183.652.978.404	3.323.498.541.162	3.459.369.563.844	3.879.748.341.122	2.728.526.370.875	2.581.024.925.622	3.062.286.647.907
475	5.212.089	206.264.630.868.316	113.354.610.329.619	33.678.856.324	2.772.022.500	13.415.778.102	49.671.328.384	272.410.661.070	84.814.086.976	72.806.359.722	569.079.380.943
477	35.179.878.969	7.198.521.305	4.656.168.048.092	48.811.182	640.545.481	142.539.094.392	182.508.584.100	190.461.825.829	138.230.507.351	368.153.254.259	368.242.904.762
478	130.778.426.689	426.922.515.141.254	655.837.377.911.789	75.487.013.001	1.737.577.421.584	177.852.819.076	562.476.000.256	1.706.020.845.618	1.698.670.380.816	153.720.013.209	111.201.803.437
Total	354.400.543.763.163	77.258.125.894.644.300	2.995.399.756.458.550	4.984.185.819.344	15.086.879.553.650	1.548.329.264.211	2.148.093.910.944	4.011.539.698.235	4.628.884.949.030	6.180.522.507.529	46.497.692.059.059
CNAE	MÉDIA ESTRATO	MÉDIA ESTRG	MÉDIA LESTF	MÉDIA UF	REG_ESTF	REG_ESTF_BRAZIL	REG_ESTF_REGIONAL	REG_ESTN	REG_L4	REG_UF	Total
451	13.668.047.655.780	11.884.611.621.112	19.526.264.213.601	12.141.059.489.263	134.243.474.237.877	11.010.756.923.884	9.869.568.946.455	171.605.606.007	11.233.018.508.774	2.667.277.221.131	18.535.971.115.527
453	88.695.308.697	28.671.998.521	36.441.385.546	552.873.570.335	7.774.804.585	375.746.104	5.415.985.920	96.799.979.296.840	236.959.408.560	7.737.265.854.560	1.932.640.171.668.740
454	209.581.564.837	619.553.563.344	52.253.722.266	1.352.666.962.221	61.846.050.956	178.702.807.558	4.553.108.322	302.247.312.808	335.954.144.733	6.408.031.821.644	536.854.629.243
462	851.588.850.260	4.344.706.009.397	21.517.094.466.625	1.203.353.460.637	2.486.155.704.828	21.330.080.535.291	834.636.638.033	2.001.978.430.410	8.608.319.326.860	11.843.904.657.148	1.523.762.428.092.970
464	1.473.013.074.006	13.971.811.912.733	124.747.255.316	15.170.038.135.971	3.273.222.427.819	376.843.012.001	2.044.025.697.696	1.501.424.117.261	4.280.625.689.438	361.129.819.049	53.841.397.510.054.600
466	1.887.381.136	220.412.015.057	15.996.629.979.728	209.673.037.785	29.510.355.488.442	1.134.824.742.236	13.549.095.243.648	68.272.123.370	22.087.064.536.881	10.302.809.955.568	5.445.638.192.473
467	2.506.253.103.225	3.981.094.097.017	22.723.845.219.791	1.503.872.234.277	3.448.601.439.368	2.306.617.026.728	4.290.110.405.135	2.067.773.103.883	7.707.080.816.287	8.245.407.343.155	8.836.482.023.618
468	1.067.286.312.120	335.762.754.175	15.333.055.837.417	159.925.885.204	1.973.098.165.839	13.549.095.243.648	4.290.110.405.135	2.067.773.103.883	7.707.080.816.287	6.422.592.354.704	265.365.848.746.130
469	41.994.469.538.596	2.827.864.937.862	20.290.270.027.862	926.973.606.435	2.762.625.025.080	8.646.917.332.251	11.011.068.679.010	49.449.428.536.962	1.652.616.971.761	5.305.454.789.503	172.751.381.462.259
471	11.346.025.015	69.807.234.751	252.465.056.387	24.233.985.012	977.466.768	227.409.425.379	46.411.983.728	4.302.512.244.075	44.081.577.799	2.490.948.086.578	959.381.414.738
472	5.937.078.959	112.585.204.166	403.158.950.438	1.672.053.073.064	5.489.566.325	117.306.155.709	247.281.971.007	23.015.402.166	640.617.278.192	7.206.885.882.003	967.906.575.727
473	6.822.929.150.097	6.536.407.258.939	29.414.693.365.939	22.708.689.072.545	289.908.950.252.251	78.321.711.762	78.321.711.762	1.087.594.956.166	9.161.788.737.035	9.161.788.737.035	23.163.804.649.570
474	3.473.969.129.308	2.912.855.276.202	2.160.438.473.043	658.875.258.080	2.791.638.601.604	2.058.656.526.849	1.671.850.933.129	1.235.804.883.475	50.307.326.699	2.071.456.109.813	5.535.766.060.865
475	1.481.249.169	38.679.279.983	232.693.349.370	1.300.773.747.851	2.079.155.392.802	8.204.184.003	32.449.097.454	10.547.537.955.751	241.624.767.427	6.218.998.377.183	16.258.056.804.059
477	18.149.411.040	324.192.995.955	414.855.780.207	1.695.789.131.537	20.703.479.799	122.885.732.779	150.144.932.782	195.773.876.053	615.557.895.469	7.223.183.462.259	804.314.689.219
478	164.938.598.627	84.134.751.231	292.224.563.641	1.436.904.910.122	205.002.029.015	17.804.288.328	532.162.613.509	42.713.517.221	6.090.592.458.601	52.285.268.388.223	5.730.148.859.429
Total	4.522.473.339.430	3.018.246.305.670	9.298.807.790.449	3.918.620.410.021	202.016.563.739.585	3.822.761.899.657	2.157.112.988.045	20.686.726.175.093	4.603.967.887.030	5.730.148.859.429	3.616.827.988.286.780

Figura 5.3: EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF A2

Figura 5.4: EQM dos modelos de imputação (exceto DISTEUCOLID_REG_N10 e N20, IBGE, DISTEUCOLIDREG medN5, MÉDIA ESTN, REG ESTF e REG EST): Ano 2015 / UF 25 - PB / ESTF A2

5.1.3 ANO 2015 / UF 25 - PB / ESTF A3

Quando analisamos a Figura 5.5, novamente percebemos os modelos DISTEUCLID_REG_N10 e DISTEUCLID_REG_N20 apresentam os maiores valores de EQM, seguidos dos modelos DISTEUCLIDREG_medN5, DISTREG_N10 e N20, MÉDIA_ESTN e ESTRATO e REG_ESTF, ESTN e L4 que também apresentam erros muito grandes para este recorte e foram excluídos.

Com isso, a Figura 5.6 contém os modelos restantes e podemos ver que aqueles com menor EQM são o do IBGE, REG_ESTF_BRASIL e REG_ESTF_REGIONAL.

Figura 5.5: EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF A3

Figura 5.6: EQM dos modelos de imputação (exceto DISTEUCLI_REG_N10 e N20, DISTEUCLI-DREG_medN5, DISTREG_N10 e N20, MÉDIA_ESTN e ESTRATO e REG_ESTF, ESTN e L4): Ano 2015 / UF 25 - PB / ESTF A3

5.1.4 ANO 2015 / UF 25 - PB / ESTF C

A partir da Figura 5.7, podemos perceber que os modelos DISTEUCLID_REG_N10 e N20 mais uma vez não apresentam um bom resultado e foram retirados da análise. Além deles, podemos ver que a CNAE 468 (Comércio atacadista especializado em outros produtos) aparenta ser um *outlier* dentre as CNAE e, por isso, optou-se por também excluir essa linha do modelo para melhor podermos analisá-lo.

Portanto, a Figura 5.8 apresenta os resultados dos EQM sem os modelos DISTEUCLID_REG_N10 e N20 e sem a CNAE 468. Nela podemos verificar que os modelos IBGE, DISTREG_medN5 e N10, DISTREG_N5 e N10 e REG_ESTF_REGIONAL apresentam os menores erros.

Figura 5.7: *EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF C*

CNAE	IRGE	DISTECLIDREG_med10	DISTECLIDREG_med5	DISTREG_med10	DISTREG_med5	DISTREG_M10	DISTREG_M20	MEDIA_ESTRB	MEDIA_ESTN	MEDIA_ESTRATO
451	1.919,045.252.272.560	8.078.183.230.295.800	8.104.756.463.947.020	1.013.597.147.798.640	685.700.677.976.769	1.402.390.688.722.070	1.132.450.558.920.830	73.491.678.788.864	6.481.437.988.422.900	1.929.495.438.710.790
453	30.901.665.330.724	71.746.950.355.670	72.058.927.495.000	2.018.297.153.052.224	2.172.107.077.976.249	18.553.949.510.589	1.163.621.598.921.598	28.206.010.498.647	60.321.124.707.297	11.810.943.476.550
454	41.167.132.609	3.429.272.077.440.170	3.429.163.039.611.140	1.868.149.074.006	4.674.053.745.936	41.082.623.485	26.262.595.104.284	990.077.426.229.670	3.319.424.986.531.460	1.206.666.991.785.080
463	1.012.203.002.889	59.079.376.788.721	55.311.062.116.900	1.309.596.252.736	1.709.956.238.659	45.363.060.440.566	44.447.755.011.020	920.966.320.191.070	14.052.183.944.000	889.604.561.798.001
464	20.637.904.066.816	1.920.682.127.065.340	338.426.171.627.769	656.083.788.727.556	476.251.927.300.809	606.188.559.084.930	42.049.124.981.981	237.240.394.277.747	913.174.496.986.978	51.600.500.183.059
466	364.528.552.644	1.358.309.084.010.030	1.355.920.071.752.400	1.323.052.962.702.030	1.341.211.095.257.400	1.104.187.575.735.130	1.100.763.208.331.510	156.448.044.787.360	1.232.291.550.377.910	748.496.092.324.656
467	101.063.412.180.900	2.095.324.405.121	952.630.656.784	457.964.971.085.872	578.970.382.935.874	521.707.803.663.544	371.788.335.324.571	627.022.688.645.613	75.628.680.300.163	864.698.932.793.421
471	17.073.603.556	125.851.017.735.556	143.940.822.081.135	651.390.710.392.900	1.097.580.080.970.820	410.436.966.163.503	635.468.013.930.064	635.468.013.930.064	151.506.666.635.810	7.383.633.240.364.310
473	17.091.640.225	142.851.140.641.225	143.940.822.081.135	2.468.816.819.025	3.701.091.087.684	924.590.279.875	2.571.968.285.138	333.377.647.691.000	116.586.666.327.611	51.128.040.608.440
474	90.880.543.296	3.742.905.446.281	3.897.655.165.504	414.710.240.400	412.494.054.009	24.493.534.000	640.808.190.868	10.493.453.499.519	2.827.311.732.199	4.188.237.666.000
477	939.662.687.044	196.511.711.555.432	206.268.140.857.329	4.120.597.535.550	675.291.141.121	18.999.453.499.377	65.378.687.761.643	85.399.043.664.572	76.247.461.174.194	72.965.370.371.225
478	273.894.688.455.696	838.147.112.543.670	839.140.511.331.396	171.956.722.353.529	90.312.235.725.625	196.745.577.274.769	123.661.074.627.457	324.611.179.266.388	822.109.678.089.491	413.384.752.227.329
475	4.256.480.639.376	187.288.010.873.041	183.723.196.129.209	29.824.350.463.056	23.130.463.023.396	16.070.786.568.404	4.664.866.322.299	4.664.866.322.299	162.496.558.799.323	18.536.656.507.837
476	32.217.378.044	90.334.810.800	64.528.003.169	1.361.348.732.592	6.064.689.261.661	3.812.633.659.098	2.873.410.217.797	24.342.254.340.695	33.338.424.664	13.546.659.321.522
477	10.901.480.244.121	94.905.147.876	47.555.147.876	2.317.739.321.952	4.843.137.708.369	1.128.667.863.623	1.350.524.923.922	213.623.668.666.666	1.825.476.072	58.900.206.369.711
478	7.462.022.689	231.564.026.521	316.082.332.944	11.295.658.892.702	26.066.405.947.729	37.786.560.976.760	13.224.960.356.986	49.405.672.143.177	178.498.354.344	44.645.415.777.630
Total	147.701.998.374.576	987.282.242.538.694	927.560.700.827.414	272.900.713.121.436	265.377.853.531.665	274.022.602.225.097	251.057.261.182.811	1.459.675.93.853.440	839.270.200.488.435	862.193.246.349.194
CNAE	MEDIA_IRGE	MEDIA_DISTREG	MEDIA_DISTREG_med5	DISTREG_med10	DISTREG_med5	DISTREG_M10	DISTREG_M20	MEDIA_ESTRB	MEDIA_ESTN	MEDIA_ESTRATO
451	324.919.221.784.669	3.812.856.595.453.810	7.873.183.850.718.380	1.514.394.719.221.730	952.155.133.578.401	838.665.965.087.522	1.572.869.410.322.530	1.921.714.326.431.410	2.559.169.090.469.930	2.788.446.497.886.980
453	10.576.653.397.189	93.331.390.879.633	51.129.654.042.764.810	254.600.042.092.045	80.931.984.084.973	17.804.506.930.991	7.675.280.977.707	56.231.514.120.564	378.108.020.000	59.982.128.721.273
454	910.629.148.452.279	312.356.688.632.105	3.270.452.854.234.810	3.259.129.029.951	9.039.188.668	24.451.061.546.775	2.429.663.845.377	625.945.919.123.482	267.472.955.456.401	966.556.286.048.300
463	945.239.765.115.016	844.416.885.182.935	41.979.798.602.092	157.866.682.311.817	2.529.458.027.702.160	141.855.444.864.946	112.613.798.492.493	119.098.036.927.485	12.212.752.268.029	522.558.736.687.489
464	79.502.229.421.569	29.763.462.005	1.274.929.306.682.660	81.941.331.872.081	168.746.652.353.004	327.220.007.238.113	274.371.981.159.000	279.501.152.117.156	605.064.683.307.337	428.549.842.949.731
466	179.976.375.682.566	23.267.201.705	1.273.265.663.835.310	920.166.759.955.289	3.027.987.049.279.740	1.364.599.648.713.250	1.194.821.409.426.640	179.955.187.592.615	1.094.269.482.225.760	1.016.637.506.249.710
467	646.163.736.372.266	1.178.70.635.544.770	1.443.487.664.007	437.746.773.621.974	365.004.001.259.238	357.948.948.389.931	544.814.819.957.527	984.666.542.496.006	421.830.565.662.734	444.192.170.979.257
469	6.332.342.986.922.050	622.333.006.068.617	112.014.902.853.742	4.292.817.932.145.630	166.653.270.655.364	16.087.946.930.317	4.011.984.907.554.640	1.303.412.593.369.290	460.222.491.216.379	2.236.925.238.793.430
471	546.805.609.801.691	7.117.213.713.351	113.644.600.891.149	876.978.178.130	53.224.246.292.524	19.674.386.482.906	597.907.377.553	9.887.103.799.515	69.641.651	81.628.406.875.782
472	12.546.793.979.708	161.134.326.924.804	380.103.463.414	52.581.984.437	17.434.174.743.834	5.998.182.721.315	132.934.957.584	8.887.172.793.045	33.759.985.304.542	14.055.723.993.045
473	136.215.422.681.480	20.529.693.664	173.327.873.036.769	26.636.793.800.302	14.029.656.031.485	23.444.455.364.399	25.544.296.444.607	27.327.902.762.351	6.061.123.860.167	61.061.034.466.469
474	345.525.842.94.702	208.064.939.395.454	769.187.364.732.118	97.991.481.981.933	122.994.843.908.448	14.457.999.236.722	98.943.617.396.097	4.966.364.577.874	9.216.881.253.783	304.851.040.999.594
475	8.185.462.465.727	900.803.450.541	152.801.488.794.087	17.597.277.170.294	641.877.394	14.420.524.624.314	24.270.242.624.314	55.791.119.857.737	142.419.669.095.773	56.884.421.930.516
476	30.550.317.474.080	207.086.912.156.704	1.166.658.535.014	8.634.419.806.707	3.565.335.851.660	1.834.995.467.039	6.040.660.685.549	39.999.453.882.771	89.211.571.861.541	23.203.775.782.694
477	273.120.755.506.122	194.514.742.169	405.014.742.169	1.337.652.291.600	81.833.315.031.600	198.587.615.104.824	2.883.610.191.167	28.967.326.934.024	64.888.526.728.581	58.772.561.152.836
478	34.601.913.392.639	198.148.988.016.933	156.027.284.014	10.514.524.926.401	3.597.034.868.567	6.078.430.574.758	10.199.395.157.334	20.397.107.775.007	363.506.949.725.909	571.390.417.868.004
Total	676.056.565.611.717	546.735.578.133.249	944.302.288.176.083	489.152.265.924.020	468.614.921.089.752	201.260.019.967.989	493.137.077.755.007	387.911.061.113.709	363.506.949.725.909	571.390.417.868.004

5.1.5 ANO 2015 / UF 25 - PB / ESTF G

Mais uma vez podemos notar na Figura 5.9 que os modelos DISTEUCRID_REG_N10 e N20 têm resultados ineficientes, portanto são descartados da análise, como segue na Figura 5.10.

Na Figura 5.10, podemos notar também que o modelo IBGE é o que tem o melhor resultado dentre os demais.

Figura 5.9: EQM dos modelos de imputação: Ano 2015 / UF 25 - PB / ESTF G

Figura 5.10: EQM dos modelos de imputação (exceto DISTEUCID_REG_N10 e N20): Ano 2015 / UF 25 - PB / ESTF G

5.1.6 ANO 2015 / UF 32 - ES / ESTF A1

Seguindo a forma de análise feita nas Seções 5.1.1 a 5.1.5, todas as apresentações de resultados a seguir serão iniciadas com os resultados de todos os modelos e são retirados da análise os modelos que forem considerados discrepantes em relação aos demais até um ponto que fique mais clara a relação entre eles.

Conforme podemos observar na Figura 5.11, os modelos DISTEUCLID_REG_N10 e N20, DISTEUCLIDREG_medN5, DISTREG_N10 e N20, MÉDIA_ESTN e REG_ESTN foram os que tiveram os maiores erros quadráticos médios e, conseqüentemente, foram retirados do modelo, conforme pode ser visto na Figura 5.12.

Na Figura 5.12, podemos ver que o modelo REG_ESTF_REGIONAL foi o modelo com menor erro quadrático médio dentre os modelos analisados.

Figura 5.11: EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF A1

5.1.7 ANO 2015 / UF 32 - ES / ESTF A2

Agora, através da Figura 5.13, podemos notar que os modelos que apresentaram os maiores erros quadráticos médios são DISTEUCRID_REG_N10 e N20, MÉDIA_ESTN e REG_ESTN e UF e, com isso, esses modelos foram retirados dessa análise, juntamente com a CNAE 462 (Comércio atacadista de matérias-primas agrícolas e animais vivos) pelo mesmo motivo. Dessa forma, a Figura 5.14 não inclui esses modelos e nem essa CNAE.

Na Figura 5.14 podemos perceber que o modelo IBGE, MÉDIA_ESTBR e REG_ESTF são os que apresentam os melhores resultados.

[illegible]

Figura 5.13: *EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF A2*

5.1.8 ANO 2015 / UF 32 - ES / ESTF A3

A partir da Figura 5.15, é possível ver que os modelos DISTEUCLID_REG_N10 e N20, DISTREG_N20, MÉDIA_ESTN, REG_ESTF, ESTN e UF foram os modelos que apresentaram pior desempenho e foram retirados da análise. A CNAE 462 foi considerada um *outlier* frente às outras CNAE, conforme segue na Figura 5.16.

Com isso, pela Figura 5.16, podemos notar que os modelos IBGE, REG_ESTF_BRASIL e L4 são os que apresentaram os menores erros quadráticos médios.

5.1.9 ANO 2015 / UF 32 - ES / ESTF C

Ao analisarmos a Figura 5.17, vemos que os modelos DISTEUCLED_REG_N10 e N20, MÉDIA_ESTRATO e REG_ESTF estão com erros quadráticos médios altos e foram retirados da análise, como segue na Figura 5.18.

Quando analisamos a Figura 5.18, vemos que o modelo IBGE é o que tem o menor EQM.

Figura 5.18: *EQM dos modelos de imputação (exceto DISTEUCLED_REG_N10 e N20, MÉDIA ESTRATO e REG ESTF): Ano 2015 / UF 32 - ES / ESTF C*

5.1.10 ANO 2015 / UF 32 - ES / ESTF G

Na Figura 5.20, podemos notar que não estão presentes os modelos DISTEUCLID_REG_N10 e N20, DISTREG_N10 e N20, MÉDIA_ESTN, REG_ESTN, L4 e UF, pois apresentaram erros muito altos na Figura 5.19 e, logo, foram retirados da análise.

Podemos também notar na Figura 5.20 que os modelos IBGE, DISTEUCLIDREG_medN5 e N10, MÉDIA_UF e REG_ESTF são os que apresentaram os menores EQM dentre todos os modelos.

Figura 5.19: EQM dos modelos de imputação: Ano 2015 / UF 32 - ES / ESTF G

Figura 5.20: EQM dos modelos de imputação (exceto DISTEUCID_REG_N10 e N20, DISTREG_N10 e N20, MÉDIA ESTN, REG ESTN, L4 e UF): Ano 2015 / UF 32 - ES / ESTF G

5.2 Diferença percentual dos totais estimados x total real

A seguir, nas Tabelas 5.1 a 5.21, são disponibilizados os Totais de *RECEITA* e $\widehat{RECEITA}$ e as respectivas VP entre esses valores.

5.2.1 Modelos de imputação de médias de grupos

As Tabelas 5.2, 5.3, 5.4, 5.5, 5.6 e 5.7 apresentam os totais e as variações dos modelos derivados de médias de grupos. Neles podemos notar que as variações não são marcantes para os modelos cujo ESTF não varie muito. Já para os modelos MÉDIA_ESTN e MÉDIA_UF as variações apresentam impactos consideráveis.

Tabela 5.1: Totais de *RECEITA* estimado versus real e VP do modelo atual do IBGE

ANO	UF	$\widehat{RECEITA}$	<i>RECEITA</i>	VP (%)
2014	25 - PB	24.661.288.856	24.638.384.694	0,09
2014	32 - ES	95.702.820.700	95.817.904.914	0,12
2014	41 - PR	236.298.904.666	235.837.708.150	0,20
2014	43 - RS	218.659.158.075	218.222.556.331	0,20
2015	25 - PB	27.356.875.316	26.029.841.795	5,10
2015	32 - ES	98.255.774.784	99.256.864.660	1,01
2015	41 - PR	250.731.272.484	249.379.286.270	0,54
2015	43 - RS	229.866.663.639	230.759.908.487	0,39

Tabela 5.2: Totais de *RECEITA* estimado versus real e VP do modelo MÉDIA_ESTRATO

ANO	UF	$\widehat{RECEITA}$	<i>RECEITA</i>	VP (%)
2014	25 - PB	24.630.927.318	24.638.384.694	0,03
2014	32 - ES	96.213.926.124	95.817.904.914	0,41
2014	41 - PR	239.127.510.086	235.837.708.150	1,39
2014	43 - RS	215.974.597.689	218.222.556.331	1,03
2015	25 - PB	26.062.185.588	26.029.841.795	0,12
2015	32 - ES	98.888.481.347	99.256.864.660	0,37
2015	41 - PR	248.591.812.809	249.379.286.270	0,32
2015	43 - RS	233.377.098.923	230.759.908.487	1,13

Tabela 5.3: Totais de *RECEITA* estimado versus real e VP do modelo MÉDIA_ESTN

ANO	UF	$\widehat{RECEITA}$	<i>RECEITA</i>	VP (%)
2014	25 - PB	26.505.668.359	24.638.384.694	7,58
2014	32 - ES	103.099.945.195	95.817.904.914	7,60
2014	41 - PR	263.351.643.686	235.837.708.150	11,67
2014	43 - RS	226.384.176.738	218.222.556.331	3,74
2015	25 - PB	27.166.292.207	26.029.841.795	4,37
2015	32 - ES	104.205.331.414	99.256.864.660	4,99
2015	41 - PR	267.370.294.546	249.379.286.270	7,21
2015	43 - RS	243.270.039.618	230.759.908.487	5,42

Tabela 5.4: Totais de *RECEITA* estimado versus real e VP do modelo *MÉDIA_UF*

ANO	UF	<i>RECÊITA</i>	<i>RECEITA</i>	VP (%)
2014	25 - PB	26.359.547.072	24.638.384.694	6,99
2014	32 - ES	112.142.620.735	95.817.904.914	17,04
2014	41 - PR	265.517.160.925	235.837.708.150	12,58
2014	43 - RS	230.877.369.418	218.222.556.331	5,80
2015	25 - PB	27.467.086.215	26.029.841.795	5,52
2015	32 - ES	106.944.493.619	99.256.864.660	7,75
2015	41 - PR	275.887.624.007	249.379.286.270	10,63
2015	43 - RS	245.539.170.458	230.759.908.487	6,40

Tabela 5.5: Totais de *RECEITA* estimado versus real e VP do modelo *MÉDIA_L4ESTF*

ANO	UF	<i>RECÊITA</i>	<i>RECEITA</i>	VP (%)
2014	25 - PB	24.400.755.136	24.638.384.694	0,96
2014	32 - ES	96.107.667.847	95.817.904.914	0,30
2014	41 - PR	240.884.755.352	235.837.708.150	2,14
2014	43 - RS	216.045.234.511	218.222.556.331	1,00
2015	25 - PB	26.220.910.286	26.029.841.795	0,73
2015	32 - ES	98.754.833.670	99.256.864.660	0,51
2015	41 - PR	250.642.680.956	249.379.286.270	0,51
2015	43 - RS	232.999.587.325	230.759.908.487	0,97

Tabela 5.6: Totais de *RECEITA* estimado versus real e VP do modelo *MÉDIA_ESTREG*

ANO	UF	<i>RECÊITA</i>	<i>RECEITA</i>	VP (%)
2014	25 - PB	24.792.210.492	24.638.384.694	0,62
2014	32 - ES	95.293.763.465	95.817.904.914	0,55
2014	41 - PR	238.927.215.054	235.837.708.150	1,31
2014	43 - RS	215.541.498.489	218.222.556.331	1,23
2015	25 - PB	26.105.707.533	26.029.841.795	0,29
2015	32 - ES	97.740.631.313	99.256.864.660	1,53
2015	41 - PR	250.698.476.828	249.379.286.270	0,53
2015	43 - RS	231.716.243.113	230.759.908.487	0,41

Tabela 5.7: Totais de *RECEITA* estimado versus real e VP do modelo *MÉDIA_ESTBR*

ANO	UF	<i>RECÊITA</i>	<i>RECEITA</i>	VP (%)
2014	25 - PB	24.991.617.899	24.638.384.694	1,43
2014	32 - ES	95.097.114.396	95.817.904.914	0,75
2014	41 - PR	240.816.857.930	235.837.708.150	2,11
2014	43 - RS	216.957.594.167	218.222.556.331	0,58
2015	25 - PB	26.349.230.798	26.029.841.795	1,23
2015	32 - ES	97.566.097.385	99.256.864.660	1,70
2015	41 - PR	252.585.714.972	249.379.286.270	1,29
2015	43 - RS	232.897.090.125	230.759.908.487	0,93

5.2.2 Modelos de imputação por função de regressão

Aqui são apresentados os modelos que utilizam somente uma função regressão, conforme segue nas Tabelas 5.8, 5.9, 5.10, 5.11, 5.12 e 5.13.

Nessas tabelas, podemos notar, assim como na Seção 5.2.1, que as variações não são muito altas quando mudamos o modelo, mas não alteramos o ESTF. A exceção talvez esteja somente no estado do Espírito Santo para o modelo REG_ESTF cuja variação chega a 7%.

Tabela 5.8: Totais de RECEITA estimado versus real e VP do modelo REG_ESTF

ANO	UF	$\widehat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	24.657.983.969	24.638.384.694	0,08
2014	32 - ES	89.063.595.228	95.817.904.914	7,05
2014	41 - PR	247.302.380.985	235.837.708.150	4,86
2014	43 - RS	219.774.376.241	218.222.556.331	0,71
2015	25 - PB	26.050.696.727	26.029.841.795	0,08
2015	32 - ES	96.800.428.265	99.256.864.660	2,47
2015	41 - PR	248.058.577.007	249.379.286.270	0,53
2015	43 - RS	224.863.792.199	230.759.908.487	2,56

Tabela 5.9: Totais de RECEITA estimado versus real e VP do modelo REG_ESTN

ANO	UF	$\widehat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	23.154.496.386	24.638.384.694	6,02
2014	32 - ES	101.234.719.295	95.817.904.914	5,65
2014	41 - PR	220.114.930.370	235.837.708.150	6,67
2014	43 - RS	261.637.016.621	218.222.556.331	19,89
2015	25 - PB	28.021.075.201	26.029.841.795	7,65
2015	32 - ES	103.590.583.402	99.256.864.660	4,37
2015	41 - PR	232.871.769.365	249.379.286.270	6,62
2015	43 - RS	242.613.811.679	230.759.908.487	5,14

Tabela 5.10: Totais de RECEITA estimado versus real e VP do modelo REG_UF

ANO	UF	$\widehat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	28.975.016.496	24.638.384.694	17,60
2014	32 - ES	247.600.279.586	95.817.904.914	158,41
2014	41 - PR	196.549.102.853	235.837.708.150	16,66
2014	43 - RS	123.102.263.832	218.222.556.331	43,59
2015	25 - PB	30.745.072.330	26.029.841.795	18,11
2015	32 - ES	185.701.031.054	99.256.864.660	87,09
2015	41 - PR	190.647.517.378	249.379.286.270	23,55
2015	43 - RS	160.898.959.871	230.759.908.487	30,27

Tabela 5.11: *Totais de RECEITA estimado versus real e VP do modelo REG_L4*

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	24.072.656.185	24.638.384.694	2,30
2014	32 - ES	99.546.206.394	95.817.904.914	3,89
2014	41 - PR	239.923.357.026	235.837.708.150	1,73
2014	43 - RS	216.026.947.211	218.222.556.331	1,01
2015	25 - PB	27.168.670.999	26.029.841.795	4,38
2015	32 - ES	98.545.355.287	99.256.864.660	0,72
2015	41 - PR	250.068.248.868	249.379.286.270	0,28
2015	43 - RS	231.103.715.253	230.759.908.487	0,15

Tabela 5.12: *Totais de RECEITA estimado versus real e VP do modelo REG_ESTF_REGIONAL*

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	24.375.448.570	24.638.384.694	1,07
2014	32 - ES	97.706.771.414	95.817.904.914	1,97
2014	41 - PR	236.777.487.064	235.837.708.150	0,40
2014	43 - RS	214.414.849.121	218.222.556.331	1,74
2015	25 - PB	25.982.533.300	26.029.841.795	0,18
2015	32 - ES	96.493.383.076	99.256.864.660	2,78
2015	41 - PR	250.114.837.227	249.379.286.270	0,29
2015	43 - RS	229.325.390.011	230.759.908.487	0,62

Tabela 5.13: *Totais de RECEITA estimado versus real e VP do modelo REG_ESTF_BRASIL*

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	24.487.681.836	24.638.384.694	0,61
2014	32 - ES	96.111.157.575	95.817.904.914	0,31
2014	41 - PR	238.203.160.553	235.837.708.150	1,00
2014	43 - RS	214.697.144.748	218.222.556.331	1,62
2015	25 - PB	26.019.173.537	26.029.841.795	0,04
2015	32 - ES	97.349.630.572	99.256.864.660	1,92
2015	41 - PR	249.141.890.654	249.379.286.270	0,10
2015	43 - RS	231.168.626.163	230.759.908.487	0,18

5.2.3 Modelos utilizando cálculo da distância do IBGE entre as empresas

Nesta seção serão apresentados os totais dos modelos de imputação cujo cálculo de distância utilizado foi o mesmo adotado no modelo de imputação do IBGE nas Tabelas 5.14 a 5.17. Essa análise é importante para que seja possível verificar o impacto dos resultados dos modelos testados no que se refere às estimativas da pesquisa quanto aos domínios abordados.

Nas referidas tabelas podemos notar que os modelos DISTREG_N10 e DISTREG_N20 não tiveram resultados satisfatórios ao contrário dos modelos DISTREG_medN5 e DISTREG_medN10 nas quais apenas a UF do Espírito Santo em 2014 não apresentou bom resultado.

Tabela 5.14: Totais de RECEITA estimado versus real e VP do modelo DISTREG_N10

ANO	UF	RECÊITA	RECEITA	VP (%)
2014	25 - PB	25.917.906.996	24.638.384.694	5,19
2014	32 - ES	346.604.928.682	95.817.904.914	261,73
2014	41 - PR	261.702.260.443	235.837.708.150	10,97
2014	43 - RS	228.499.077.969	218.222.556.331	4,71
2015	25 - PB	26.935.499.616	26.029.841.795	3,48
2015	32 - ES	103.389.326.821	99.256.864.660	4,16
2015	41 - PR	267.644.875.293	249.379.286.270	7,32
2015	43 - RS	248.292.301.602	230.759.908.487	7,60

Tabela 5.15: Totais de RECEITA estimado versus real e VP do modelo DISTREG_N20

ANO	UF	RECÊITA	RECEITA	VP (%)
2014	25 - PB	23.906.990.474	24.638.384.694	2,97
2014	32 - ES	230.635.261.224	95.817.904.914	140,70
2014	41 - PR	267.637.804.320	235.837.708.150	13,48
2014	43 - RS	228.180.542.041	218.222.556.331	4,56
2015	25 - PB	29.675.175.669	26.029.841.795	14,00
2015	32 - ES	118.482.612.353	99.256.864.660	19,37
2015	41 - PR	266.691.368.935	249.379.286.270	6,94
2015	43 - RS	247.671.587.245	230.759.908.487	7,33

Tabela 5.16: Totais de RECEITA estimado versus real e VP do modelo DISTREG_medN5

ANO	UF	RECÊITA	RECEITA	VP (%)
2014	25 - PB	24.070.683.370	24.638.384.694	2,30
2014	32 - ES	121.264.699.935	95.817.904.914	26,56
2014	41 - PR	240.130.265.681	235.837.708.150	1,82
2014	43 - RS	213.788.397.647	218.222.556.331	2,03
2015	25 - PB	25.936.636.113	26.029.841.795	0,36
2015	32 - ES	96.394.270.881	99.256.864.660	2,88
2015	41 - PR	252.095.197.810	249.379.286.270	1,09
2015	43 - RS	228.673.868.869	230.759.908.487	0,90

Tabela 5.17: Totais de RECEITA estimado versus real e VP do modelo DISTREG_medN10

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	23.900.506.153	24.638.384.694	2,99
2014	32 - ES	165.274.986.520	95.817.904.914	72,49
2014	41 - PR	242.858.896.346	235.837.708.150	2,98
2014	43 - RS	213.492.892.744	218.222.556.331	2,17
2015	25 - PB	25.612.025.691	26.029.841.795	1,61
2015	32 - ES	96.607.171.544	99.256.864.660	2,67
2015	41 - PR	246.594.158.196	249.379.286.270	1,12
2015	43 - RS	229.217.839.577	230.759.908.487	0,67

5.2.4 Modelos utilizando distância euclidiana entre as empresas

Já nas Tabelas 5.18, 5.19, 5.20 e 5.21, são apresentados os totais dos modelos de imputação em que o cálculo da distância entre as empresas baseia-se na distância euclidiana. Nelas podemos notar o mesmo comportamento dos modelos com o cálculo de distância do IBGE, ou seja, os modelos DISTREG_N10 e DISTREG_N20 não tiveram resultados satisfatórios enquanto os modelos DISTREG_medN5 e DISTREG_medN10 tiveram, exceto para o estado do Espírito Santo em 2014.

Tabela 5.18: Totais de RECEITA estimado versus real e VP do modelo DISTEUCLID_REG_N10

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	63.915.850.150	24.638.384.694	159,42
2014	32 - ES	261.999.781.404	95.817.904.914	173,44
2014	41 - PR	198.062.214.702	235.837.708.150	16,02
2014	43 - RS	139.020.667.443	218.222.556.331	36,29
2015	25 - PB	- 562.736.156	26.029.841.795	102,16
2015	32 - ES	14.606.531.992	99.256.864.660	85,28
2015	41 - PR	226.108.111.711	249.379.286.270	9,33
2015	43 - RS	177.203.341.953	230.759.908.487	23,21

Tabela 5.19: Totais de RECEITA estimado versus real e VP do modelo DISTEUCLID_REG_N20

ANO	UF	$\hat{RECEITA}$	RECEITA	VP (%)
2014	25 - PB	13.560.232.949	24.638.384.694	44,96
2014	32 - ES	68.883.846.571	95.817.904.914	28,11
2014	41 - PR	263.091.318.960	235.837.708.150	11,56
2014	43 - RS	213.840.576.545	218.222.556.331	2,01
2015	25 - PB	56.725.640.263	26.029.841.795	117,93
2015	32 - ES	- 1.200.650.326.875	99.256.864.660	1309,64
2015	41 - PR	609.145.296.066	249.379.286.270	144,26
2015	43 - RS	323.001.203.045	230.759.908.487	39,97

Tabela 5.20: Totais de RECEITA estimado versus real e VP do modelo DISTEUCIDREG_medN5

ANO	UF	RECÊITA	RECEITA	VP (%)
2014	25 - PB	23.294.433.828	24.638.384.694	5,45
2014	32 - ES	125.532.112.940	95.817.904.914	31,01
2014	41 - PR	242.146.951.572	235.837.708.150	2,68
2014	43 - RS	208.641.837.129	218.222.556.331	4,39
2015	25 - PB	26.156.662.678	26.029.841.795	0,49
2015	32 - ES	97.474.447.489	99.256.864.660	1,80
2015	41 - PR	249.421.194.623	249.379.286.270	0,02
2015	43 - RS	227.093.164.542	230.759.908.487	1,59

Tabela 5.21: Totais de RECEITA estimado versus real e VP do modelo DISTEUCIDREG_medN10

ANO	UF	RECÊITA	RECEITA	VP (%)
2014	25 - PB	23.161.166.412	24.638.384.694	6,00
2014	32 - ES	163.548.315.930	95.817.904.914	70,69
2014	41 - PR	233.270.175.542	235.837.708.150	1,09
2014	43 - RS	207.288.500.572	218.222.556.331	5,01
2015	25 - PB	24.566.066.428	26.029.841.795	5,62
2015	32 - ES	96.412.032.826	99.256.864.660	2,87
2015	41 - PR	244.507.931.436	249.379.286.270	1,95
2015	43 - RS	223.705.236.400	230.759.908.487	3,06

No próximo capítulo serão apresentadas as conclusões acerca dos resultados expostos neste capítulo.

Capítulo 6

Conclusões

6.1 Análise dos resultados

Para auxiliar na leitura das análises dos resultados foram criadas as Figuras 6.1 e 6.2, onde se encontram os resultados dos EQM dos modelos abertos por UF e ESTF nas empresas de 2014 e 2015, respectivamente; as Figuras 6.3 e 6.4, que apresentam os resultados dos EQM abertos apenas por UF nos anos de 2014 e 2015, respectivamente; e a Figura 6.5, a qual apresenta os resultados das variações percentuais de forma resumida.

Nas Figuras de 6.1 a 6.4 estão destacados em verde os modelos com menores erros quadráticos médios, considerando os domínios da figura em questão, sendo destacados com maior intensidade de verde e com um “x” triplo o modelo com o MENOR EQM. Os resultados cuja raiz quadrada do EQM varie em no máximo 15% do menor EQM também tiveram a mesma marcação; se a variação for de entre 15% e 30%, então marcou-se com um “x” duplo e um tom de verde intermediário; e variações entre 30% e 50%, marcou-se apenas um “x” com um verde mais brando. Variações superiores a 50% não foram sinalizadas (espaços em branco). Cada coluna da figura representa um domínio analisado.

Já na Figura 6.5, as menores variações estão destacadas em verde e as maiores em vermelho. Quanto mais intenso o verde, menor a VP e quanto maior a intensidade de vermelho, maior a VP. Nessa figura, cada linha representa um domínio analisado.

Inicialmente, iremos considerar as Figuras 6.1 e 6.2, pois nelas é possível verificar que o modelo de imputação do IBGE é de longe o melhor em comparação com os demais para a base utilizada neste trabalho.

Para os estados da Paraíba e do Espírito Santo, tanto em 2014 quanto em 2015, podemos perceber que o modelo do IBGE perde um pouco de eficiência nos estratos finais amostrados (A1, A2 e A3), nos quais os modelos com melhor desempenho são os de média dos estratos a nível

Brasil, da regressão dos estratos da região geográfica da empresa e regressão do estrato a nível Brasil (MÉDIA_ESTBR, REG_ESTF_BRASIL e REG_ESTF_REGIONAL, respectivamente). Isto significa que os modelos com abrangência mais específica, contudo com maior número de empresas possível, se adequa bem ao recorte.

Para o estrato final C do estado da Paraíba podemos ver que todas as variações dos modelos de imputação através da função distância utilizada pelo IBGE (DISTREG) também têm bons resultados, enquanto no Espírito Santo o modelo do IBGE se destaca isoladamente, assim como no estrato final G da Paraíba. Já no ESTF G do Espírito Santo os resultados nos dois anos concordam apenas quanto ao bom desempenho do modelo do IBGE.

No estado do Paraná, o modelo de média dos estratos da linha de divulgação (MÉDIA_L4ESTF) também apresentou bons resultados juntamente com o modelo do IBGE para os estratos finais A1 e A2, demonstrando mais uma vez que um modelo com abrangência mais específica e com um volume maior de empresas pode ser um bom modelo. Já para os estratos finais A3 e C, o modelo do IBGE é o único com resultados satisfatórios, enquanto para o ESTF G o modelo da mediana dos 10 vizinhos mais próximos pela função distância do IBGE (DISTREG_medN10) e o modelo da média geral da UF (MÉDIA_UF) são aqueles com os melhores desempenhos se analisarmos os dois anos em conjunto.

Já no estado do Rio Grande do Sul, os resultados apontam que para os estratos finais C e G, o modelo do IBGE é novamente o que possui melhor desempenho. Contudo, nos demais estratos finais os modelos regressão da linha de divulgação e mediana dos 5 e 10 vizinhos mais próximos pela função distância do IBGE (REG_L4, DISTREG_medN5 e DISTREG_medN10, respectivamente) também têm bons resultados quanto aos erros quadráticos médios.

Ou seja, de maneira geral, os estratos finais A1, A2 e A3 podem apresentar mais comumente modelos de imputação tão satisfatórios quanto os do IBGE. Já nos estratos finais que contemplam as grandes empresas isso não é tão comum, com exceção do estado da Paraíba cujo número e porte das empresas é menor do que os dos demais estados utilizados na análise.

Também é possível notar que os modelos mediana dos 5 e 10 vizinhos mais próximos pela função distância do IBGE (DISTREG_medN5 e DISTREG_medN10, respectivamente) tiveram alguns resultados satisfatórios, o que era esperado por serem uma extensão do modelo do IBGE, porém com algumas adaptações. Os modelos de médias do estrato da região geográfica e da linha de divulgação da empresa (MÉDIA_ESTREG e MÉDIA_L4ESTF, respectivamente) vêm em seguida e a vantagem em sua utilização está no fato de não dependerem de variáveis auxiliares.

Como atualmente a relação dos valores de *RECEITA* das empresas estão cada vez menos

relacionados diretamente ao número de funcionários destas, fazer uma estimativa da *RECEITA* utilizando a informação de *PO* como variável auxiliar não altera necessariamente a qualidade das estimativas. Essa relação torna-se ainda mais conflituosa quando notamos que em alguns casos as informações prestadas à RAIS divergem consideravelmente das informadas na pesquisa. Dessa forma, o modelo implementado pelo IBGE continua sendo a melhor opção dentre os demais.

Figura 6.1: Resultados do EQM em 2014

	2015																	
	25 - PB					32 - ES					41 - PR				43 - RS			
	A1	A2	A3	C	G	A1	A2	A3	C	G	A1	A2	A3	C	G			
IBGE	xxx		xxx	xxx	xxx													
DISTEUCUD_REG_N10																		
DISTEUCUD_REG_N20																		
DISTEUCUDREG_medN10																		
DISTREG_medN10	x	xxx		x	xxx					xxx			x					
DISTREG_medN5	x	xx		x	x					xxx			xxx	xxx	xxx			
DISTREG_N10	xx			x	xx					xxx			xxx	xxx	xxx			
DISTREG_N20																		
MÉDIA_ESTBR	xx				xx													
MÉDIA_ESTN					x													
MÉDIA_ESTRATO	x				x					x			xxx					
MÉDIA_ESTREG	x	x			x					x			x	xxx				
MÉDIA_L4ESTF					x					xxx			xxx	xxx				
MÉDIA_UF					xx					xx			xxx	xxx				
REG_ESTF					xx					xx			xx	xx				
REG_ESTF_BRASIL					x													
REG_ESTF_REGIONAL	xx	xx		xx	xxx													
REG_ESTN					xxx					x								
REG_L4																		
REG_UF					xxx					xx				xx				

xxx

xx

x

Até 15% da raiz quadrada do menor EOM

De 15 a 30% da raiz quadrada do menor EOM

De 30 a 50% da raiz quadrada do menor EOM

Figura 6.2: Resultados do EQM em 2015

Ao analisarmos os resultados dos erros quadráticos médios consolidados por UF, desconsiderando a abertura por ESTF, podemos notar que o modelo de imputação do IBGE se mantém com melhor resultado em relação aos demais. A exceção é o estado da Paraíba, no ano de 2015, no qual todas as variações do modelo que utiliza uma função distância do IBGE (DISTREG) e o modelo de regressão do estrato final da região geográfica da empresa (REG_ESTF_REGIONAL) destacam-se em comparação ao modelo do IBGE, conforme as Figuras 6.3 e 6.4.

	2014			
	25 - PB	32 - ES	41 - PR	43 - RS
IBGE	xxx	xxx	xxx	xxx
DISTEUCLID_REG_N10				
DISTEUCLID_REG_N20				
DISTEUCLIDREG_medN10				
DISTEUCLIDREG_medN5				
DISTREG_medN10	x			
DISTREG_medN5				
DISTREG_N10				
DISTREG_N20	x			
MÉDIA_ESTBR				
MÉDIA_ESTN				
MÉDIA_ESTRATO				
MÉDIA_ESTREG	x			
MÉDIA_L4ESTF	xx			
MÉDIA_UF				
REG_ESTF				
REG_ESTF_BRASIL				
REG_ESTF_REGIONAL				
REG_ESTN				
REG_L4				
REG_UF				

xxx	Até 15% da raiz quadrada do menor EQM
xx	De 15 a 30% da raiz quadrada do menor EQM
x	De 30 a 50% da raiz quadrada do menor EQM

Figura 6.3: Resultados do EQM em 2014 consolidados por UF

Com isso, podemos verificar que o modelo de imputação do IBGE continua sendo o mais eficiente para os estados do Espírito Santo, Paraná e Rio Grande do Sul, considerando as premissas adotadas neste trabalho, com o recorte adotado sob a perspectiva do EQM, enquanto para o estado da Paraíba outros modelos de imputação podem ser admitidos sem perda de eficiência.

Talvez essa diferença de resultados seja decorrente da forma como a classificação econômica das

	2015			
	25 - PB	32 - ES	41 - PR	43 - RS
IBGE		xxx	xxx	xxx
DISTEUCLID_REG_N10				
DISTEUCLID_REG_N20				
DISTEUCLIDREG_medN10				
DISTEUCLIDREG_medN5				
DISTREG_medN10	xxx			
DISTREG_medN5	xxx			
DISTREG_N10	xx			
DISTREG_N20	xxx			
MÉDIA_ESTBR				
MÉDIA_ESTN				
MÉDIA_ESTRATO				
MÉDIA_ESTREG				
MÉDIA_L4ESTF				
MÉDIA_UF				
REG_ESTF				
REG_ESTF_BRASIL	x			
REG_ESTF_REGIONAL	xxx			
REG_ESTN				
REG_L4	x			
REG_UF	x			

xxx	Até 15% da raiz quadrada do menor EQM
xx	De 15 a 30% da raiz quadrada do menor EQM
x	De 30 a 50% da raiz quadrada do menor EQM

Figura 6.4: Resultados do EQM em 2015 consolidados por UF

empresas se agrega de maneira distinta entre esses estados.

Apesar do estado do Espírito Santo ter a mesma agregação de CNAE da Paraíba, o restante de sua região geográfica tem uma agregação diferente e, como muitos dos modelos de imputação utilizam esses outros estados desta região geográfica para suas estimativas, as imputações podem ficar com a característica dos demais estados.

Entretanto, quando analisamos os resultados sob a perspectiva das variações percentuais dos totais reais contra os estimados, apresentados na Figura 6.5, podemos perceber que os modelos de imputação de média e regressão dos agrupamentos ESTF, ESTF_REGIONAL e ESTF_BRASIL (estrato, estrato da região geográfica e estratos a nível Brasil, respectivamente) não divergem de forma significativa dos resultados originais. O mesmo pode ser verificado nos modelos da mediana dos 5 e 10 vizinhos mais próximos pela função distância do IBGE (DISTREG_medN5 e DIS-

TREG_medN10, respectivamente) e pela função distância euclidiana (DSTEUCIDREG_medN5 e DSTEUCIDREG_medN10, respectivamente) com exceção do estado do Espírito Santo no ano de 2014, enquanto nos demais modelos podem haver impactos expressivos nas estimativas finais.

ANO	UF	IBGE	MÉDIA_ESTRATO	MÉDIA_ESTN	MÉDIA_UF	MÉDIA_L4ESTF	MÉDIA_ESTREG	MÉDIA_ESTBR	REG_ESTF	REG_ESTN	REG_UF	REG_L4	REG_ESTF_REGIONAL	REG_ESTF_BRASIL	DISTREG_N10	DISTREG_N20	DISTREG_medN5	DISTREG_medN10	DISTEUCLID_REG_N10	DISTEUCLID_REG_N20	DISTEUCLIDREG_medN5	DISTEUCLIDREG_medN10
2014	25 - PB	0,09	0,03	7,58	6,99	0,96	0,62	1,43	0,08	6,02	17,60	2,30	1,07	0,61	5,19	2,97	2,30	2,99	159,42	44,96	5,45	6,00
2014	32 - ES	0,12	0,41	7,60	17,04	0,30	0,55	0,75	7,05	5,65	158,41	3,89	1,97	0,31	261,73	140,70	26,56	72,49	173,44	28,11	31,01	70,69
2014	41 - PR	0,20	1,39	11,67	12,58	2,14	1,31	2,11	4,86	6,67	16,66	1,73	0,40	1,00	10,97	13,48	1,82	2,98	16,02	11,56	2,68	1,09
2014	43 - RS	0,20	1,03	3,74	5,80	1,00	1,23	0,58	0,71	19,89	43,59	1,01	1,74	1,62	4,71	4,56	2,03	2,17	36,29	2,01	4,39	5,01
2015	25 - PB	5,10	0,12	4,37	5,52	0,73	0,29	1,23	0,08	7,65	18,11	4,38	0,18	0,04	3,48	14,00	0,36	1,61	102,16	117,93	0,49	5,62
2015	32 - ES	1,01	0,37	4,99	7,75	0,51	1,53	1,70	2,47	4,37	87,09	0,72	2,78	1,92	4,16	19,37	2,88	2,67	85,28	1309,64	1,80	2,87
2015	41 - PR	0,54	0,32	7,21	10,63	0,51	0,53	1,29	0,53	6,62	23,55	0,28	0,29	0,10	7,32	6,94	1,09	1,12	9,33	144,26	0,02	1,95
2015	43 - RS	0,39	1,13	5,42	6,40	0,97	0,41	0,93	2,56	5,14	30,27	0,15	0,62	0,18	7,60	7,33	0,90	0,67	23,21	39,97	1,59	3,06

VP (%)

Figura 6.5: Resultados do VP em 2014 e 2015 consolidados por UF

6.2 Considerações finais

Decorrente das análises dos resultados feitas na Seção 6.1, não há fortes evidências que incentivem a alteração do modelo de imputação atualmente utilizado na PAC-IBGE. O modelo aplicado pelo IBGE consegue na grande maioria dos casos discutidos no presente trabalho apresentar desempenho melhor ou igual aos dos demais modelos investigados.

Contudo, também foi possível verificar que a utilização de outros modelos, apesar de apresentarem alto EQM, não distorcem as estimativas finais dos totais da variável *RECEITA*, portanto não sendo totalmente desaconselhável a sua implementação em decorrência do seu impacto nesses casos.

Assim, na eventual necessidade de imputação do questionário de alguma empresa e não querendo o uso de uma variável auxiliar no modelo a ser utilizado, os modelos de média dos estratos da região geográfica da empresa (MÉDIA_ESTREG) e da linha de divulgação da empresa (MÉDIA_L4ESTF) podem ser boas opções para tais casos, pois tanto na análise por EQM quanto por VP apresentaram bons resultados e não fazem uso de variáveis auxiliares em sua implementação. Os modelos da mediana dos 5 e 10 vizinhos mais próximos da empresa (DISTREG_medN5 e DISTREG_medN10, respectivamente) também tiveram resultados satisfatórios em ambas as análises e poderiam ser boas opções no caso da falta do modelo do IBGE.

6.3 Sugestões para trabalhos futuros

Os modelos de imputação utilizados neste trabalho não foram empregados à exaustão, ou seja, há ainda vários outros modelos que poderiam ter sido testados para avaliar sua eficiência frente ao atual. Não obstante, o estudo e a comparação de outros novos modelos viria a ser uma contribuição muito bem-vinda à PAC-IBGE, como também agregaria valor ao presente trabalho.

No entanto, devido à limitação de escopo da dissertação aqui descrita, não foi possível tratar de todos os temas alusivos às não-respostas em pesquisas, sobretudo as econômicas. Sendo assim, o estudo concentrou-se em analisar apenas os temas diretamente relacionados com os Métodos de Imputação.

Desta maneira, a seguir estão descritas algumas sugestões para trabalhos futuros que, assim como as análises aqui apresentadas, também são de grande relevância para as pesquisas econômicas do IBGE.

6.3.1 Tratamento para não-respostas e mortes de empresas

Torna-se necessário estudar tratamentos para as empresas que foram não-respostas ou morreram e seus impactos na pesquisa que não estejam diretamente relacionados com a necessidade de imputação, como por exemplo: o motivo delas serem classificadas como tal e possíveis correções de cadastro na base de dados que impacta na seleção das amostras.

6.3.2 Calibração

De acordo com Pessoa [2005], o uso de fatores de calibração é uma prática comum nos órgãos produtores de estatística com o objetivo de melhorar as estimativas de suas pesquisas.

Ainda assim, o calibrador em alguns casos pode causar um alto peso concentrado em uma unidade, sendo exigido reavaliar o seu uso para tais casos, havendo eventualmente a necessidade de imputação de algum questionário do estrato para redistribuição dos pesos amostrais.

Sendo assim, há a necessidade de se estudar novos calibradores e verificar se o atual método de calibração encontra-se, ainda, aceitável. Outra proposta para futuras pesquisas seria alterar as variáveis utilizadas na calibração e avaliar se essa mudança traria melhoria nas estimativas.

6.3.3 Plano amostral

Redefinir o desenho amostral também seria outro estudo oportuno e neste caso, dois desenhos de amostras estratificadas deveriam ser propostos: o primeiro seria condicional a um estudo da atividade econômica e, conseqüentemente, a redefinição do porte das empresas, ou seja, redefinir o corte do número de pessoas ocupadas para estipular o que seria uma empresa pertencente ao A1, A2, A3 e estrato certo; ou manter a faixa de pessoas ocupadas iguais a todas as atividades econômicas, contudo alterando o tamanho da faixa atual, ou seja, aumentando ou diminuindo o número de pessoas ocupadas nas empresas para determinar se elas pertencem ao estrato A1.

6.3.4 Avaliar impacto das imputações na variância do estimador

Toda imputação realizada possui erros associados e, atualmente, tais erros não são levados em consideração no cálculo das estimativas das pesquisas econômicas. Todavia, todo erro associado à estimativa calculada deveria ser avaliado e informado. Logo, o estudo do impacto da inclusão dos erros associados à imputação é um tema bastante importante para a PAC.

As propostas aqui apresentadas, acrescidas de outras inerentes ao processo de avaliação do que aqui ora se apresenta como um todo, podem vir a ser alvo de uma continuidade de estudos, haja vista o vínculo deste pesquisador com o IBGE e, em decorrência disto, as próprias necessidades da instituição quanto à aferição e melhoria dos seus métodos de pesquisa. De todo modo, acredita-se que o trabalho até aqui desenvolvido esteja em consonância com as premissas institucionais de pesquisa do IBGE, bem como com o grau de exigência do programa que acolheu este pesquisador e permitiu, com a anuência de sua orientadora, a sua realização.

Anexo A

Questionário Simplificado da PAC

Instituto Brasileiro de Geografia e Estatística
Diretoria de Pesquisas
Coordenação de Serviços e Comércio
Gerência de Pesquisas

PESQUISA ANUAL DE COMÉRCIO - 2015

QUESTIONÁRIO SIMPLIFICADO

www.ibge.gov.br
ibge@ibge.gov.br

01	IDENTIFICAÇÃO DO QUESTIONÁRIO (Uso do Órgão Regional)		
01	CÓDIGO DO MUNICÍPIO DA UC		
UF		MUNICÍPIO	DIST/SUBDIST
02	CADASTRO DO TÉCNICO DE PESQUISAS	03	CONTROLE
			1

OBRIGATORIEDADE E SIGILO DAS INFORMAÇÕES - a legislação vigente, de acordo com o Decreto Federal nº 73.177 de 20 de novembro de 1973 e a Lei nº 5.534 de 14 de novembro de 1968, modificada pela Lei nº 5.878 de 11 de maio de 1973, dispõe sobre a obrigatoriedade e sigilo das informações coletadas pelo IBGE, as quais se destinam, exclusivamente, a fins estatísticos e não poderão ser objeto de certidão e nem terão eficácia jurídica como meio de prova.

PRAZO DE ENTREGA - conforme descrito no recibo entregue pelo Técnico de Pesquisas.

PROPÓSITO DA PESQUISA ANUAL DE COMÉRCIO - coletar dados econômico-financeiros necessários à formulação de políticas públicas e programas sociais, bem como à estimação pelo Sistema de Contas Nacionais dos agregados macroeconômicos, em especial o Produto Interno Bruto - PIB.

I - INFORMAÇÕES CADASTRAIS

02	IDENTIFICAÇÃO DA SEDE DA EMPRESA		
PAC		PAC	
CNPJ da empresa		Sufixo	DV
03	DADOS CADASTRAIS		
SEDE DA EMPRESA (MATRIZ)			
01	FIRMA OU RAZÃO SOCIAL		
02	NOME FANTASIA (se não possuir, registre S/D)		02A SITE DA EMPRESA (se não possuir, registre N/T)
03	LOGRADOURO (rua, avenida, rodovia, etc.)	04 NÚMERO	05 COMPLEMENTO (bloco, grupo, andar, sala, km, etc.)
06	BAIRRO/DISTRITO/LOCALIDADE	13 NOME DO MUNICÍPIO	12 SIGLA UF 07 CEP
08	DDD 09 TELEFONE	10 RAMAL 11 FAX	14 E-MAIL DO RESPONSÁVEL PELA EMPRESA
15	DESCRIÇÃO DA PRINCIPAL ATIVIDADE (MAIOR RECEITA) DA EMPRESA		16 CÓDIGO CNAE 2.0
UNIDADE DE COLETA			
Endereço da Unidade da Empresa (Sede/Filial/Grupo Empresarial) designada a preencher o questionário. Não informar endereço do contador terceirizado/contratado			
17	SUFIXO / DV - DA UNIDADE DE COLETA	18 LOGRADOURO (rua, avenida, rodovia, etc.)	19 NÚMERO 20 COMPLEMENTO (bloco, grupo, andar, sala, km, etc.)
21	BAIRRO/DISTRITO/LOCALIDADE	28 NOME DO MUNICÍPIO	27 SIGLA UF 22 CEP
23	DDD 24 TELEFONE	25 RAMAL 26 FAX	29 E-MAIL DO RESPONSÁVEL PELO PREENCHIMENTO
04	DADOS CADASTRAIS COMPLEMENTARES		
01	SITUAÇÃO CADASTRAL EM 31-12-2015	02	MUDANÇAS NA ESTRUTURA DA EMPRESA NO ANO DE 2015
03	CNPJ DE LIGAÇÃO DA EMPRESA		
01 - Em operação 03 - Paralisada com informação de receita 04 - Extinta com informação de receita		00 - Não houve mudanças 01 - Fusão ou cisão total 02 - Cisão parcial 03 - Incorporação de outra(s) empresa(s) 03A - Incorporada por outra empresa 06 - Alteração de CNPJ por outros motivos (esclareça em "OBSERVAÇÕES")	
01A	DATA DA OCORRÊNCIA para os itens 03 e 04	Sucessora/Antecessora - Arrendatária/Arrendada	
MÊS ANO		Sucessora/Antecessora - Arrendatária/Arrendada	
01 02			
04	PRINCIPAIS MERCADORIAS REVENDIDAS POR ORDEM DE PREDOMINÂNCIA NA RECEITA, INCLUSIVE POR REPRESENTAÇÃO COMERCIAL	05	CLASSE PRINCIPAL DE COMÉRCIO (predominância na receita)
1) 2) 3)		1 - Comércio Varejista 2 - Comércio Atacadista 3 - Representação Comercial	
		06	
		FORMA DE TRIBUTAÇÃO UTILIZADA PELA EMPRESA	
		1 - Lucro Real 2 - Lucro Presumido ou Arbitrado 3 - Sistema "Simples Nacional" 4 - Imune ou Isenta	

II - INFORMAÇÕES DA EMPRESA

ATENÇÃO : Unidade de Investigação: neste questionário devem ser registrados os dados da empresa como um todo.
Informações contábeis: devem se referir às de competência do ano civil (janeiro a dezembro) e serem prestadas de acordo com a Legislação Societária.
Forma de preenchimento: registre os dados, de forma legível, a caneta. O preenchimento dos valores deve ser em **REAIS e SEM CENTAVOS.**
Antes de preencher o questionário, leia as instruções em anexo.

05	NÚMERO DE PESSOAS OCUPADAS	CÓD	EM 31-03	CÓD	EM 30-06	CÓD	EM 30-09	CÓD	EM 31-12
	Pessoal Assalariado	01		05		09		13	
	Pessoal Não-Assalariado								
	Proprietário e sócios com atividade na empresa	02		06		10		14	
	Membros da família sem remuneração	03		07		11		15	
	Total	04		08		12		16	

06	DEMONSTRATIVO DA RECEITA NO ANO	CÓD	VALORES EM REAIS
	Receita bruta		
	Revenda de mercadorias	17	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Comissões sobre vendas de representação comercial e agentes do comércio	17A	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Outras atividades (indústria, serviços, transporte, etc.)	18	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Deduções (ICMS, ISS, PIS/PASEP, SIMPLES NACIONAL, etc.) (-)	19	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Receita líquida (Cód. 17 + 17A + 18 - 19)	20	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Demais receitas operacionais e outras receitas	21	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Total (Cód. 20 + 21)	22	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00

07	COMPRAS, ESTOQUES E CUSTO DAS MERCADORIAS REVENDIDAS (Não inclua ICMS e o IPI quando recuperáveis. Preencher somente se o cód.17 estiver preenchido)	CÓD	VALORES EM REAIS
	Compras de mercadorias para revenda no ano	23	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Estoque em 31-12-2014	24	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Estoque em 31-12-2015	25	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Custo das mercadorias revendidas (valor contábil apurado e levado à DRE - Cód. 23 + 24 - 25)	25A	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00

08	GASTOS COM PESSOAL NO ANO Total pago ou devido	CÓD	VALORES EM REAIS
	Salários e outras remunerações (inclua 13º salário, férias, horas extras, comissões, participação nos lucros pagas aos empregados, honorários, etc.) - não incluir bolsa paga a estagiários (Cód. 39)	26	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Retiradas pró-labore do proprietário ou sócios com atividade na empresa (vide instrução)	27	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Contribuições para a Previdência Social (parte da empresa, ou seja, não inclua a contribuição descontada dos salários dos empregados, nem os recolhimentos relativos aos autônomos)	28	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	FGTS (se tiver empregados e não preencher, justificar em OBSERVAÇÕES) - não incluir os 50% do FGTS pagos por motivo de dispensa (Cód. 30)	29	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Indenizações trabalhistas, rescisões contratuais e por dispensa incentivada - inclusive os 50% do FGTS	30	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Benefícios concedidos (transporte, alimentação, auxílio-educação, planos de saúde, auxílio-doença, seguro de vida em grupo, treinamento, etc.)	31	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	PIS sobre folha de pagamento	31A	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00
	Total	32	<div style="border-bottom: 1px solid black; width: 100px;"></div> ,00

AUTENTICAÇÃO			
Após verificar se as informações foram prestadas em REAIS e SEM CENTAVOS , se estão corretas e sem rasuras			
Data de entrega: ____/____/____		Data de devolução: ____/____/____	
Responsável pelo preenchimento:		Cargo:	
Tel.:	Ramal:	Fax:	
Assinatura: <div style="text-align: center;">Certifico que as informações aqui contidas são completas e estão de acordo com as demonstrações financeiras da empresa.</div>			
Assinatura do Técnico de Pesquisas:			
OBSERVAÇÕES			
<div>.....</div> <div>.....</div> <div>.....</div>			
NOTA			
A empresa pode optar por prestar informações por questionário eletrônico, cujo software está disponível no site do IBGE: www.ibge.gov.br			
COMENTÁRIOS E/OU SUGESTÕES			
Que comentários sobre a pesquisa e sugestões de melhorias o(a) senhor(a) teria a fazer?			
<div>.....</div> <div>.....</div> <div>.....</div>			
Obrigado pela dedicação em preencher este questionário. Por favor, retenha uma cópia para seu controle. Os resultados finais desta pesquisa serão divulgados na <i>homepage</i> do IBGE.			
INSTRUÇÕES DE PREENCHIMENTO			
<p>Este questionário aplica-se às empresas com atividade predominantemente comercial, ou seja, cujo maior faturamento advém da atividade comercial, registradas no Cadastro Nacional das Pessoas Jurídicas - CNPJ - do Ministério da Fazenda.</p> <p>Como atividade comercial entende-se a compra de mercadorias para revenda, sem transformações, representação comercial e agentes do comércio.</p> <p>I - INFORMAÇÕES CADASTRAIS</p> <p>Preencha os Dados Cadastrais da Sede da Empresa e da Unidade de Coleta utilizando somente endereços pertencentes à empresa. Não preencha com o endereço do contador. Caso seja Unidade única, repetir o endereço da Sede.</p> <p>CAPÍTULO 03 - DADOS CADASTRAIS</p> <p>Item 02 - Nome fantasia - informar o nome pelo qual a empresa é conhecida. Caso a empresa possua lojas, filiais ou locais de venda com nomes fantasia diferenciados, registre o nome fantasia pela qual a empresa é mais conhecida. Caso a empresa não possua nome fantasia, registre S/D (Sem Denominação).</p> <p>Item 15 - Descrição da atividade de maior receita - informe com o maior nível de detalhamento possível a atividade desenvolvida pela empresa. Exemplo: comércio varejista de roupas, comércio atacadista de produtos alimentícios, comércio varejista de peças para automóveis, mercearia, farmácia, açougue, ótica, papelaria, livraria, sapataria, loja de material de construção, representação comercial de material de construção, representação de veículos automotores, representação de motocicletas, representação de combustíveis, etc.</p> <p>Item 16 - Código CNAE 2.0 - informe o código equivalente à descrição da atividade de maior faturamento da empresa. O código pode ser consultado na Relação de Atividades, em anexo, neste questionário ou pela internet no endereço http://www.cnae.ibge.gov.br.</p> <p>Item 17 - Sufixo e DV da Unidade de Coleta - corresponde ao endereço da Unidade designada para preencher o questionário. Se o endereço for de outra empresa do Grupo Empresarial, deverá ser preenchido com o Sufixo 9999 e o Dígito Verificador (DV) será, automaticamente, calculado pelo sistema.</p> <p>ATENÇÃO: não informar endereço do contador terceirizado/contratado.</p> <p>CAPÍTULO 04 - DADOS CADASTRAIS COMPLEMENTARES</p> <p>Item 01 - Situação cadastral em 31-12-15 - registre nas quadrículas o código da situação da empresa, conforme definido abaixo:</p> <p>01 - Em operação - se a empresa estava funcionando em 31.12.15.</p>			

03 - Paralisada com informação de receita - se a empresa estava com sua(s) atividade(s) paralisadas em 31.12.15, porém, tenha funcionado durante um período do ano e auferido receita.

04 - Extinta com informação de receita - se a empresa estava com sua(s) atividade(s) encerradas em 31.12.15, porém, tenha funcionado durante um período do ano e auferido receita.

Item 01A - Data da Ocorrência (somente para as Situações Cadastrais 03 e 04) - informar o mês e o ano da paralisação ou extinção da empresa.

Itens 02 e 03 - Mudanças na Estrutura da Empresa - caso tenha ocorrido algum processo de mudança na estrutura da empresa em 2015, devem ser preenchidos o código correspondente e o **CNPJ de Ligação**, respectivamente. Caso não tenha ocorrido no ano base de 2015, preencher com **00 - Não houve mudanças**.

Item 04 - Principais Mercadorias Revendidas - descreva, detalhadamente, as três principais mercadorias revendidas de acordo com a predominância na receita bruta do comércio.

Item 05 - Classe principal de comércio - preencha este item, de acordo com a predominância da receita, seguindo as definições das classes de comércio:

1 - Comércio Varejista - venda de produtos novos ou usados destinados, predominantemente, a pessoas físicas, para consumo, uso pessoal ou doméstico, independente da natureza e quantidade vendida.

2 - Comércio Atacadista - venda de produtos que serão utilizados no processo produtivo, destinam-se, portanto, a instituições públicas, exportações, revendedores, indústrias, profissionais autônomos, agricultores, prestadores de serviços (restaurantes, bares, oficinas, etc.).

3 - Representação comercial - vendas de produtos a base de comissão realizadas por representantes comerciais e/ou agentes do comércio, que se destinam a instituições públicas, exportações, revendedores, indústrias, profissionais autônomos, agricultores, prestadores de serviços (restaurantes, bares, oficinas, etc.).

Item 06 - Forma de tributação - registre na quadrícula o código correspondente a forma de tributação utilizada: **1 - Lucro Real**; **2 - Lucro Presumido ou Arbitrado**; **3 - Simples Nacional** - se a empresa estiver na condição de Microempresa - ME ou Empresa de Pequeno Porte - EPP, na forma da Lei Complementar nº 123/2006 que instituiu, a partir de 01-07-2007, novo tratamento tributário simplificado, também conhecido como Simples Nacional ou Super Simples. Tal regime substituiu, a partir de 01-07-2007, o Simples Federal, Lei 9.317/1996, que foi revogada a partir daquela data; ou **4 - Imune ou Isenta**.

II - INFORMAÇÕES DA EMPRESA

CAPÍTULO 05 - PESSOAL OCUPADO

Informar o número de pessoas efetivamente ocupadas em **31-03**, **30-06**, **30-09** e **31-12-15**, independente de terem ou não vínculo empregatício, desde que tenham sido remuneradas pela empresa, com exceção dos membros da família sem remuneração.

Inclua as pessoas afastadas em gozo de férias ou por motivo de licenças. Não considerar as pessoas que se encontram afastadas por licenças, ou pelo seguro por acidentes por mais de 15 dias. Não incluir os estagiários contratados pela empresa.

Pessoal assalariado - informar o número de pessoas que recebem salários independente de serem ou não ligadas à atividade comercial, tais como: vendedores, caixas, chefes, supervisores de vendas, gerente de vendas, carregadores, balconistas, pessoal ligado à atividade administrativa, estoquistas, empacotadores, frentistas, encarregados de compra, motoristas e ajudantes (caso a empresa possua frota própria para transporte de mercadorias), etc.

Pessoal não-assalariado - informar o proprietário e os sócios com atividade na empresa e membros da família sem carteira assinada, remuneração fixa ou salário.

CAPÍTULO 06 - DEMONSTRATIVO DA RECEITA NO ANO

17 e 18 - Receita bruta - informar as receitas brutas provenientes da revenda de mercadorias e de outras atividades executadas pela empresa, não deduzidas dos impostos incidentes sobre as vendas, das vendas canceladas, dos abatimentos e dos descontos incondicionais. Caso o valor do **Cód. 18** seja maior que o do **Cód. 17 + 17A**, favor discriminar em **OBSERVAÇÕES** qual atividade proporcionou esta receita.

17A - Comissões sobre vendas de representação comercial e agentes do comércio - informar a receita bruta proveniente do recebimento de comissões sobre as vendas efetuadas através de representação comercial e/ou agentes do comércio.

19 - SIMPLES NACIONAL - caso a empresa tenha optado pelo sistema de tributação SIMPLES NACIONAL, registre no **Cód. 19** o total do pagamento anual unificado dos impostos e contribuições relativos ao ano de 2015. O sistema SIMPLES NACIONAL implica no pagamento, em uma única guia de recolhimento, dos seguintes impostos e contribuições: Imposto de Renda Pessoa Jurídica - IRPJ, PIS, COFINS, IPI, Previdência Social, ICMS, ISS e Contribuição Social sobre o Lucro Líquido - CSLL.

21 - Demais receitas operacionais e outras receitas - inclua neste código as demais receitas operacionais, tais como: receitas financeiras,

variações monetárias ativas, receita de aluguéis, recuperação de despesas, resultado positivo em participações societárias, etc. Informar também as outras receitas, que são resultantes de operações não incluídas nas atividades principais e acessórias da empresa, tais como: o lucro na alienação de bens, ou seja, o resultado positivo obtido na venda de bens do ativo imobilizado.

CAPÍTULO 07 - COMPRAS, ESTOQUES E CUSTOS DAS MERCADORIAS REVENDIDAS

23 - Compras de mercadorias para revenda no ano - informar o valor das mercadorias adquiridas para revenda no mercado interno e externo, pagas à vista ou a prazo.

Inclua no custo das mesmas as despesas de seguro e de transporte até o estabelecimento, além dos impostos não recuperáveis. Informar o valor líquido das compras, ou seja, não inclua as devoluções, abatimentos, descontos obtidos, ICMS e IPI quando recuperáveis.

24 - Estoque em 31-12-14 - corresponde às mercadorias não vendidas no ano de 2014, inventariadas e contabilizadas, passando então para o exercício pesquisado como estoque inicial.

25 - Estoque em 31-12-15 - corresponde às mercadorias não vendidas no ano de 2015, inventariadas e contabilizadas no exercício pesquisado como estoque final.

25A - Custo das mercadorias revendidas - informar o valor contábil apurado e lançado na Demonstração do Resultado do Exercício. Observar o fechamento com a seguinte equação: **Códigos (23 + 24 - 25)**

CAPÍTULO 08 - GASTOS COM PESSOAL NO ANO

26 - Salários e outras remunerações - informar o total das importâncias pagas a título de salários fixos, comissões sobre vendas, horas extras, participações nos lucros pagas aos empregados, ajudas de custo, 13º salário, abono financeiro de 1/3 das férias, etc. Não deduza as parcelas correspondentes às cotas de Previdência e Assistência Social ou de consignação de interesse do empregado (aluguel de casa, contas de cooperativas, etc.). Não inclua: diárias pagas a empregados em viagens; participações ou comissões pagas a profissionais autônomos; indenizações pagas pelas empresas por motivos de dispensa (aviso prévio, 50% do FGTS, férias proporcionais, planos de demissão voluntária, etc.), que devem ser informados no **Cód. 30**; e bolsa paga a estagiários (**Cód. 39**).

27 - Retiradas pró-labore do proprietário ou sócios - considere as importâncias pagas a título de pró-labore dos sócios e do proprietário **com atividade na empresa**.

28 - Contribuições para a Previdência Social - informar as despesas da empresa e de competência do ano de referência da pesquisa, independente de terem sido pagas ou não. Não incluir a parte descontada dos salários dos empregados, nem os recolhimentos relativos aos autônomos.

29 - FGTS - informar os valores de competência do ano de referência da pesquisa, independente de terem sido pagos ou não. Caso tenha empregados o preenchimento é obrigatório, com exceção dos casos em que não existe vínculo empregatício. Lembramos que a opção pelo SIMPLES não desobriga a empresa de preencher este código. Não incluir os 50% do FGTS pagos por indenizações por dispensa, que deve ser informado no **Cód. 30**.

30 - Indenizações trabalhistas, rescisões contratuais e por dispensa incentivada - correspondem às obrigações da empresa por ocasião da dispensa do empregado, tais como: aviso prévio, 50% do FGTS, 13º salário e férias proporcionais, planos de demissão voluntária (incentivo a demissões), acordos judiciais, etc.

31A - PIS sobre folha de pagamento - informar as contribuições incidentes sobre o total da folha de pagamento de seus empregados, quando a empresa for contribuinte nesta modalidade.

CAPÍTULO 09 - DESPESAS NO ANO

Não devem ser registradas, neste capítulo, as despesas e os custos registrados em capítulos anteriores, tais como: custo das mercadorias revendidas, gastos com pessoal (encargos sociais, salários, retiradas e outras remunerações), etc.

34 - Serviços prestados por terceiros - informar as despesas pagas a profissionais liberais ou autônomos, tais como: contadores, advogados, despachantes, bombeiros hidráulicos, eletricitas, pedreiros, pintores, etc; e também a empresas prestadoras de serviços, tais como: contabilidade, informática, segurança, transporte de valores, manutenção, etc.

38 - Despesas financeiras - informar as despesas relativas a juros de empréstimos, financiamentos, descontos de títulos e descontos concedidos a clientes por pagamento antecipado de duplicatas e outros títulos.

38B - PIS/COFINS sobre outras receitas - informar as contribuições incidentes sobre as receitas declaradas no **Capítulo 06 - Demais receitas (Cód. 21)**.

39 - Demais despesas operacionais e outras despesas - devem ser informadas as despesas operacionais não registradas nos Códigos 33 a 38D, tais como: contribuições a associações de classe, assinaturas de revistas e publicações, indenizações trabalhistas sem motivo de demissão, despesas com cartórios, despesas indedutíveis, doações, donativos, títulos incobráveis, quebra de estoque, bonificações e garantias, pedágio, etc. Informar também as outras despesas, ou seja, as despesas resultantes de transações não incluídas nas atividades principais ou acessórias da empresa, tais como: o prejuízo na alienação de bens, ou seja, o resultado negativo obtido na venda de bens do ativo imobilizado, que consiste

na diferença entre o valor de venda do bem e o valor contábil (custo histórico, corrigido e depreciado). Quando esta despesa ultrapassar 30% do **Cód. 40**, discrimine os três principais valores nos campos reservados para esse fim.

Inclua despesas com redução nos valores dos ativos - *impairment*, de acordo com a legislação contábil instituída pelas leis 11.638/07 e 11.941/09.

Devem ser informados em outros códigos: Provisão para Imposto de Renda (IRPJ) e Contribuição Social sobre o Lucro Líquido (CSLL) (**Cód. 38C**); Outras Provisões constituídas (contingências, provisão para devedores duvidosos, perdas com clientes, perdas de estoque, etc.)(**Cód. 38D**); REFIS (apenas o valor dos juros de atualização será considerado no questionário e informado como despesas financeiras) e juros sobre Capital Próprio (empresas que não sejam S/A) (**Cód. 38**).

Não informar em nenhum capítulo do questionário: Juros sobre Capital Próprio de companhias abertas (a deliberação CVM nº 207/1996 determina às companhias abertas (S/A) a contabilização desses juros diretamente em conta de lucros acumulados sem afetar o resultado do exercício); valores de depreciação e amortização acumuladas; e recolhimentos da Previdência Social relativos aos autônomos.

CAPÍTULO 09A - AQUISIÇÕES E BAIXAS DO ATIVO IMOBILIZADO NO ANO

40A - Aquisições e produção própria - considere como aquisições todos os bens adquiridos e incorporados ao Ativo Imobilizado da empresa, pelo custo histórico, durante o ano. Não inclua os encargos financeiros decorrentes de financiamento. Registre também os bens oriundos da produção própria, e incorporados ao Ativo Imobilizado no ano.

40B - Melhorias - considere como melhorias as benfeitorias e os melhoramentos que tenham aumentado a vida útil dos bens e que tenham sido incorporados ao Ativo Imobilizado no ano.

40C - Baixas - registre todos os bens do Ativo Imobilizado baixados pelo valor contábil, isto é, pelo custo histórico corrigido e depreciado, durante o ano.

CAPÍTULO 10 - FORMA DE REALIZAÇÃO DAS VENDAS - SISTEMA DE COMERCIALIZAÇÃO

Deverá ser informado o percentual, segundo a forma de comercialização, da receita bruta de revenda no ano, informada no **Capítulo 06, Cód. 17**.

41 - Lojas (vendas no balcão), estabelecimentos, postos de combustíveis, boxes em mercados, depósitos, galpões, armazéns e salas - vendas realizadas em estabelecimentos localizados em prédios circundados por paredes e com entrada independente.

42 - Quiosques, trailers e outros estabelecimentos situados em local fixo, fora da loja, em estradas, praças, rodoviárias, corredores de shopping centers, etc. - vendas realizadas em estabelecimentos localizados em espaços públicos, como praças, calçadas, ruas de pedestre ou corredores de shopping centers.

43 - Correio - vendas realizadas através de correio com apoio ou não de catálogos que contêm a descrição dos produtos e seus preços.

44 - Porta a porta (domicílio), postos móveis ou por ambulantes - vendas realizadas por vendedores que se deslocam às casas dos consumidores potenciais (com apoio ou não de catálogos que contêm a descrição dos produtos e seus preços), se deslocam pela cidade ou se fixam nas ruas.

45 - Internet - vendas realizadas através de um *site* da rede internacional de computadores.

46 - Televendas - vendas realizadas por telefone onde a empresa toma a iniciativa de ligar para o cliente em potencial ou disponibiliza um número para contato.

47 - Outros - forma de comercialização não incluída nos códigos anteriores, favor discriminar no campo específico.

III - DADOS DE REGIONALIZAÇÃO

Devem ser informados pela empresa, ainda que esta tenha exercido suas atividades em apenas uma Unidade da Federação. Os **Códigos 59, 65 e 77** devem corresponder aos **Códigos 16, (26 + 27) e (17 + 17A)**, respectivamente. Nos endereços da empresa onde foram exercidas, exclusivamente, atividades administrativas e de apoio (depósitos, almoxarifados, escritórios contábeis, etc.), informar apenas pessoal ocupado e salários.

Como Unidades Locais com receita de revenda, considerar todos os endereços onde funcionam lojas, filiais, locais de venda, postos de venda, etc., ou seja, todos os endereços da empresa que efetuam revenda de mercadorias.

A Receita Bruta de Revenda deve ser informada na Unidade da Federação onde ocorrer o fato gerador da receita, ou seja, na Unidade da Federação onde foi realizada a revenda das mercadorias.

RELAÇÃO DE ATIVIDADES NO ÂMBITO DA PAC	
CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0	
CÓD	DESCRIÇÃO
	COMÉRCIO A VAREJO, POR ATACADO E REPRESENTAÇÃO COMERCIAL DE VEÍCULOS AUTOMOTORES, MOTOCICLETAS, PEÇAS E ACESSÓRIOS
4511-1	Comércio a varejo e por atacado de veículos automotores novos e usados - inclusive reboque e semi-reboques, ambulâncias, veículos betoneiras - não incluir bicicletas (4649-4 e 4763-6), tratores agrícolas (4661-3) e motocicletas (4541-2)
4512-9	Representações comerciais e agentes do comércio de veículos automotores - inclusive a venda sob consignação a varejo e por atacado - não incluir a representação comercial e agentes do comércio de peças e acessórios novos e usados para veículos automotores (4530-7)
4530-7	Representações comerciais, agentes do comércio e comércio a varejo e por atacado de peças e acessórios, novos e usados, para veículos automotores - não incluir para motocicletas (4541-2) e para tratores agrícolas (4661-3)
4541-2	Comércio por atacado e a varejo de motocicletas, peças e acessórios novos e usados
4542-1	Representações comerciais e agentes do comércio de motocicletas, peças e acessórios - inclusive a venda sob consignação, a varejo e por atacado.
	REPRESENTANTES COMERCIAIS E AGENTES DO COMÉRCIO EXCETO DE VEÍCULOS AUTOMOTORES
4611-7	Matérias-primas agrícolas <i>in natura</i> e animais vivos (cana-de-açúcar, algodão, café em coco, bovinos, suínos, lã, couro bruto, chifres, ossos, crinas, etc.)
4612-5	Combustíveis, minerais, produtos siderúrgicos e químicos
4613-3	Madeira, material de construção, ferragens, material elétrico e material para pintura
4614-1	Máquinas, equipamentos, embarcações e aeronaves
4615-0	Eletrodomésticos, móveis e artigos de uso doméstico - não incluir máquinas e equipamentos de informática e de comunicação (4614-1)
4616-8	Têxteis, vestuário, calçados e artigos de viagens (fios naturais ou artificiais, fibras têxteis, artigos e complementos do vestuário e artigos em couro curtido).
4617-6	Produtos alimentícios, bebidas e fumo - não incluir ração para animais (4618-4); fumo não beneficiado (4611-7)
4618-4	Produtos especializados não especificados anteriormente (produtos farmacêuticos, cosméticos, perfumaria, jóias, brinquedos, bicicletas, papel de parede, ração, distribuidores de jornais e revistas para o comércio varejista sob contrato de comissão) - não incluir entregas de jornais e revistas em domicílio (5320-2); veículos automotores e auto-peças (45)
4619-2	Representantes comerciais e agentes do comércio de mercadorias em geral não especializados (representante sem predominância de algum tipo de mercadoria ou grupos de mercadorias específicas)
	COMÉRCIO ATACADISTA DE MATÉRIAS-PRIMAS AGRÍCOLAS, ANIMAIS VIVOS; PRODUTOS ALIMENTÍCIOS PARA ANIMAIS
4621-4	Comércio atacadista de café em grão, em coco ou verde
4622-2	Comércio atacadista de soja
4623-1	Comércio atacadista de animais vivos (bovinos, ovinos, suínos etc.), forragens, rações, couros, lã, peles, sêmen bovino e outros subprodutos de origem animal, algodão, fumo em folha não beneficiado, cacau, sementes, flores, plantas, sisal, cereais em bruto, fibras têxteis não beneficiadas - não incluir produtos hortifrutigranjeiros (4633-8) e cereais beneficiados (4632-0), aves vivas, coelhos e pequenos animais vivos para alimentação (4633-8)
	COMÉRCIO ATACADISTA ESPECIALIZADO EM PRODUTOS ALIMENTÍCIOS, BEBIDAS E FUMO
4631-1	Comércio atacadista de leite e laticínios
4632-0	Comércio atacadista de cereais e leguminosas beneficiados, farinhas, amidos e féculas
4633-8	Comércio atacadista de hortifrutigranjeiros - inclusive aves vivas, ovos e pequenos animais vivos para alimentação
4634-6	Comércio atacadista de carnes, produtos da carne, pescados e frutos do mar - inclusive produtos de salsicharia, aves abatidas e derivados
4635-4	Comércio atacadista de bebidas alcoólicas e não-alcoólicas
4636-2	Comércio atacadista de produtos do fumo (cigarros, cigarrilhas, fumo em rolo ou em corda beneficiado, inclusive isqueiros, cachimbos e demais artigos de tabacaria) - não incluir fumo em folha não beneficiado (4623-1)
4637-1	Comércio atacadista especializado em produtos alimentícios não especificados anteriormente (café torrado, moído e solúvel, açúcares, balas, chocolates, mel, sorvetes, produtos de padaria, massas, sucos, conservas de frutas e legumes, óleos e gorduras, condimentos, alimentos congelados, etc.)

CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0

CÓD	DESCRIÇÃO
4639-7	Comércio atacadista de produtos alimentícios em geral compreende a comercialização dos produtos das classes 4631-1 a 4637-1, porém sem predominância de nenhum deles
COMÉRCIO ATACADISTA DE PRODUTOS DE CONSUMO NÃO-ALIMENTAR	
4641-9	Comércio atacadista de tecidos, artefatos de tecidos, artigos de cama, mesa e banho e artigos de armarinho - não incluir colchoaria, persianas e cortinas (4649-4)
4642-7	Comércio atacadista de artigos do vestuário e acessórios - inclusive para uso profissional e de segurança do trabalho, roupas esportivas fardamento e uniforme
4643-5	Comércio atacadista de calçados e artigos de viagem - calçados de qualquer material, inclusive esportivo, bolsas, malas e outros artigos para viagens - não incluir calçados ortopédicos sob medida (3250-7)
4644-3	Comércio atacadista de produtos farmacêuticos, para uso humano e veterinário inclusive da flora medicinal - não incluir cosméticos, produtos de perfumaria e de higiene pessoal (4646-0)
4645-1	Comércio atacadista de instrumentos e materiais para uso médico, cirúrgico, hospitalares e de laboratórios, próteses e artigos de ortopedia e produtos odontológicos - não incluir equipamentos médico-cirúrgicos hospitalares (4664-8)
4646-0	Comércio atacadista de cosméticos, produtos de perfumaria e de higiene pessoal, inclusive fraldas descartáveis
4647-8	Comércio atacadista de artigos de escritório e de papelaria; livros, jornais e outras publicações; papel, papelão e seus artefatos - não incluir papel e papelão em bruto, embalagens de papel e papelão (4686-9), papel e papelão usados e aparas de papel (4687-7), suprimentos para informática (4651-6)
4649-4	Comércio atacadista de equipamentos elétricos, eletrônicos pessoal e de uso domésticos, móveis e outros artigos de usos pessoal e doméstico não especificados anteriormente (bicicletas, triciclos, artigos de tapeçaria, persianas, cortinas, lustres, luminárias, filmes para fotos, CDs, fitas e discos, produtos de higiene e limpeza domiciliar, jóias, relógios, bijuterias, pedras preciosas e semipreciosas lapidadas, artigos de cutelaria, artigos para habitação de vidro, cristal, porcelana, borracha, plástico, metal, madeira, vime, bambu e outros similares; brinquedos; instrumentos musicais; artigos de óptica, papel de parede, artigos para caça, pesca e <i>camping</i> , etc.) - não incluir equipamento e material de informática (4651-6), equipamentos de comunicação (4652-4), equipamentos de ginástica (4669-9) e de fisioterapia (4664-8), peças e acessórios para aparelhos eletrodomésticos (4689-3)
COMÉRCIO ATACADISTA DE EQUIPAMENTOS E PRODUTOS DE TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO	
4651-6	Comércio atacadista de computadores, periféricos e suprimentos de informática (computadores, programas de computador não-customizáveis, <i>toner</i> , fitas, disquetes ópticos, etc) - não incluir componentes eletrônicos para equipamentos de informática (4652-4), máquinas para automação gerencial e comercial (4665-6), manutenção e reparação de computadores (9511-8)
4652-4	Comércio atacadista de componentes eletrônicos e equipamentos de telefonia e comunicação (válvulas e tubos eletrônicos, semicondutores, <i>microchips</i> e circuitos integrados, circuitos impressos, telefones, intercomunicadores, secretária eletrônica e similares)
COMÉRCIO ATACADISTA DE MÁQUINAS, APARELHOS E EQUIPAMENTOS, EXCLUSIVE DE TECNOLOGIA DE INFORMAÇÃO E COMUNICAÇÃO	
4661-3	Comércio atacadista de máquinas, aparelhos e equipamentos para uso agropecuário; partes e peças; arados, cultivadores, tratores para uso na agricultura e silvicultura, cortadores de grama, inclusive máquinas e equipamentos para avicultura e apicultura
4662-1	Comércio atacadista de máquinas, equipamentos para terraplenagem, mineração e construção - não incluir partes e peças de tratores para uso na agricultura e silvicultura (4661-3)
4663-0	Comércio atacadista de máquinas, aparelhos e equipamentos para uso industrial; partes e peças - máquinas operatrizes e máquinas ferramentas para uso nos diversos ramos de indústria - não incluir para mineração e construção (4662-1)
4664-8	Comércio atacadista de máquinas, aparelhos e equipamentos para uso odonto-médico-hospitalares e laboratoriais; partes e peças - mobiliário médico-hospitalares e odontológico, equipamentos para laboratórios, para monitoração médica, médico-cirúrgicos e para fisioterapia
4665-6	Comércio atacadista de máquinas e equipamentos para uso comercial; partes e peças - balanças, balcões frigoríficos, refrigeradores, vitrines, registradoras, máquinas de fatiar, emissoras de cupom fiscal, leitora óptica, máquinas para automação comercial
4669-9	Comércio atacadista de máquinas, aparelhos e equipamentos não especificados anteriormente; partes e peças - bombas, compressores e carneiros hidráulicos, motores e transformadores elétricos, sistema para controle de incêndio, instrumentos e equipamentos de medida, máquinas de costura para qualquer uso, máquinas e equipamentos para usos técnicos e profissional, máquinas e equipamentos para escritório, robôs.

CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0

CÓD	DESCRIÇÃO
COMÉRCIO ATACADISTA DE MADEIRA, FERRAGENS, FERRAMENTAS, MATERIAL ELÉTRICO E MATERIAL DE CONSTRUÇÃO	
4671-1	Comércio atacadista de madeira e produtos derivados - madeira em bruto perfilada ou serrada, tábuas, pranchas, dormentes, barrotes e similares, pré-moldados de madeira para construção - não incluir fabricação de pré-moldados com montagem (1622-6)
4672-9	Comércio atacadista de ferragens e ferramentas (ferragens para construção, ferramentas manuais) - não incluir produtos siderúrgicos e metalúrgicos para construção (vergalhões, arames e pregos)(4679-6), chapas de aço, bobinas (4685-1) e máquinas ferramentas industriais (4669-9)
4673-7	Comércio atacadista de material elétrico - fios, cabos, condutores elétricos, lâmpadas, tomadas, chaves elétricas - não incluir lustres, luminárias (4649-4) e peças e acessórios para aparelhos de usos doméstico e pessoal, elétricos e eletrônicos (4689-3)
4674-5	Comércio atacadista de cimento
4679-6	Comércio atacadista especializado de materiais de construção não especificados anteriormente e de materiais de construção em geral - tintas e materiais para pintura, mármore e granitos, vidros, espelhos e molduras, areia, pedra, tijolos, tubos e canos de água, aparelhos sanitários, vergalhões, arames, pregos, esquadrias metálicas, divisórias, azulejos, pisos, piscinas e equipamentos correlatos, etc.
COMÉRCIO ATACADISTA ESPECIALIZADO EM OUTROS PRODUTOS	
4681-8	Comércio atacadista de combustíveis sólidos, líquidos e gasosos - inclusive lubrificantes, lenha, carvão vegetal, carvão mineral, petróleo cru em bruto - não incluir gás natural veicular para postos de combustíveis (3520-4) e GLP (4682-6) e a venda a varejo de GNV (4731-8)
4682-6	Comércio atacadista de Gás Liquefeito de Petróleo (GLP)
4683-4	Comércio atacadista de defensivos agrícolas, adubos, fertilizantes e corretivos do solo - não incluir produtos farmacêuticos para uso veterinário (4644-3)
4684-2	Comércio atacadista de produtos químicos e petroquímicos (resinas e elastômeros, solventes, cloro, oxigênio, carga para extintor de incêndio, álcool etílico, artigos pirotécnicos, fósforo de segurança, adesivos, essências e concentrados aromáticos) - não incluir agroquímicos (4683-4)
4685-1	Comércio atacadista de produtos siderúrgicos e metalúrgicos (chapas de aço, bobinas etc) - não incluir artigos para construção civil (4679-6)
4686-9	Comércio atacadista de papel e papelão em bruto e de embalagem
4687-7	Comércio atacadista de resíduos e sucatas metálicas e não metálicas - aparas de papel e papelão, sacos usados, plásticos, vidros usados
4689-3	Comércio atacadista especializado de outros produtos intermediários não especificados anteriormente - fios e fibras têxteis beneficiadas, produtos de extração mineral (minerais ferrosos e não ferrosos, sal gema e sal marinho, ouro e outros metais preciosos), inclusive peças e acessórios para aparelhos de uso doméstico e pessoal - não incluir combustíveis (4681-8) artefatos de borracha para veículos (4530-7) e para uso residencial (4649-4)
COMÉRCIO ATACADISTA NÃO-ESPECIALIZADO	
4691-5	Comércio atacadista de mercadorias em geral, com predominância de produtos alimentícios (não especializados, tais como, supermercado/hipermercado atacadista, distribuidores de cesta básica)
4692-3	Comércio atacadista de mercadorias em geral, com predominância de insumos agropecuários - não incluir o comércio especializado na venda de insumos agropecuários (ração no 4623-1; produtos veterinários no 4644-3; implementos agrícolas no 4661-3; adubos, fertilizantes para agricultura no 4683-4)
4693-1	Comércio atacadista de mercadorias em geral, sem predominância de produtos alimentícios ou de insumos agropecuários
COMÉRCIO VAREJISTA NÃO-ESPECIALIZADO	
4711-3	Comércio varejista de mercadorias em geral, com predominância de produtos alimentícios - hipermercado (área de venda superior a 5.000 m² e supermercado (área de venda entre 300 e 5.000 m²)
4712-1	Comércio varejista de mercadorias em geral, com predominância de produtos alimentícios (minimercados, mercearias, armazéns, empório, estivas, quitanda, com área de venda inferior a 300 m²)
4713-0	Comércio varejista de mercadorias em geral, sem predominância de produtos alimentícios - lojas de departamentos ou magazines, bazar, lojas de variedades, lojas duty-free de aeroportos internacionais, "lojas de 1,99", etc.
COMÉRCIO VAREJISTA DE PRODUTOS ALIMENTÍCIOS, BEBIDAS E FUMO	
4721-1	Comércio varejista de produtos de padaria, laticínios, frios e carnes conservadas, doces, balas, bombons e semelhantes, conservas de frutas, legumes, verduras e similares

CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0

CÓD	DESCRIÇÃO
4722-9	Comércio varejista de carnes e pescados (açougues e peixarias - inclusive pequenos animais abatidos, miúdos, vísceras e similares, pescado crustáceos e moluscos, frescos, congelados, conservados ou frigorificados)
4723-7	Comércio varejista de bebidas alcoólicas e não-alcoólicas - não incluir bares, botequins e lanchonetes (5611-2)
4724-5	Comércio varejista de hortifrutigranjeiros, inclusive aves vivas, coelhos e outros pequenos animais para alimentação
4729-6	Comércio varejista de produtos alimentícios em geral ou especializados em produtos alimentícios não especificados anteriormente - artigos de tabacaria, comidas congeladas, produtos naturais e dietéticos, gelo, mel, sorvetes embalados em potes e similares, lojas de conveniência, lojas de delicatessen, etc.
COMÉRCIO VAREJISTA DE COMBUSTÍVEIS PARA VEÍCULOS AUTOMOTORES	
4731-8	Comércio varejista de combustíveis para veículos automotores - inclusive Gás Natural Veicular (GNV) - não incluir Gás Liquefeito de Petróleo (GLP) (4784-9) e lubrificantes (4732-6)
4732-6	Comércio varejista de lubrificantes para uso automotivo e para outros usos
COMÉRCIO VAREJISTA DE MATERIAL DE CONSTRUÇÃO	
4741-5	Comércio varejista de tintas e materiais para pintura - inclusive tintas automotivas
4742-3	Comércio varejista de material elétrico (fios cabos, condutores elétricos, chaves elétricas, interruptores, tomadas e similares) - não incluir peças e acessórios para aparelhos de uso domésticos e pessoal (4757-1) para veículos (45), lustres, luminárias e abajures (4754-7)
4743-1	Comércio varejista de vidros planos e de segurança e espelhos - não incluir para veículos (4530-7) e especializado em molduras de quadros (4789-0)
4744-0	Comércio varejista de ferragens, madeira e seus artefatos, material hidráulico e materiais de construção não especificados anteriormente e materiais de construção em geral sem especialização
COMÉRCIO VAREJISTA DE EQUIPAMENTOS DE INFORMÁTICA E COMUNICAÇÃO; EQUIPAMENTOS E ARTIGOS DE USO DOMÉSTICO	
4751-2	Comércio varejista especializado de equipamentos e suprimentos de informática (computadores e periféricos, programas de computador não-customizáveis, partes e peças para equipamentos de informática, recarga de cartuchos associado ao comércio) - não incluir equipamentos para escritório (4789-0)
4752-1	Comércio varejista de equipamentos de telefonia e comunicação - telefones, fax, intercomunicadores e similares, peças e partes para equipamentos de telefonia
4753-9	Comércio varejista especializado de eletrodomésticos e equipamentos de áudio e vídeo - não incluir equipamentos de informática (4751-2), de comunicação (4752-1) e peças e acessórios para eletroeletrônicos domésticos (4757-1)
4754-7	Comércio varejista de móveis para qualquer uso, colchoaria e artigos de iluminação residencial - não incluir artigos para habitação (4759-8)
4755-5	Comércio varejista especializado em tecidos, artigos de cama, mesa e banho e armarinho - não incluir cortinas, tapeçaria (4759-8)
4756-3	Comércio varejista especializado de instrumentos musicais e acessórios, inclusive músicas impressas
4757-1	Comércio varejista especializado de peças e acessórios para aparelhos eletroeletrônicos para uso doméstico - inclusive antenas - não incluir partes e peças para equipamentos de informática (4751-2) e comunicação (4752-1)
4759-8	Comércio varejista de artigos de uso doméstico não especificados anteriormente (artigos de tapeçaria, cortinas e persianas, artigos para habitação, de vidro, cristal, porcelana, borracha, plástico, metal, madeira, vime, bambu, e outros similares, artigos de cutelaria, toldos e similares, papel de parede, sistema de segurança residencial - não incluir molduras e quadros e flores para ornamentação, objetos de arte (4789-0) antiguidades (4785-7)
COMÉRCIO VAREJISTA DE ARTIGOS CULTURAIS, RECREATIVOS E ESPORTIVOS	
4761-0	Comércio varejista de livros, jornais, revistas e papelaria - não incluir suprimentos de informática (4751-2) e livros usados (4785-7)
4762-8	Comércio varejista de discos, fitas de áudio, vídeo, CDs e DVDs gravados ou não
4763-6	Comércio varejista de artigos recreativos e esportivos (brinquedos e jogos eletrônicos, bicicletas e triciclos, artigos de caça, pesca e <i>camping</i> , artigos do vestuário e acessórios especializado para a prática de esportes, embarcações e outros veículos recreativos, motores para embarcações de esporte e lazer; peças e acessórios referente a esta classe - não incluir armas e munições (4789-0) e calçados esportivos (4782-2)

RELAÇÃO DE ATIVIDADES NO ÂMBITO DA PAC		(conclusão)
CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0		
CÓD	DESCRIÇÃO	
	COMÉRCIO VAREJISTA DE PRODUTOS FARMACÊUTICOS, PERFUMARIA E COSMÉTICOS, ÓPTICOS E ORTOPÉDICOS	
4771-7	Comércio varejista de produtos farmacêuticos para uso humano e veterinário, inclusive homeopáticos, fitoterápicos e da flora medicinal - não incluir produtos odontológicos (4645-1), cosméticos, perfumaria e de higiene pessoal (4772-5)	
4772-5	Comércio varejista de cosméticos, produtos de perfumaria, de toucador e de higiene pessoal	
4773-3	Comércio varejista de artigos médicos e ortopédicos (muletas, cadeiras de rodas, aparelhos auditivos, kits diagnósticos etc) - não incluir equipamentos médico-cirúrgicos hospitalares (4664-8)	
4774-1	Comércio varejista de artigos de óptica - óculos, lentes	
	COMÉRCIO VAREJISTA DE PRODUTOS NOVOS NÃO ESPECIFICADOS ANTERIORMENTE E DE PRODUTOS USADOS	
4781-4	Comércio varejista de artigos do vestuário novos de qualquer material, acessórios e complementos do vestuário - não incluir roupas para segurança pessoal (4642-7); roupas e artigos do vestuário usados (4785-7), calçados e artigos de viagem (4782-2), artigos esportivos (4763-6)	
4782-2	Comércio varejista de calçados e artigos de viagem	
4783-1	Comércio varejista de jóias e relógios inclusive pedras preciosas e semi-preciosas lapidadas, inclusive peças e partes - não incluir bijuteria (4789-0)	
4784-9	Comércio varejista de Gás Liquefeito de Petróleo - (GLP) em botijões	
4785-7	Comércio varejista de artigos usados - móveis, eletrodomésticos, objetos de arte, roupas e calçados, moedas e selos de coleção, livros e revistas, material de demolição, sacos, tambores - não incluir veículos usados, partes e peças (45)	
	COMÉRCIO VAREJISTA DE OUTROS PRODUTOS NOVOS NÃO ESPECIFICADOS ANTERIORMENTE	
4789-0	Comércio varejista de outros produtos novos não especificados anteriormente - suvenires, bijuterias, flores, plantas, vasos, material de limpeza, sementes e mudas para jardinagem, artesanato, objetos de arte; animais vivos e alimentos para animais de estimação; brindes; produtos eróticos; artigos para festas; artigos funerários; fogos de artifício, equipamentos para escritório, artigos fotográficos e para filmagem, armas e munições, artigos religiosos e de culto, rede de dormir, carvão e lenha, extintores, cartões telefônicos, cargas e preparados para incêndio, quinquilharias para usos agrícolas, etc.)	

Anexo B

Questionário Completo da PAC

Instituto Brasileiro de Geografia e Estatística
Diretoria de Pesquisas
Coordenação de Serviços e Comércio
Gerência de Pesquisas

PESQUISA ANUAL DE COMÉRCIO - 2015

QUESTIONÁRIO COMPLETO

www.ibge.gov.br
ibge@ibge.gov.br

01	IDENTIFICAÇÃO DO QUESTIONÁRIO (Uso do Órgão Regional)		
01	CÓDIGO DO MUNICÍPIO DA UC		
	UF	MUNICÍPIO	DIST/SUBDIST
02	CADASTRO DO TÉCNICO DE PESQUISAS	03	CONTROLE
			2

OBRIGATORIEDADE E SIGILO DAS INFORMAÇÕES - a legislação vigente, de acordo com o Decreto Federal nº 73.177 de 20 de novembro de 1973 e a Lei nº 5.534 de 14 de novembro de 1968, modificada pela Lei nº 5.878 de 11 de maio de 1973, dispõe sobre a obrigatoriedade e sigilo das informações coletadas pelo IBGE, as quais se destinam, exclusivamente, a fins estatísticos e não poderão ser objeto de certidão e nem terão eficácia jurídica como meio de prova.

PRAZO DE ENTREGA - conforme descrito no recibo entregue pelo Técnico de Pesquisas.

PROPÓSITO DA PESQUISA ANUAL DE COMÉRCIO - coletar dados econômico-financeiros necessários à formulação de políticas públicas e programas sociais, bem como à estimação pelo Sistema de Contas Nacionais dos agregados macroeconômicos, em especial o Produto Interno Bruto - PIB.

I - INFORMAÇÕES CADASTRAIS

02	IDENTIFICAÇÃO DA SEDE DA EMPRESA									
PAC		CNPJ da empresa	Sufixo	DV	PAC					
03	DADOS CADASTRAIS									
SEDE DA EMPRESA (MATRIZ)										
01	FIRMA OU RAZÃO SOCIAL									
02	NOME FANTASIA (se não possuir, registre S/D)				02A	SITE DA EMPRESA (se não possuir, registre N/T)				
03	LOGRADOURO (rua, avenida, rodovia, etc.)		04	NÚMERO	05	COMPLEMENTO (bloco, grupo, andar, sala, km, etc.)				
06	BAIRRO/DISTRITO/LOCALIDADE		13	NOME DO MUNICÍPIO		12	SIGLA UF	07	CEP	
08	DDD	09	TELEFONE	10	RAMAL	11	FAX	14	E-MAIL DO RESPONSÁVEL PELA EMPRESA	
15	DESCRIÇÃO DA PRINCIPAL ATIVIDADE (MAIOR RECEITA) DA EMPRESA								16	CÓDIGO CNAE 2.0
UNIDADE DE COLETA										
Endereço da Unidade da Empresa (Sede/Filial/Grupo Empresarial) designada para preencher o questionário. Não informar endereço do contador terceirizado/contratado										
17	SUFIXO / DV - DA UNIDADE DE COLETA	18	LOGRADOURO (rua, avenida, rodovia, etc.)	19	NÚMERO	20	COMPLEMENTO (bloco, grupo, andar, sala, km, etc.)			
21	BAIRRO/DISTRITO/LOCALIDADE		28	NOME DO MUNICÍPIO		27	SIGLA UF	22	CEP	
23	DDD	24	TELEFONE	25	RAMAL	26	FAX	29	E-MAIL DO RESPONSÁVEL PELO PREENCHIMENTO	
04	DADOS CADASTRAIS COMPLEMENTARES									
01	SITUAÇÃO CADASTRAL EM 31-12-2015		02	MUDANÇAS NA ESTRUTURA DA EMPRESA NO ANO DE 2015			03	CNPJ DE LIGAÇÃO DA EMPRESA		
01 - Em operação 03 - Paralisada com informação de receita 04 - Extinta com informação de receita			00 - Não houve mudanças 01 - Fusão ou cisão total 02 - Cisão parcial 03 - Incorporação de outra(s) empresa(s) 03A - Incorporada por outra empresa 06 - Alteração de CNPJ por outros motivos (esclareça em "OBSERVAÇÕES")				Sucessora/Antecessora - Arrendatária/Arrendada			
01A DATA DA OCORRÊNCIA para os Itens 03 e 04							Sucessora/Antecessora - Arrendatária/Arrendada			
MÊS			ANO							
01			02							
04	PRINCIPAIS MERCADORIAS REVENDIDAS POR ORDEM DE PREDOMINÂNCIA NA RECEITA, INCLUSIVE POR REPRESENTAÇÃO COMERCIAL						05	FORMA DE TRIBUTAÇÃO UTILIZADA PELA EMPRESA		
1)						1- Lucro Real				
2)						2 - Lucro Presumido ou Arbitrado				
3)						3 - Sistema "Simples Nacional"				
						4 - Imune ou Isenta				

II - INFORMAÇÕES DA EMPRESA

ATENÇÃO: Unidade de Investigação: **neste questionário devem ser registrados os dados da empresa como um todo.**

Informações contábeis: devem se referir às de competência do ano civil (janeiro a dezembro) e serem prestadas de acordo com a Legislação Societária.

Forma de preenchimento: registre os dados, de forma legível, a caneta. O preenchimento dos valores deve ser em **REAIS e SEM CENTAVOS.**

Antes de preencher o questionário, leia as instruções em anexo.

05	NÚMERO DE PESSOAS OCUPADAS	CÓD	EM 31-03	CÓD	EM 30-06	CÓD	EM 30-09	CÓD	EM 31-12
Pessoal Assalariado									
Ligado à atividade comercial ou de representação (gerentes de vendas, supervisores, vendedores, balconistas, caixas, representante comercial, empacotadores, frentistas, etc.)		001		006		011		016	
Não ligado à atividade comercial (presidente, diretores, pessoal administrativo e ligado a outras atividades, etc.)		002		007		012		017	
Pessoal Não-Assalariado									
Proprietário e sócios com atividade na empresa		003		008		013		018	
Membros da família sem remuneração.....		004		009		014		019	
Total.....		005		010		015		020	
06	DEMONSTRATIVO DA RECEITA NO ANO					CÓD	VALORES EM REAIS		
Receita Bruta (Se cód. 022 a 025 > 021 + 021A, discrimine a atividade correspondente em OBSERVAÇÕES)									
Revenda de mercadorias (no caso de cooperativa, inclua as vendas a cooperados e de mercadorias de cooperados)						021			,00
Comissões sobre vendas de representação comercial e agentes de comércio						021A			,00
Venda de produtos de fabricação própria						022			,00
Serviços de manutenção e reparação (veículos, eletrodomésticos, computadores, telefones, relógios, etc.) ...						023			,00
Royalties de franquias						024			,00
Outras atividades (lanchonetes, restaurantes, armazenagem, transporte, exploração de estacionamento, instalação e manutenção elétrica ou hidráulica, agropecuária, etc.)						025			,00
Deduções (-)									
Vendas Canceladas, abatimentos e descontos incondicionais (-)						026			,00
ICMS sobre vendas (-)						027			,00
PIS/PASEP (-)						028			,00
Outros impostos e contribuições sobre vendas (IPI, ISS, SIMPLES NACIONAL, COFINS) (-)						029			,00
Receita Líquida (021 + 021A + 022 + 023 + 024 + 025 - 026 - 027 - 028 - 029)						030			,00
07	DEMAIS RECEITAS NO ANO					CÓD	VALORES EM REAIS		
Aluguel de imóveis, máquinas e equipamentos						031			,00
Receitas financeiras (juros, descontos obtidos, etc.)						032			,00
Variações monetárias ativas						033			,00
Resultado positivo em participações societárias e em sociedades em conta de participação						034			,00
Demais receitas operacionais (recuperação de despesas, etc.)						035			,00
Outras receitas - lucro na alienação de bens do ativo imobilizado, etc.....						036			,00
Total (031 + 032 + 033 + 034 + 035 + 036)						037			,00

08	RECEITA LÍQUIDA DE REVENDA NO ANO, SEGUNDO CLASSE DE COMÉRCIO Preencher somente se o Código 021 estiver preenchido		CÓD	VALORES EM REAIS	
Varejo (vendas ao público em geral).....			038	<input type="text"/>	<input type="text"/>
Atacado (exportações, vendas a instituições públicas, comerciais, indústrias, profissionais liberais, autônomos, agricultores, prestadores de serviços - restaurantes, bares, oficinas, etc.).....			039	<input type="text"/>	<input type="text"/>
Total			040	<input type="text"/>	<input type="text"/>
09	DESTINO DAS VENDAS DE MERCADORIAS Preencher somente se o Código 021 estiver preenchido		CÓD	PERCENTUAL	
Exportação direta - verificar correspondência com o Cód. 039			041	<input type="text"/>	<input type="text"/>
Vendas a outras empresas - verificar correspondência com o Cód. 039			042	<input type="text"/>	<input type="text"/>
Vendas para órgãos, empresas e instituições públicas - verificar correspondência com o Cód. 039			043	<input type="text"/>	<input type="text"/>
Vendas ao público em geral - verificar correspondência com o Cód. 038 (varejo)			044	<input type="text"/>	<input type="text"/>
Total			045	100%	
10	COMPRAS NO ANO Não inclua o ICMS e o IPI quando recuperáveis. Preencher somente se o Código 021 estiver preenchido		CÓD	VALORES EM REAIS	
Mercadorias para revenda (inclusive mercadorias de cooperados. Inclua os fretes e seguros sobre as compras e não inclua as devoluções e os impostos recuperáveis)			046	<input type="text"/>	<input type="text"/>
Matérias-primas para fabricação própria (deve preencher Cód. 022)			047	<input type="text"/>	<input type="text"/>
Material de embalagem e outros materiais (de reposição, peças, etc.)			048	<input type="text"/>	<input type="text"/>
Total			049	<input type="text"/>	<input type="text"/>
11	ESTOQUES	CÓD	EM 31-12-2014	CÓD	EM 31-12-2015
			VALORES EM REAIS	VALORES EM REAIS	
Mercadorias para revenda (inclusive merc. de cooperados).....		050	<input type="text"/>	055	<input type="text"/>
Produtos acabados e em elaboração (ver Cód. 022)		051	<input type="text"/>	056	<input type="text"/>
Matérias-primas para fabricação própria (ver Cód. 022)		052	<input type="text"/>	057	<input type="text"/>
Material de embalagem e outros materiais (de reposição, peças, etc.)		053	<input type="text"/>	058	<input type="text"/>
Total		054	<input type="text"/>	059	<input type="text"/>
12	MÉTODO DE VALORAÇÃO DOS ESTOQUES	13	CUSTO DAS MERCADORIAS REVENIDAS Somente se o Cód.021 estiver preenchido	CÓD	VALORES EM REAIS
1- <input type="text"/> - Custo médio 2- <input type="text"/> - PEPS 3- <input type="text"/> - UEPS 4- <input type="text"/> - Outro		Custo das mercadorias revendidas (valor contábil apurado e levado à DRE)		060	<input type="text"/>
		(Cód. 046 + 050 - 055)			
14	GASTOS COM PESSOAL NO ANO Total pago ou devido		CÓD	VALORES EM REAIS	
Salários e outras remunerações (inclusive 13º salário, férias, horas extras, etc.) - não incluir bolsa paga a estagiários (Cód. 090)			061	<input type="text"/>	<input type="text"/>
Participação nos lucros (somente a parte distribuída aos empregados) e honorários da diretoria.....			062	<input type="text"/>	<input type="text"/>
Retiradas pró-labore do proprietário ou sócios com atividade na empresa (vide instrução)			063	<input type="text"/>	<input type="text"/>
Contribuições para a Previdência Social (parte da empresa, ou seja, não inclua a parte descontada dos salários dos empregados, nem os recolhimentos relativos aos autônomos)			064	<input type="text"/>	<input type="text"/>
FGTS (se tiver empregados e não preencher, justificar em OBSERVAÇÕES) - não incluir os 50% do FGTS pagos por motivo de dispensa (Cód. 067)			065	<input type="text"/>	<input type="text"/>
Contribuições para a previdência privada (parte da empresa, ou seja, não incluir a parte descontada dos salários dos empregados)			066	<input type="text"/>	<input type="text"/>
Indenizações trabalhistas, rescisões contratuais e por dispensa incentivada - inclusive os 50% do FGTS			067	<input type="text"/>	<input type="text"/>
Benefícios concedidos aos empregados (transporte, alimentação, auxílio-educação, planos de saúde, auxílio-doença, seguro de vida em grupo, treinamento, etc.).....			068	<input type="text"/>	<input type="text"/>
PIS sobre folha de pagamento			068A	<input type="text"/>	<input type="text"/>
Total			069	<input type="text"/>	<input type="text"/>

16	DESPESAS FINANCEIRAS E DE PARTICIPAÇÕES NO ANO				CÓD	VALORES EM REAIS			
Despesas com arrendamento mercantil (<i>leasing</i>) de máquinas, equipamentos e veículos.....					092				
Despesas financeiras - inclusive <i>factoring</i> , taxa de juros de longo prazo e despesas bancárias					093				
Variações monetárias passivas					094				
Comissões pagas a administradoras de cartões de crédito					095				
Resultado negativo em participações societárias e em sociedades em conta de participação					096				
Total					097				
17	DEPRECIÇÃO, AMORTIZAÇÃO, OUTRAS DESPESAS E CONSTITUIÇÃO DAS PROVISÕES NO ANO				CÓD	VALORES EM REAIS			
Depreciação e amortização (valores relativos ao ano - não inclua depreciação e amortização acumulada em nenhum código)					098				
Outras despesas (<i>impairment</i> , prejuízo na alienação de bens do ativo imobilizado, etc.).					099				
Provisão para IRPJ e CSLL e pagamentos relativos a esses tributos					100				
Outras provisões constituídas (contingências, devedores duvidosos, perdas com clientes, perdas com estoque, etc.)					101				
Total					102				
18	RESULTADO DO EXERCÍCIO APÓS A PROVISÃO PARA IMPOSTO DE RENDA E CONTRIBUIÇÃO SOCIAL SOBRE LUCRO LÍQUIDO Preencher se for optante do Lucro Real, Presumido ou Arbitrado				CÓD	VALORES EM REAIS			
Lucro (030 + 037) > [(049 + 054 - 059) + (069 + 091 + 097 + 102)]					103				
Prejuízo (030 + 037) < [(049 + 054 - 059) + (069 + 091 + 097 + 102)]					104				
19	AQUISIÇÕES (exceto <i>leasing</i>), PRODUÇÃO PRÓPRIA E MELHORIAS DO ATIVO IMOBILIZADO NO ANO VALORES EM REAIS								
	CÓD	AQUISIÇÕES DE TERCEIROS	CÓD	PRODUÇÃO PRÓPRIA REALIZADA NO ATIVO IMOBILIZADO	CÓD	MELHORIAS			
Terrenos	105				253				
Edificações	105A		248		254				
Máquinas, equipamentos e instalações - exclusive processamento de dados ...	106		249		255				
Microcomputadores e equipamentos de telecomunicações.....	106A		250		256				
Meios de transporte	107				257				
Outros (móveis, utensílios, etc.)	108		251		258				
Total	109		252		259				
19A	BAIXAS DO ATIVO IMOBILIZADO NO ANO				CÓD	VALORES EM REAIS			
Terrenos					110				
Edificações					110A				
Máquinas, equipamentos e instalações - exclusive processamento de dados					111				
Microcomputadores e equipamentos de telecomunicações.....					111A				
Meios de transporte					112				
Outros (móveis, utensílios, etc.)					113				
Total					114				
19B	ATIVO IMOBILIZADO		CÓD	VALOR EM REAL 2014	CÓD	VALOR EM REAL 2015			
Total do Ativo Imobilizado			260		261				
20	FORMA DE REALIZAÇÃO DAS VENDAS - SISTEMA DE COMERCIALIZAÇÃO Preencher somente se o Código 021 estiver preenchido				CÓD	PERCENTUAL Relativo a Receita Bruta - (Código 021)			
Lojas (vendas no balcão), estabelecimentos, postos de combustíveis, boxes em mercados, depósitos, galpões, armazéns e salas					115				
Quiosques, <i>trailers</i> e outros estabelecimentos situados em local fixo fora da loja em estradas, praças, rodoviárias, corredores de <i>shopping centers</i> , etc.					116				
Correio					117				
Porta a porta (domicílio), postos móveis ou por ambulantes					118				
Internet					119				
Televendas					120				
Outros (especificar):					121				
Total					122	100 %			

III - DADOS DE REGIONALIZAÇÃO

UNIDADES DA FEDERAÇÃO	21		DE TODA EMPRESA			
	OS CÓDIGOS 150 E 178 DEVEM CORRESPONDER AOS CÓDIGOS 020 E (061 + 062 + 063), RESPECTIVAMENTE					
	PESSOAL OCUPADO EM 31/12/15		SALÁRIOS, RETIRADAS E OUTRAS REMUNERAÇÕES NO ANO			
	CÓD	Nº DE PESSOAS	CÓD	VALORES EM REAIS		
NORTE						
Rondônia	123		151			,00
Acre	124		152			,00
Amazonas	125		153			,00
Roraima	126		154			,00
Pará	127		155			,00
Amapá	128		156			,00
Tocantins	129		157			,00
NORDESTE						
Maranhão	130		158			,00
Piauí	131		159			,00
Ceará	132		160			,00
Rio Grande do Norte ..	133		161			,00
Paraíba	134		162			,00
Pernambuco	135		163			,00
Alagoas	136		164			,00
Sergipe	137		165			,00
Bahia	138		166			,00
SUDESTE						
Minas Gerais	139		167			,00
Espírito Santo	140		168			,00
Rio de Janeiro	141		169			,00
São Paulo	142		170			,00
SUL						
Paraná	143		171			,00
Santa Catarina	144		172			,00
Rio Grande do Sul	145		173			,00
CENTRO-OESTE						
Mato Grosso do Sul ..	146		174			,00
Mato Grosso	147		175			,00
Goiás	148		176			,00
Distrito Federal	149		177			,00
Total	150		178			,00

22		DA ATIVIDADE DE COMÉRCIO			
O CÓDIGO 234 DEVE CORRESPONDER AO CÓDIGO 021 + 021A					
UNIDADES LOCAIS COM RECEITA DE REVENDA OU DE REPRES. COMERCIAL		RECEITA BRUTA DE REVENDA OU DE REPRESENTAÇÃO COMERCIAL NO ANO			
CÓD	Nº DE UNID. LOCAIS	CÓD	VALORES EM REAIS		
179		207			,00
180		208			,00
181		209			,00
182		210			,00
183		211			,00
184		212			,00
185		213			,00
186		214			,00
187		215			,00
188		216			,00
189		217			,00
190		218			,00
191		219			,00
192		220			,00
193		221			,00
194		222			,00
195		223			,00
196		224			,00
197		225			,00
198		226			,00
199		227			,00
200		228			,00
201		229			,00
202		230			,00
203		231			,00
204		232			,00
205		233			,00
206		234			,00

AUTENTICAÇÃO		
Após verificar se as informações foram prestadas em REAIS e SEM CENTAVOS , se estão corretas e sem rasuras		
Data de entrega: ____/____/____	Data de devolução: ____/____/____	
Responsável pelo preenchimento:	Cargo:	
Tel.:	Ramal:	Fax:
Assinatura: Certifico que as informações aqui contidas são completas e estão de acordo com as demonstrações financeiras da empresa.		
Assinatura do Técnico de Pesquisas:		
OBSERVAÇÕES		
.....		
NOTA		
A empresa pode optar por prestar informações por questionário eletrônico, cujo software está disponível no <i>site</i> do IBGE: www.ibge.gov.br		
COMENTÁRIOS E/OU SUGESTÕES		
Que comentários sobre a pesquisa e sugestões de melhorias o(a) senhor(a) teria a fazer?		
Obrigado pela dedicação em preencher este questionário. Por favor, retenha uma cópia para seu controle. Os resultados finais desta pesquisa serão divulgados na <i>homepage</i> do IBGE.		
INSTRUÇÕES DE PREENCHIMENTO		
<p>Este questionário aplica-se às empresas com atividade predominantemente comercial, ou seja, cujo maior faturamento advém da atividade comercial, registradas no Cadastro Nacional das Pessoas Jurídicas - CNPJ - do Ministério da Fazenda.</p> <p>Como atividade comercial entende-se a compra de mercadorias para revenda, sem transformações, a representação comercial e agentes do comércio.</p> <p>I - INFORMAÇÕES CADASTRAIS</p> <p>Preencha os Dados Cadastrais da Sede da Empresa e da Unidade de Coleta utilizando somente endereços pertencentes à empresa. Não preencha com o endereço do contador terceirizado. Caso seja Unidade única, repetir o endereço da Sede.</p> <p>CAPÍTULO 03 - DADOS CADASTRAIS</p> <p>Item 02 - Nome fantasia - informar o nome pelo qual a empresa é conhecida. Caso a empresa possua lojas, filiais ou locais de venda com nome fantasia diferenciados, registre o nome fantasia pelo qual a empresa é mais conhecida. Caso a empresa não possua nome fantasia, registre S/D (Sem Denominação).</p> <p>Item 15 - Descrição da atividade de maior receita - informar com o maior nível de detalhamento possível a atividade desenvolvida pela empresa. Exemplos: comércio varejista de roupas, comércio atacadista de produtos alimentícios, comércio varejista de peças para automóveis, farmácia, açougue, ótica, papelaria, livraria, supermercado, posto de combustíveis, sapataria, loja de material de construção, representação comercial de material de construção, de veículos automotores, de combustíveis, etc.</p> <p>Item 16 - Código CNAE 2.0 - informe o código correspondente à descrição da atividade de maior faturamento da empresa. O código pode ser consultado na Relação de Atividades, em anexo, neste questionário ou pela internet no endereço http://www.cnae.ibge.gov.br.</p> <p>Item 17 - Sufixo e DV da Unidade de Coleta - corresponde ao endereço da Unidade designada para preencher o questionário. Se o endereço for de outra empresa do Grupo Empresarial, deverá ser preenchido com o Sufixo 9999 e o Dígito Verificador (DV) será, automaticamente, calculado pelo sistema.</p> <p>ATENÇÃO: não informar endereço do contador terceirizado/contratado.</p> <p>CAPÍTULO 04 - DADOS CADASTRAIS COMPLEMENTARES</p> <p>Item 01 - Situação cadastral em 31-12-15 - registre nas quadrículas o código da situação da empresa, conforme definido abaixo:</p> <p>01 - Em operação - se a empresa estava funcionando em 31.12.15.</p> <p>03 - Paralisada com informação de receita - se a empresa estava com sua(s) atividade(s) paralisadas em 31.12.15, porém tenha funcionado durante um período do ano e auferido receita.</p> <p>04 - Extinta com informação de receita - se a empresa estava com sua(s) atividade(s) encerradas em 31.12.15, porém tenha funcionado durante um período do ano e auferido receita.</p>		

Item 01A - Data da Ocorrência (somente para as Situações Cadastrais 03 e 04) - informar o mês e o ano da paralisação ou extinção da empresa.

Itens 02 e 03 - Mudanças na Estrutura da Empresa - caso tenha ocorrido algum processo de mudança na estrutura da empresa em 2015, devem ser preenchidos com o código correspondente e o CNPJ de ligação, respectivamente. Caso não tenha ocorrido no ano de 2015, preencher com **00 - Não houve mudanças**.

Item 04 - Principais Mercadorias Revendidas - descreva, detalhadamente, as três principais mercadorias revendidas de acordo com a predominância na receita bruta do comércio.

Item 05 - Forma de tributação - registre na quadrícula o código correspondente a forma de tributação utilizada: **1 - Lucro Real**; **2 - Lucro Presumido ou Arbitrado**; **3 - Simples Nacional**, se a empresa estiver na condição de Microempresa - ME ou Empresa de Pequeno Porte - EPP, na forma da Lei Complementar nº 123/2006 que instituiu, a partir de 01-07-2007, novo tratamento tributário simplificado também conhecido como Simples Nacional ou Super Simples. Tal regime substituiu, a partir de 01-07-07, o Simples Federal (Lei 9.317/1996) que foi revogada a partir daquela data; ou **4 - Imune ou Isenta**.

II - INFORMAÇÕES DA EMPRESA

CAPÍTULO 05 - NÚMERO DE PESSOAS OCUPADAS

Informar o número de pessoas efetivamente ocupadas em **31-03, 30-06, 30-09 e 31-12-15**, independente de terem ou não vínculo empregatício, desde que tenham sido remuneradas pela empresa, com exceção dos membros da família sem remuneração.

Inclua as pessoas afastadas em gozo de férias ou por motivo de licenças. Não considerar as pessoas que se encontram afastadas por licenças, ou pelo seguro por acidentes por mais de 15 dias. Não incluir os estagiários contratados pela empresa.

Pessoal ligado à atividade comercial - informar o número de pessoas alocadas nas operações de revenda e distribuição de mercadorias, tais como: vendedores, balconistas, estoquistas, gerente de vendas, empacotadores, frentistas, caixas, carregadores, supervisores de venda, encarregados de compra, motoristas e ajudantes (caso a empresa possua frota própria para transporte de mercadorias), etc.

Pessoal não-ligado à atividade comercial - informar o número de pessoas que não atuaram diretamente no processo de comercialização, tais como: presidente, diretores, pessoal administrativo, pessoal ligado à atividade industrial e a outras atividades.

Proprietário e sócios - informar todos os sócios e proprietário com atividade na empresa.

CAPÍTULO 06 - DEMONSTRATIVO DA RECEITA NO ANO

021 a 025 - Receita Bruta - considere as receitas brutas provenientes da revenda de mercadorias, de comissões sobre vendas de representação comercial e agentes de comércio, da venda de produtos de fabricação própria, de serviços de manutenção e reparação, de royalties de franquia e de outras atividades, sem deduções dos impostos e contribuições que incidam diretamente sobre essas receitas (ICMS, IPI, ISS, PIS, COFINS, etc.), das vendas canceladas, abatimentos e descontos incondicionais, que devem ser informados nos **Cód. 026 a 029**.

021A - Comissões sobre vendas de representantes comerciais e agentes do comércio - informar as receitas brutas provenientes do recebimento de comissões sobre as vendas efetuadas através de representação comercial e/ou agentes do comércio.

No caso específico das Cooperativas considere como receita de revenda a venda de produtos dos cooperados, bem como a venda de produtos que a própria Cooperativa efetuou aos cooperados, tais como: sementes, fertilizantes, rações, produtos veterinários, ferramentas, gêneros alimentícios, combustíveis, etc.

Se quaisquer dos códigos de **022 a 025** superar em valor o **Cód. 021 + 021A**, favor discriminar em **OBSERVAÇÕES** qual atividade gerou esta receita.

029 - Sistema SIMPLES NACIONAL - caso a empresa tenha optado pelo sistema de tributação SIMPLES NACIONAL, registre no código 029 o total do pagamento anual unificado dos impostos e contribuições relativos ao ano de 2015. O sistema SIMPLES NACIONAL implica o pagamento, em uma única guia de recolhimento, dos seguintes impostos e contribuições: Imposto de Renda Pessoa Jurídica - IRPJ, PIS, COFINS, IPI, Previdência Social, ICMS, ISS e Contribuição Social sobre o Lucro Líquido - CSLL.

CAPÍTULO 07 - DEMAIS RECEITAS NO ANO

032 - Receitas financeiras - informar as receitas realizadas no período-base relativas a juros recebidos, descontos obtidos, lucro na operação de resgate e prêmio de resgate de títulos e debêntures. Inclua também neste código o valor total dos rendimentos auferidos com aplicações em títulos de correção pré-fixada, bem como a atualização monetária pré-fixada.

033 - Variações monetárias ativas - considere como variações monetárias ativas as variações nas taxas de câmbio e as variações monetárias pós-fixadas decorrentes de atualizações dos direitos de crédito, calculadas com base em índices ou coeficientes aplicáveis por disposição legal ou contratual.

034 - Resultado positivo em participações societárias e em sociedades em conta de participação - informar o resultado positivo em participações em controladas e coligadas obtido pelo método da equivalência patrimonial.

035 - Demais receitas operacionais - considere as demais receitas operacionais, tais como: recuperação de despesas, recuperação de fretes, etc.

036 - Outras receitas - informar as receitas resultantes de operações não incluídas nas atividades principais e acessórias da empresa, tais como: o lucro na alienação de bens, ou seja, o resultado positivo obtido na venda de bens do ativo imobilizado, representado pela diferença entre o valor de venda e o valor contábil (custo histórico, corrigido e depreciado).

CAPÍTULO 08 - RECEITA LÍQUIDA DE REVENDA NO ANO, SEGUNDO CLASSE DE COMÉRCIO

Informar a receita líquida de revenda de acordo com as seguintes definições:

038 - Varejo - venda de produtos novos ou usados destinados predominantemente a pessoas físicas, para consumo, uso pessoal ou doméstico, independente da natureza e quantidade vendida.

039 - Atacado - venda de produtos que serão utilizados no processo produtivo, isto é, destinam-se a instituições públicas, exportações, revendedores, indústrias, profissionais autônomos, agricultores, prestadores de serviços (restaurantes, bares, oficinas, etc.).

CAPÍTULO 10 - COMPRAS NO ANO

046 - Mercadorias para revenda - informar o valor das mercadorias adquiridas para a revenda no mercado interno e externo, pagas à vista ou a prazo. Inclua no custo das mesmas as despesas de seguro e de transporte até o estabelecimento, os impostos não recuperáveis, taxas aduaneiras, além de comissões pagas a terceiros pela aquisição das mercadorias. Informar o valor líquido das compras, ou seja, não inclua as devoluções, abatimentos e descontos obtidos, o ICMS e o IPI quando recuperáveis.

No caso específico das Cooperativas considere como compra de mercadorias para revenda, a compra de produtos dos cooperados, bem como a compra de produtos para a venda aos cooperados.

CAPÍTULO 11 - ESTOQUES

050 - Estoque em 31-12-14 - corresponde às mercadorias não vendidas no ano de 2014, inventariadas e contabilizadas, passando então para o exercício pesquisado como **estoque inicial**.

055 - Estoque em 31-12-15 - corresponde às mercadorias não vendidas no ano de 2015, inventariadas e contabilizadas no exercício pesquisado como **estoque final**.

Os códigos referentes aos estoques de produtos acabados e em elaboração (**051 e 056**) e de matérias-primas para fabricação própria (**052 e 057**), devem ser preenchidos caso a empresa seja mista, isto é, tenha atividade comercial e industrial. O preenchimento de um desses códigos, implica o preenchimento do **Código 022**.

CAPÍTULO 12 - MÉTODO DE VALORAÇÃO DOS ESTOQUES

Informar o método ou critério utilizado pela empresa para determinar o custo das mercadorias estocadas, tendo em vista que a empresa poderá adquirir um mesmo tipo de mercadorias em datas diferentes, pagando por elas preços variados. Os métodos de valoração ou avaliação dos estoques que foram destacados neste capítulo estão definidos abaixo, no entanto, existem outros métodos de valoração que poderão ser utilizados. Caso a empresa utilize outro método, informe-o no campo de observações:

PEPS - a sigla PEPS significa "primeiro que entra, primeiro que sai", e é também conhecido como FIFO, iniciais da frase em inglês "first in, first out". Adotando este critério para valoração dos seus estoques, a empresa atribuirá às mercadorias estocadas os custos mais recentes.

UEPS - a sigla UEPS significa "último que entra, primeiro que sai", e é também conhecida como LIFO, iniciais da frase em inglês "last in, first out". Adotando este critério para valoração dos seus estoques, a empresa sempre atribuirá às suas mercadorias em estoque os custos mais antigos, guardadas as devidas proporções com as mercadorias que entraram e saíram da empresa.

Custo Médio - é o método mais utilizado no Brasil e consiste em avaliar os estoques pelo custo médio de aquisição apurado em cada entrada de material, ponderado pelas quantidades adicionadas e pelas anteriormente existentes na empresa.

CAPÍTULO 13 - CUSTO DAS MERCADORIAS REVENDIDAS

060 - Custo das mercadorias revendidas - Informar o valor contábil apurado e lançado na Demonstração do Resultado do Exercício. Observar o fechamento com a seguinte equação: **Códigos (046 + 050 - 055)**.

CAPÍTULO 14 - GASTOS COM PESSOAL NO ANO

061 - Salários e outras remunerações - informar o total das importâncias pagas a título de salários fixos, comissões sobre vendas, horas extras, ajudas de custo, 13º salário, abono financeiro de 1/3 das férias, etc. Não deduza as parcelas correspondentes às cotas de Previdência e Assistência Social (INSS) ou de consignação de interesse dos empregados (aluguel de casa, contas de cooperativas, etc.). Não inclua diárias pagas a empregados em viagens e participações ou comissões pagas a profissionais autônomos. Não inclua as indenizações pagas pelas empresas por motivos de dispensa (aviso prévio, 50% do FGTS, férias proporcionais, planos de demissão voluntária, etc.), que devem ser informados no **Cód. 067**. Não incluir bolsa paga a estagiários (**Cód. 090**).

062 - Participação nos lucros (somente a parte distribuída aos empregados) e honorários da diretoria - informar a parcela do lucro líquido distribuída aos funcionários, bem como os honorários da diretoria e dos membros dos conselhos fiscal, consultivo ou deliberativo.

063 - Retiradas pró-labore do proprietário ou sócios - considere as importâncias pagas a título de pró-labore dos sócios e do proprietário com atividade na empresa.

064 - Contribuições para a Previdência Social - informar as despesas da empresa e de competência do ano de referência da pesquisa, independente de terem sido pagas ou não. Não incluir a parte descontada dos salários dos empregados.

065 - FGTS - informar os valores de competência do ano de referência da pesquisa, independente de terem sido pagos ou não. Caso tenha empregados o preenchimento é obrigatório, com exceção dos casos em que não existe vínculo empregatício. Lembramos que a opção pelo SIMPLES NACIONAL não desobriga a empresa de preencher este código. Não incluir os 50% do FGTS pago por dispensa, que deve ser informado no **Cód. 067**.

067 - Indenizações trabalhistas, rescisões contratuais e por dispensa incentivada - correspondem às obrigações da empresa por ocasião da dispensa do empregado, tais como: aviso prévio, 50% do FGTS, 13º salário e férias proporcionais, planos de demissão voluntária (incentivo a demissões), acordos judiciais, etc.

068A - PIS sobre folha de pagamento - informar as contribuições incidentes sobre o total da folha de pagamento de seus empregados, quando a empresa for contribuinte nesta modalidade.

CAPÍTULO 15 - DESPESAS OPERACIONAIS NO ANO

Observar que não devem ser registradas neste capítulo, as despesas e custos registrados em capítulos anteriores, tais como: custo das mercadorias revendidas, gastos com pessoal, compras de matérias-primas, etc; bem como as Provisões para Imposto de Renda e Contribuição Social.

071 - Condomínios - inclusive taxas de administração de *shopping*, galerias, etc. - devem ser incluídas as taxas de administração no caso da empresa ter lojas localizadas em *shopping centers*.

072 - Comissões pagas a representantes comerciais e outros intermediários - comissões pagas a representantes e outros intermediários ligados à revenda de mercadorias. Não incluir as comissões sobre vendas pagas aos empregados da empresa, que devem ser informadas como salários.

073 - Publicidade e propaganda (feiras, eventos, lançamento de produtos, promoções de vendas, material promocional e marketing) - registrar as despesas relativas à divulgação e à promoção externa de produtos/serviços, atividades ou fatos de interesse de uma empresa, através da sua veiculação nos meios de comunicação. Por exemplo: gastos com eventos, feiras, brindes, anúncios em jornais, revistas, rádios, televisões e vias públicas, campanha promocional da empresa, etc.

074 - Combustíveis, peças e acessórios consumidos nos meios de transporte próprios, inclusive o consumo de forma geral (veículos, geradores, empilhadeiras, atividade administrativa, etc.) - registrar todas as despesas relativas ao consumo de combustíveis, sejam na frota própria, veículos alugados, equipamentos utilizados e nas atividades administrativas.

075 - Serviços prestados por profissionais liberais ou autônomos (pessoas físicas) - inclusive os carreteiros, despachantes, etc. - registrar as despesas pagas ou creditadas a profissionais liberais ou autônomos que não estiveram realizando revenda de mercadorias, tais como: bombeiros hidráulicos, eletricitas, diaristas em geral, contadores, advogados, etc. As despesas pagas a profissionais autônomos ligados à revenda de mercadorias devem ser informadas no **Cód. 072**.

076 - Serviços técnico-profissionais - registrar as despesas pagas ou creditadas a empresas prestadoras de serviços, tais como: consultoria em informática, serviços de auditoria, contábeis, jurídicos, consultoria, pesquisa de mercado, etc. Observe que as despesas pagas a empresas prestadoras de serviços de vigilância, segurança e transporte de valores devem ser informadas no **Cód. 077**.

Assim como as despesas pagas às empresas prestadoras de serviços de manutenção e reparação de imóveis, veículos em geral, instalações, máquinas e equipamentos devem estar informadas no **Cód. 080**.

078 - Fretes e carretos (não inclua os fretes sobre as compras - Cód. 046) - registrar as despesas pagas ou creditadas a empresas prestadoras de serviços de fretes e carretos.

079 - Mão de obra contratada temporariamente junto a empresas locadoras de mão de obra - registrar as despesas pagas à empresas locadoras de mão de obra pelo fornecimento temporário de mão de obra, tais como: secretárias, contínuos, pessoal de escritório, recepcionistas, telefonistas, etc.

081 - Outros serviços prestados por empresas - registrar as despesas pagas ou creditadas a empresas prestadoras de serviços que ainda não foram informadas nos Códigos referentes a serviços prestados por empresas (**Cód. 076 a 080**). Por exemplo: limpeza, zeladoria, portaria, dedetização, cobranças, etc.

085 - Prêmio de seguros - registrar as parcelas de prêmios de seguros de competência do ano da pesquisa, relativos aos bens de propriedade da empresa, tais como imóveis, veículos, mercadorias em estoque, bem como de responsabilidade civil.

089 - Royalties pelo uso de marcas, patentes e franquias - informar as despesas anuais decorrentes da utilização de marcas de terceiros, bem como aquelas oriundas de contratos de franquia. Exemplo: percentual sobre o faturamento pago ao franqueador.

089B - PIS/COFINS sobre outras receitas - informar as contribuições incidentes sobre as receitas integrantes do **Capítulo 07 - Outras Receitas (Cód. 031 a 036)**.

090 - Demais despesas operacionais - devem ser informadas as despesas operacionais não registradas nos Cód. 070 a 089, tais como: contribuições a associações de classe, assinaturas de revistas e publicações, indenizações trabalhistas não vinculadas com demissões, despesas com cartórios, despesas indedutíveis, doações, donativos, títulos incobráveis, quebra de estoque, bonificações e garantias, etc. Quando esta despesa ultrapassar 30% do **Cód. 091**, discrimine os três principais valores nos campos reservados para este fim.

Devem ser informados em outros códigos: Pagamentos de IRPJ e CSLL (**Cód. 100**); Provisão para Imposto de Renda e Contribuição Social sobre o Lucro Líquido (**Cód. 100**); Outras provisões constituídas (contingências, férias, provisão para devedores duvidosos, 13º salário, perdas de estoque, perdas com clientes, etc.) (**Cód. 101**); REFIS - apenas o valor dos juros de atualização será considerado no questionário e informado como despesas financeiras (**Cód. 093**); Juros sobre Capital Próprio (empresas que não sejam S/A) informar no **Cód. 093**.

Não informar em nenhum capítulo do questionário: Juros sobre Capital Próprio de companhias abertas (a deliberação CVM nº 207/1996 determina às companhias abertas (S/A) a contabilização desses juros diretamente em conta de lucros acumulados, sem afetar o resultado do exercício); valores de depreciação e amortização acumuladas; e recolhimentos da Previdência Social relativos aos autônomos.

CAPÍTULO 16 - DESPESAS FINANCEIRAS E DE PARTICIPAÇÕES NO ANO

093 - Despesas financeiras - informar as despesas relativas a juros de empréstimos, financiamentos, descontos de títulos, comissões e despesas bancárias e os descontos concedidos a clientes por pagamento antecipado de duplicata e outros títulos, o deságio na colocação de debêntures ou outros títulos.

094 - Variações monetárias passivas - considere as despesas com variações cambiais e a correção monetária passiva.

095 - Comissões pagas a administradoras de cartão de crédito - informar as comissões pagas às administradoras de cartão de crédito pela utilização do cartão como forma de recebimento das vendas efetuadas pela empresa, através de seus estabelecimentos comerciais.

096 - Resultado negativo em participações societárias e em sociedades em conta de participação - informar o resultado negativo em participações em controladas e coligadas obtido pelo método da equivalência patrimonial.

CAPÍTULO 17 - DEPRECIAÇÃO, AMORTIZAÇÃO, OUTRAS DESPESAS E CONSTITUIÇÃO DAS PROVISÕES NO ANO

098 - Depreciação e amortização - informar as parcelas anuais consideradas como despesas ou custo na Demonstração do Resultado do Exercício. Não registrar a depreciação e amortização acumulada.

099 - Outras despesas - informar as despesas resultantes de transações não incluídas nas atividades principais ou acessórias da empresa, tais como: o prejuízo na alienação de bens, ou seja, o resultado negativo obtido na venda de bens do ativo imobilizado, que consiste na diferença entre o valor de venda do bem e o valor contábil (custo histórico, corrigido e depreciado).

Inclua despesas com redução nos valores dos ativos - *impairment*, de acordo com a legislação contábil instituída pelas leis 11.638/07 e 11.941/09.

100 - Provisão para IRPJ, CSLL e pagamentos relativos a esses tributos - informar as despesas contabilizadas com a constituição das provisões para Imposto de Renda e Contribuição Social sobre o Lucro Líquido e os pagamentos relativos ao IRPJ e CSLL.

101 - Outras provisões constituídas - informar as despesas contabilizadas com a constituição de provisões operacionais e não operacionais, como, por exemplo: perdas prováveis na realização de investimentos, contingências e perdas de estoque, etc.

CAPÍTULO 18 - RESULTADO DO EXERCÍCIO APÓS A PROVISÃO PARA O IMPOSTO DE RENDA E CONTRIBUIÇÃO SOCIAL SOBRE O LUCRO LÍQUIDO

103/104 - Lucro/Prejuízo - informar o valor contábil apurado na Demonstração do Resultado do Exercício e relativo ao resultado após as Provisões para Imposto de Renda e Contribuição Social. Este capítulo só deverá ser preenchido pelas empresas optantes do Lucro Real, Presumido ou Arbitrado. As optantes pelo Simples Nacional estão dispensadas deste preenchimento.

CAPÍTULO 19 - AQUISIÇÕES (exceto *leasing*) PRODUÇÃO PRÓPRIA E MELHORIAS DO ATIVO IMOBILIZADO NO ANO

105 a 108 - Aquisições de Terceiros (exceto *leasing*) - considere como aquisições todos os bens adquiridos e incorporados ao Ativo Imobilizado da empresa, pelo custo histórico, durante o ano. Não inclua os encargos financeiros decorrentes de financiamento.

248 a 251 - Produção Própria - considere os bens oriundos da produção própria, e incorporados ao Ativo Imobilizado no ano.

253 a 258 - Melhorias - considere como melhorias as benfeitorias e os melhoramentos que tenham aumentado a vida útil dos bens, e que tenham sido incorporados ao Ativo Imobilizado no ano.

CAPÍTULO 19A - BAIXAS DO ATIVO IMOBILIZADO NO ANO

110 a 113 - Baixas - registre todos os bens do Ativo Imobilizado baixados pelo valor contábil, isto é, custo histórico corrigido e depreciado, durante o ano.

CAPÍTULO 19B - ATIVO IMOBILIZADO

260 e 261 - Total do Ativo Imobilizado - registre o saldo do Ativo Imobilizado evidenciado no Balanço Patrimonial dos anos 2013 e 2014, respectivamente.

CAPÍTULO 20 - FORMA DE REALIZAÇÃO DAS VENDAS - SISTEMA DE COMERCIALIZAÇÃO

Deverá ser informado o percentual, segundo a forma de comercialização, relativo a receita bruta de revenda no ano, informada no capítulo 06, código 021.

115 - Lojas (vendas no balcão), estabelecimentos, postos de combustíveis, boxes em mercados, depósitos, galpões, armazéns e salas - vendas realizadas em estabelecimentos localizados em prédios circundados por paredes e com entrada independente.

116 - Quiosques, trailers e outros estabelecimentos situados em local fixo fora da loja em estradas, praças, rodoviárias, corredores de shopping centers, etc. - vendas realizadas em estabelecimentos localizados em espaços públicos, como praças, calçadas, ruas de pedestre ou corredores de *shopping centers*.

117 - Correio - vendas realizadas através de correio com apoio ou não de catálogos que contêm a descrição dos produtos e seus preços.

118 - Porta a porta (domicílio), postos móveis ou por ambulantes - vendas realizadas por vendedores que se deslocam às casas dos consumidores potenciais (com apoio ou não de catálogos que contêm a descrição dos produtos e seus preços), se deslocam pela cidade ou se fixam nas ruas.

119 - Internet - vendas realizadas através de um site da rede internacional de computadores.

120 - Televendas - vendas realizadas por telefone onde a empresa toma a iniciativa de ligar para o cliente em potencial ou disponibiliza um número para contato.

121 - Outros - forma de comercialização não incluída nos códigos anteriores. Discriminar no campo específico.

III - DADOS DE REGIONALIZAÇÃO

Devem ser informados pela empresa, ainda que esta tenha exercido suas atividades em apenas uma Unidade da Federação. Os **Códigos 150, 178 e 234** devem corresponder aos **Códigos 020, (061 + 062 + 063) e (021 + 021A)**, respectivamente. Nos endereços da empresa onde foram exercidas, exclusivamente, atividades administrativas e de apoio (depósitos, almoxarifados, escritórios contábeis, etc.) informar apenas pessoal ocupado e salários.

Como Unidades Locais com receita de revenda, considere todos os endereços em que funcionaram lojas, filiais, locais de venda, pontos de vendas etc., ou seja, todos os endereços da empresa que efetuaram revenda de mercadorias.

A Receita Bruta de Revenda deve ser informada na Unidade da Federação onde ocorrer o fato gerador da receita, ou seja, na Unidade da Federação onde for realizada a revenda das mercadorias.

RELAÇÃO DE ATIVIDADES NO ÂMBITO DA PAC	
CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0	
CÓD	DESCRIÇÃO
COMÉRCIO A VAREJO, POR ATACADO E REPRESENTAÇÃO COMERCIAL DE VEÍCULOS AUTOMOTORES, MOTOCICLETAS, PEÇAS E ACESSÓRIOS	
4511-1	Comércio a varejo e por atacado de veículos automotores novos e usados - inclusive reboque e semi-reboques, ambulâncias, veículos betoneiras - não incluir bicicletas (4649-4 e 4763-6), tratores agrícolas (4661-3) e motocicletas (4541-2)
4512-9	Representações comerciais e agentes do comércio de veículos automotores - inclusive a venda sob consignação a varejo e por atacado - não incluir a representação comercial e agentes do comércio de peças e acessórios novos e usados para veículos automotores (4530-7)
4530-7	Representações comerciais, agentes do comércio e comércio a varejo e por atacado de peças e acessórios, novos e usados, para veículos automotores - não incluir para motocicletas (4541-2) e para tratores agrícolas (4661-3)
4541-2	Comércio por atacado e a varejo de motocicletas, peças e acessórios novos e usados
4542-1	Representantes comerciais e agentes do comércio de motocicletas, peças e acessórios - inclusive a venda sob consignação, a varejo e por atacado.
REPRESENTANTES COMERCIAIS E AGENTES DO COMÉRCIO EXCETO DE VEÍCULOS AUTOMOTORES	
4611-7	Matérias-primas agrícolas <i>in natura</i> e animais vivos (cana-de-açúcar, algodão, café em coco, bovinos, suínos, lã, couro bruto, chifres, ossos, crinas, etc.)
4612-5	Combustíveis, minerais, produtos siderúrgicos e químicos
4613-3	Madeira, material de construção, ferragens, material elétrico e material para pintura
4614-1	Máquinas, equipamentos, embarcações e aeronaves
4615-0	Eletrodomésticos, móveis e artigos de uso doméstico - não incluir máquinas e equipamentos de informática e de comunicação (4614-1)
4616-8	Têxteis, vestuário, calçados e artigos de viagens (fios naturais ou artificiais, fibras têxteis, artigos e complementos do vestuário e artigos em couro curtido).
4617-6	Produtos alimentícios, bebidas e fumo - não incluir ração para animais (4618-4); fumo não beneficiado (4611-7)
4618-4	Produtos especializados não especificados anteriormente (produtos farmacêuticos, cosméticos, perfumaria, jóias, brinquedos, bicicletas, papel de parede, ração, distribuidores de jornais e revistas para o comércio varejista sob contrato de comissão) - não incluir entregas de jornais e revistas em domicílio (5320-2); veículos automotores e auto-peças (45)
4619-2	Representantes comerciais e agentes do comércio de mercadorias em geral não especializados (representante sem predominância de algum tipo de mercadoria ou grupos de mercadorias específicas)
COMÉRCIO ATACADISTA DE MATÉRIAS-PRIMAS AGRÍCOLAS, ANIMAIS VIVOS; PRODUTOS ALIMENTÍCIOS PARA ANIMAIS	
4621-4	Comércio atacadista de café em grão, em coco ou verde
4622-2	Comércio atacadista de soja
4623-1	Comércio atacadista de animais vivos (bovinos, ovinos, suínos etc.), forragens, rações, couros, lã, peles, sêmen bovino e outros subprodutos de origem animal, algodão, fumo em folha não beneficiado, cacau, sementes, flores, plantas, sisal, cereais em bruto, fibras têxteis não beneficiadas - não incluir produtos hortifrutigranjeiros (4633-8) e cereais beneficiados (4632-0), aves vivas, coelhos e pequenos animais vivos para alimentação (4633-8)
COMÉRCIO ATACADISTA ESPECIALIZADO EM PRODUTOS ALIMENTÍCIOS, BEBIDAS E FUMO	
4631-1	Comércio atacadista de leite e laticínios
4632-0	Comércio atacadista de cereais e leguminosas beneficiados, farinhas, amidos e féculas
4633-8	Comércio atacadista de hortifrutigranjeiros - inclusive aves vivas, ovos e pequenos animais vivos para alimentação
4634-6	Comércio atacadista de carnes, produtos da carne, pescados e frutos do mar - inclusive produtos de salsicharia, aves abatidas e derivados
4635-4	Comércio atacadista de bebidas alcoólicas e não-alcoólicas
4636-2	Comércio atacadista de produtos do fumo (cigarros, cigarrilhas, fumo em rolo ou em corda beneficiado, inclusive isqueiros, cachimbos e demais artigos de tabacaria) - não incluir fumo em folha não beneficiado (4623-1)
4637-1	Comércio atacadista especializado em produtos alimentícios não especificados anteriormente (café torrado, moído e solúvel, açúcares, balas, chocolates, mel, sorvetes, produtos de padaria, massas, sucos, conservas de frutas e legumes, óleos e gorduras, condimentos, alimentos congelados, etc.)
4639-7	Comércio atacadista de produtos alimentícios em geral compreende a comercialização dos produtos das classes 4631-1 a 4637-1, porém sem predominância de nenhum deles
COMÉRCIO ATACADISTA DE PRODUTOS DE CONSUMO NÃO-ALIMENTAR	
4641-9	Comércio atacadista de tecidos, artefatos de tecidos, artigos de cama, mesa e banho e artigos de armarinho - não incluir colchoaria, persianas e cortinas (4649-4)
4642-7	Comércio atacadista de artigos do vestuário e acessórios - inclusive para uso profissional e de segurança do trabalho, roupas esportivas fardamento e uniforme
4643-5	Comércio atacadista de calçados e artigos de viagem - calçados de qualquer material, inclusive esportivo, bolsas, malas e outros artigos para viagens - não incluir calçados ortopédicos sob medida (3250-7)

CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0

CÓD	DESCRIÇÃO
4644-3	Comércio atacadista de produtos farmacêuticos, para uso humano e veterinário inclusive da flora medicinal - não incluir cosméticos, produtos de perfumaria e de higiene pessoal (4646-0)
4645-1	Comércio atacadista de instrumentos e materiais para uso médico, cirúrgico, hospitalares e de laboratórios, próteses e artigos de ortopedia e produtos odontológicos - não incluir equipamentos médico-cirúrgicos hospitalares (4664-8)
4646-0	Comércio atacadista de cosméticos, produtos de perfumaria e de higiene pessoal, inclusive fraldas descartáveis
4647-8	Comércio atacadista de artigos de escritório e de papelaria; livros, jornais e outras publicações; papel, papelão e seus artefatos - não incluir papel e papelão em bruto, embalagens de papel e papelão (4686-9), papel e papelão usados e aparas de papel (4687-7), suprimentos para informática (4651-6)
4649-4	Comércio atacadista de equipamentos elétricos, eletrônicos pessoal e de uso domésticos, móveis e outros artigos de usos pessoal e doméstico não especificados anteriormente (bicicletas, triciclos, artigos de tapeçaria, persianas, cortinas, lustres, luminárias, filmes para fotos, CDs, fitas e discos, produtos de higiene e limpeza domiciliar, jóias, relógios, bijuterias, pedras preciosas e semipreciosas lapidadas, artigos de cutelaria, artigos para habitação de vidro, cristal, porcelana, borracha, plástico, metal, madeira, vime, bambu e outros similares; brinquedos; instrumentos musicais; artigos de óptica, papel de parede, artigos para caça, pesca e <i>camping</i> , etc.) - não incluir equipamento e material de informática (4651-6), equipamentos de comunicação (4652-4), equipamentos de ginástica (4669-9) e de fisioterapia (4664-8), peças e acessórios para aparelhos eletrodomésticos (4689-3)
COMÉRCIO ATACADISTA DE EQUIPAMENTOS E PRODUTOS DE TECNOLOGIAS DE INFORMAÇÃO E COMUNICAÇÃO	
4651-6	Comércio atacadista de computadores, periféricos e suprimentos de informática (computadores, programas de computador não-customizáveis, <i>toner</i> , fitas, disquetes ópticos, etc) - não incluir componentes eletrônicos para equipamentos de informática (4652-4), máquinas para automação gerencial e comercial (4665-6), manutenção e reparação de computadores (9511-8)
4652-4	Comércio atacadista de componentes eletrônicos e equipamentos de telefonia e comunicação (válvulas e tubos eletrônicos, semicondutores, microchips e circuitos integrados, circuitos impressos, telefones, intercomunicadores, secretária eletrônica e similares)
COMÉRCIO ATACADISTA DE MÁQUINAS, APARELHOS E EQUIPAMENTOS, EXCLUSIVE DE TECNOLOGIA DE INFORMAÇÃO E COMUNICAÇÃO	
4661-3	Comércio atacadista de máquinas, aparelhos e equipamentos para uso agropecuário; partes e peças; arados, cultivadores, tratores para uso na agricultura e silvicultura, cortadores de grama, inclusive máquinas e equipamentos para avicultura e apicultura
4662-1	Comércio atacadista de máquinas, equipamentos para terraplenagem, mineração e construção - não incluir partes e peças de tratores para uso na agricultura e silvicultura (4661-3)
4663-0	Comércio atacadista de máquinas, aparelhos e equipamentos para uso industrial; partes e peças - máquinas operatrizes e máquinas ferramentas para uso nos diversos ramos de indústria - não incluir para mineração e construção (4662-1)
4664-8	Comércio atacadista de máquinas, aparelhos e equipamentos para uso odonto-médico-hospitalares e laboratoriais; partes e peças - mobiliário médico-hospitalares e odontológico, equipamentos para laboratórios, para monitoração médica, médico- cirúrgicos e para fisioterapia
4665-6	Comércio atacadista de máquinas e equipamentos para uso comercial; partes e peças - balanças, balcões frigoríficos, refrigeradores, vitrines, registradoras, máquinas de fatiar, emissoras de cupom fiscal, leitora óptica, máquinas para automação comercial
4669-9	Comércio atacadista de máquinas, aparelhos e equipamentos não especificados anteriormente; partes e peças - bombas, compressores e carneiros hidráulicos, motores e transformadores elétricos, sistema para controle de incêndio, instrumentos e equipamentos de medida, máquinas de costura para qualquer uso, máquinas e equipamentos para usos técnicos e profissional, máquinas e equipamentos para escritório, <i>robôs</i> .
COMÉRCIO ATACADISTA DE MADEIRA, FERRAGENS, FERRAMENTAS, MATERIAL ELÉTRICO E MATERIAL DE CONSTRUÇÃO	
4671-1	Comércio atacadista de madeira e produtos derivados - madeira em bruto perfilada ou serrada, tábuas, pranchas, dormentes, barrotes e similares, pré-moldados de madeira para construção - não incluir fabricação de pré-moldados com montagem (1622-6)
4672-9	Comércio atacadista de ferragens e ferramentas (ferragens para construção, ferramentas manuais) - não incluir produtos siderúrgicos e metalúrgicos para construção (vergalhões, arames e pregos)(4679-6), chapas de aço, bobinas (4685-1) e máquinas ferramentas industriais(4669-9)
4673-7	Comércio atacadista de material elétrico - fios, cabos, condutores elétricos, lâmpadas, tomadas, chaves elétricas - não incluir lustres, luminárias (4649-4) e peças e acessórios para aparelhos de usos doméstico e pessoal, elétricos e eletrônicos (4689-3)
4674-5	Comércio atacadista de cimento
4679-6	Comércio atacadista especializado de materiais de construção não especificados anteriormente e de materiais de construção em geral - tintas e materiais para pintura, mármore e granitos, vidros, espelhos e molduras, areia, pedra, tijolos, tubos e canos de água, aparelhos sanitários, vergalhões, arames, pregos, esquadrias metálicas, divisórias, azulejos, pisos, piscinas e equipamentos correlatos, etc.
COMÉRCIO ATACADISTA ESPECIALIZADO EM OUTROS PRODUTOS	
4681-8	Comércio atacadista de combustíveis sólidos, líquidos e gasosos - inclusive lubrificantes, lenha, carvão vegetal, carvão mineral, petróleo cru em bruto - não incluir gás natural veicular para postos de combustíveis (3520-4) e GLP (4682-6) e a venda a varejo de GNV (4731-8)
4682-6	Comércio atacadista de Gás Liquefeito de Petróleo (GLP)
4683-4	Comércio atacadista de defensivos agrícolas, adubos, fertilizantes e corretivos do solo - não incluir produtos farmacêuticos para uso veterinário (4644-3)

CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0

CÓD	DESCRIÇÃO
4684-2	Comércio atacadista de produtos químicos e petroquímicos (resinas e elastômeros, solventes, cloro, oxigênio, carga para extintor de incêndio, álcool etílico, artigos pirotécnicos, fósforo de segurança, adesivos, essências e concentrados aromáticos) - não incluir agroquímicos (4683-4)
4685-1	Comércio atacadista de produtos siderúrgicos e metalúrgicos (chapas de aço, bobinas etc.) - não incluir artigos para construção civil (4679-6)
4686-9	Comércio atacadista de papel e papelão em bruto e de embalagem
4687-7	Comércio atacadista de resíduos e sucatas metálicas e não metálicas - aparas de papel e papelão, sacos usados, plásticos, vidros usados
4689-3	Comércio atacadista especializado de outros produtos intermediários não especificados anteriormente - fios e fibras têxteis beneficiadas, produtos de extração mineral (minerais ferrosos e não ferrosos, sal gema e sal marinho, ouro e outros metais preciosos), inclusive peças e acessórios para aparelhos de uso doméstico e pessoal - não incluir combustíveis (4681-8) artefatos de borracha para veículos (4530-7) e para uso residencial (4649-4)
COMÉRCIO ATACADISTA NÃO-ESPECIALIZADO	
4691-5	Comércio atacadista de mercadorias em geral, com predominância de produtos alimentícios (não especializados, tais como, supermercado/hipermercado atacadista, distribuidores de cesta básica)
4692-3	Comércio atacadista de mercadorias em geral, com predominância de insumos agropecuários - não incluir o comércio especializado na venda de insumos agropecuários (ração no 4623-1; produtos veterinários no 4644-3; implementos agrícolas no 4661-3; adubos, fertilizantes para agricultura no 4683-4)
4693-1	Comércio atacadista de mercadorias em geral, sem predominância de produtos alimentícios ou de insumos agropecuários
COMÉRCIO VAREJISTA NÃO-ESPECIALIZADO	
4711-3	Comércio varejista de mercadorias em geral, com predominância de produtos alimentícios - hipermercado (área de venda superior a 5.000 m² e supermercado (área de venda entre 300 e 5.000 m²)
4712-1	Comércio varejista de mercadorias em geral, com predominância de produtos alimentícios (minimercados, mercearias, armazéns, empório, estivas, quitanda, com área de venda inferior a 300 m²)
4713-0	Comércio varejista de mercadorias em geral, sem predominância de produtos alimentícios - lojas de departamentos ou magazines, bazar, lojas de variedades, lojas <i>duty-free</i> de aeroportos internacionais, "lojas de R\$ 1,99", etc.
COMÉRCIO VAREJISTA DE PRODUTOS ALIMENTÍCIOS, BEBIDAS E FUMO	
4721-1	Comércio varejista de produtos de padaria, laticínios, frios e carnes conservadas, doces, balas, bombons e semelhantes, conservas de frutas, legumes, verduras e similares
4722-9	Comércio varejista de carnes e pescados (açougues e peixarias - inclusive pequenos animais abatidos, miúdos, vísceras e similares, pescados, crustáceos e moluscos, frescos, congelados, conservados ou frigorificados)
4723-7	Comércio varejista de bebidas alcoólicas e não-alcoólicas - não incluir bares, botequins e lanchonetes (5611-2)
4724-5	Comércio varejista de hortifrutigranjeiros, inclusive aves vivas, coelhos e outros pequenos animais para alimentação
4729-6	Comércio varejista de produtos alimentícios em geral ou especializados em produtos alimentícios não especificados anteriormente - artigos de tabacaria, comidas congeladas, produtos naturais e dietéticos, gelo, mel, sorvetes embalados em potes e similares, lojas de conveniência, lojas de delicatessen, etc.
COMÉRCIO VAREJISTA DE COMBUSTÍVEIS PARA VEÍCULOS AUTOMOTORES	
4731-8	Comércio varejista de combustíveis para veículos automotores - inclusive Gás Natural Veicular (GNV) - não incluir Gás Liquefeito de Petróleo (GLP) (4784-9) e lubrificantes (4732-6)
4732-6	Comércio varejista de lubrificantes para uso automotivo e para outros usos
COMÉRCIO VAREJISTA DE MATERIAL DE CONSTRUÇÃO	
4741-5	Comércio varejista de tintas e materiais para pintura - inclusive tintas automotivas
4742-3	Comércio varejista de material elétrico (fios cabos, condutores elétricos, chaves elétricas, interruptores, tomadas e similares) - não incluir peças e acessórios para aparelhos de uso domésticos e pessoal (4757-1) para veículos (45), lustres, luminárias e abajures (4754-7)
4743-1	Comércio varejista de vidros planos e de segurança e espelhos - não incluir para veículos (4530-7) e especializado em molduras de quadros (4789-0)
4744-0	Comércio varejista de ferragens, madeira e seus artefatos, material hidráulico e materiais de construção não especificados anteriormente e materiais de construção em geral sem especialização
COMÉRCIO VAREJISTA DE EQUIPAMENTOS DE INFORMÁTICA E COMUNICAÇÃO; EQUIPAMENTOS E ARTIGOS DE USO DOMÉSTICO	
4751-2	Comércio varejista especializado de equipamentos e suprimentos de informática (computadores e periféricos, programas de computador não-customizáveis, partes e peças para equipamentos de informática, recarga de cartuchos associado ao comércio) - não incluir equipamentos para escritório (4789-0)
4752-1	Comércio varejista de equipamentos de telefonia e comunicação - telefones, fax, intercomunicadores e similares, peças e partes para equipamentos de telefonia

RELAÇÃO DE ATIVIDADES NO ÂMBITO DA PAC		(conclusão)
CLASSIFICAÇÃO NACIONAL DE ATIVIDADES ECONÔMICAS - CNAE 2.0		
CÓD	DESCRIÇÃO	
4753-9	Comércio varejista especializado de eletrodomésticos e equipamentos de áudio e vídeo - não incluir equipamentos de informática (4751-2), de comunicação (4752-1) e peças e acessórios para eletroeletrônicos domésticos (4757-1)	
4754-7	Comércio varejista de móveis para qualquer uso, colchoaria e artigos de iluminação residencial - não incluir artigos para habitação (4759-8)	
4755-5	Comércio varejista especializado em tecidos, artigos de cama, mesa e banho e armário - não incluir cortinas, tapeçaria (4759-8)	
4756-3	Comércio varejista especializado de instrumentos musicais e acessórios, inclusive músicas impressas	
4757-1	Comércio varejista especializado de peças e acessórios para aparelhos eletroeletrônicos para uso doméstico - inclusive antenas - não incluir partes e peças para equipamentos de informática (4751-2) e comunicação (4752-1)	
4759-8	Comércio varejista de artigos de uso doméstico não especificados anteriormente (artigos de tapeçaria, cortinas e persianas, artigos para habitação, de vidro, cristal, porcelana, borracha, plástico, metal, madeira, vime, bambu, e outros similares, artigos de cutelaria, toldos e similares, papel de parede, sistema de segurança residencial - não incluir molduras e quadros e flores para ornamentação, objetos de arte (4789-0) antiguidades (4785-7)	
	COMÉRCIO VAREJISTA DE ARTIGOS CULTURAIS, RECREATIVOS E ESPORTIVOS	
4761-0	Comércio varejista de livros, jornais, revistas e papelaria - não incluir suprimentos de informática (4751-2) e livros usados (4785-7)	
4762-8	Comércio varejista de discos, fitas de áudio, vídeo, CDs e DVDs gravados ou não	
4763-6	Comércio varejista de artigos recreativos e esportivos (brinquedos e jogos eletrônicos, bicicletas e triciclos, artigos de caça, pesca e <i>camping</i> , artigos do vestuário e acessórios especializado para a prática de esportes, embarcações e outros veículos recreativos, motores para embarcações de esporte e lazer; peças e acessórios referente a esta classe - não incluir armas e munições (4789-0) e calçados esportivos (4782-2)	
	COMÉRCIO VAREJISTA DE PRODUTOS FARMACÊUTICOS, PERFUMARIA E COSMÉTICOS, ÓPTICOS E ORTOPÉDICOS	
4771-7	Comércio varejista de produtos farmacêuticos para uso humano e veterinário, inclusive homeopáticos, fitoterápicos e da flora medicinal - não incluir produtos odontológicos (4645-1), cosméticos, perfumaria e de higiene pessoal (4772-5)	
4772-5	Comércio varejista de cosméticos, produtos de perfumaria, de toucador e de higiene pessoal	
4773-3	Comércio varejista de artigos médicos e ortopédicos (muletas, cadeiras de rodas, aparelhos auditivos, kits diagnósticos etc) - não incluir equipamentos médico-cirúrgicos hospitalares (4664-8)	
4774-1	Comércio varejista de artigos de óptica - óculos, lentes	
	COMÉRCIO VAREJISTA DE PRODUTOS NOVOS NÃO ESPECIFICADOS ANTERIORMENTE E DE PRODUTOS USADOS	
4781-4	Comércio varejista de artigos do vestuário novos de qualquer material, acessórios e complementos do vestuário - não incluir roupas para segurança pessoal (4642-7); roupas e artigos do vestuário usados (4785-7), calçados e artigos de viagem (4782-2), artigos esportivos (4763-6)	
4782-2	Comércio varejista de calçados e artigos de viagem	
4783-1	Comércio varejista de jóias e relógios inclusive pedras preciosas e semi-preciosas lapidadas, inclusive peças e partes - não incluir bijuteria (4789-0)	
4784-9	Comércio varejista de Gás Liqüefeito de Petróleo - (GLP) em botijões	
4785-7	Comércio varejista de artigos usados - móveis, eletrodomésticos, objetos de arte, roupas e calçados, moedas e selos de coleção, livros e revistas, material de demolição, sacos, tambores - não incluir veículos usados, partes e peças (45)	
	COMÉRCIO VAREJISTA DE OUTROS PRODUTOS NOVOS NÃO ESPECIFICADOS ANTERIORMENTE	
4789-0	Comércio varejista de outros produtos novos não especificados anteriormente - suvenires, bijuterias, flores, plantas, vasos, material de limpeza, sementes e mudas para jardinagem, artesanato, objetos de arte; animais vivos e alimentos para animais de estimação; brindes; produtos eróticos; artigos para festas; artigos funerários; fogos de artifício, equipamentos para escritório, artigos fotográficos e para filmagem, armas e munições, artigos religiosos e de culto, rede de dormir, carvão e lenha, extintores, cartões telefônicos, cargas e preparados para incêndio, quinquilharias para usos agrícolas, etc.)	

Anexo C

Estatísticas

Tabela C.1: Número de empresas na base utilizada para testar os modelos de imputação por UF e ESTF em 2014

ESTF	UF				Total
	25 - PB	32 - ES	41 - PR	43 - RS	
A1	10	7	30	33	80
A2	18	16	61	65	160
A3	16	15	62	59	152
C	17	20	68	69	174
G	5	9	43	20	77
Total	66	67	264	246	643

Tabela C.2: Número de empresas na base utilizada para testar os modelos de imputação por UF e ESTF em 2015

ESTF	UF				Total
	25 - PB	32 - ES	41 - PR	43 - RS	
A1	7	7	32	36	82
A2	16	17	51	63	147
A3	16	17	56	56	145
C	17	19	68	69	173
G	5	9	39	26	79
Total	61	69	246	250	626

Tabela C.3: Frequência de Empresas por tipo de ação no questionário na PAC em 2014

Ação	Frequência	%
Nenhuma	82.385	98,92
Questionário imputado	342	0,41
Questionário transcrito	118	0,14
Empresa não selecionada considerada na expansão	260	0,31
Empresa fora do âmbito	174	0,21
Outras ações	4	0,00

Tabela C.4: *Frequência de Empresas por tipo de ação no questionário na PAC em 2015*

Ação	Frequência	Percent
Nenhuma	83686	99.20
Questionário imputado	231	0.27
Questionário transcrito	97	0.11
Empresa não selecionada considerada na expansão	1	0.00
Empresa fora do âmbito	345	0.41
Outras ações	4	0.00

Tabela C.5: *Código das Unidades Federativas e respectivas Regiões*

Região		UF	
1	N	11	RO
		12	AC
		13	AM
		14	RR
		15	PA
		16	AP
		17	TO
2	NE	21	MA
		22	PI
		23	CE
		24	RN
		25	PB
		26	PE
		27	AL
		28	SE
3	SE	31	MG
		32	ES
		33	RJ
		35	SP
4	S	41	PR
		42	SC
		43	RS
5	CO	50	MS
		51	MT
		52	GO
		53	DF

Tabela C.6: *Agrupamentos de classes da CNAE 2.0 nos Estados de Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina e Rio Grande do Sul trabalhados a três dígitos*

CNAE 2.0	Descrição
46.1	Representantes comerciais e agentes do comércio, exceto de veículos automotores e motocicletas
46.2	Comércio atacadista de matérias-primas agrícolas e animais vivos
4662-1, 4664-8 e 4669-9	Comércio atacadista de máquinas, aparelhos e equipamentos para usos técnico e profissional e outros usos não especificados anteriormente; partes e peças
47.2	Comércio varejista de produtos alimentícios, bebidas e fumo
47.3	Comércio varejista de combustíveis para veículos automotores

Referências Bibliográficas

- Albieri(1989)** Sônia Albieri. A ausência de resposta em pesquisas: uma aplicação de métodos de imputação. Dissertação de Mestrado, Instituto de Matemática Pura e Aplicada, Brasil. Citado na pág. 4, 23, 26, 29
- BRASIL(2003)** BRASIL. Decreto nº 4.740, de 13 de junho de 2003, aprova o estatuto e o quadro demonstrativo dos cargos em comissão e funções gratificadas da fundação instituto brasileiro de geografia e estatística - ibge, e dá outras providências @Misc. <https://ww2.ibge.gov.br/home/disseminacao/eventos/missao/estatuto.shtm>, 2003. Último acesso em 06/03/2017. Citado na pág. 5
- Census Bureau(2018)** U.S. Census Bureau. *Reengineering the Census Bureau's Annual Economic Surveys*. The National Academies Press. National Academies of Sciences, Engineering and Medicine. Citado na pág. 38
- Góes(1996)** Magdalena C. Góes. A modernização das estatísticas econômicas. Em *IV Encontro Nacional de Produtores e Usuários de Informações Sociais, Econômicas e Territoriais*, Rio de Janeiro, RJ. 9 v. em 24. Citado na pág. 10
- IBGE(2017a)** IBGE. Missão, principais funções, histórico e estrutura do ibge @Misc. <https://ww2.ibge.gov.br/home/disseminacao/eventos/missao/default.shtm>, 2017a. Último acesso em 06/03/2017. Citado na pág. 5
- IBGE(2019)** IBGE. Princípios fundamentais das estatísticas oficiais do ibge @Misc. https://ww2.ibge.gov.br/home/disseminacao/eventos/missao/principios_fundamentais_estatisticas.shtm, 2019. Último acesso em 16/02/2019. Citado na pág. 3
- IBGE(2007)** Comissão Nacional de Classificação IBGE. *Classificação Nacional de Atividades Econômicas, versão 2.0*. CDDI/IBGE, 1st edição. Citado na pág. 7
- IBGE(2018)** Coordenação de Serviços e Comércio IBGE. *Pesquisa Anual de Comércio*. Série Relatórios Metodológicos - Volume 28. CDDI/IBGE, 1st edição. Citado na pág. 19
- IBGE(2000)** Departamento de Comércio e Serviço IBGE. *Pesquisa Anual de Comércio*. Série Relatórios Metodológicos - Volume 12. CDDI/IBGE, 2nd edição. Citado na pág. 7, 9, 17
- IBGE(2017b)** Departamento de Comércio e Serviço IBGE. *Pesquisa Anual de Comércio*. Volume 27. CDDI/IBGE. Citado na pág. 11, 14
- Kalton e Kasprzyk(1982)** Grahah Kalton e Daniel Kasprzyk. Imputing for missing survey responses. Em *Proceedings of the Section on Survey Research Methods*, Washington, DC. American Statistical Association. Citado na pág. 23, 26

- Little e Rubin(2002)** Roderick J. A. Little e Donald B. Rubin. *Statistical Analysis with Missing Data*. John Wiley Sons, Inc, 2nd edição. Citado na pág. [25](#), [32](#)
- Lohr(2010)** Sharon L. Lohr. *Sampling: Design and Analysis. Advanced Series*. Brooks/Cole, Cengage Learning, 2nd edição. Citado na pág. [3](#), [23](#), [24](#), [25](#), [26](#)
- Office for National Statistics(2018)** U.K. Office for National Statistics. *Annual Business Survey technical report*. Office for National Statistics. Citado na pág. [38](#)
- Pessoa(2005)** Djalma Pessoa. *Estimadores de Calibração em Pesquisas Amostrais*. CO-MEQ/IBGE. Citado na pág. [109](#)
- Silva(2017)** Pedro L. N. Silva. Crítica e imputação. Em *Pós-graduação em População, Território e Estatísticas Públicas*, Rio de Janeiro, RJ. Escola Nacional de Ciências Estatísticas. Notas de aula. Material em pdf. Citado na pág. [26](#)
- Snijkers et al.(2013)** Ger Snijkers, Gustav Haraldsen, Jacqui Jones e Diane Willimack. *Designing and Conducting Business Surveys*. Wiley Series in Survey Methodology. Wiley, 1st edição. Citado na pág. [18](#)
- Särndal e Lundström(2005)** Carl-Erik Särndal e Sixten Lundström. *Estimation in Surveys with Nonresponse*. John Wiley Sons Ltd, 1st edição. Citado na pág. [18](#), [26](#), [32](#)