

# NOTA Técnica

## Método das Peneiras Sucessivas na Extração de Dados da Nota Fiscal Eletrônica no Distrito Federal

Brasília-DF, maio de 2017

**GOVERNO DO DISTRITO FEDERAL**  
**Rodrigo Rollemberg**  
Governador

**Renato Santana**  
Vice-Governador

**SECRETARIA DE ESTADO DE PLANEJAMENTO, ORÇAMENTO  
E GESTÃO DO DISTRITO FEDERAL - SEPLAG**  
**Leany Barreiro de Sousa Lemos**  
Secretária

**COMPANHIA DE PLANEJAMENTO DO DISTRITO FEDERAL - CODEPLAN**  
**Lucio Remuzat Rennó Júnior**  
Presidente

**Martinho Bezerra de Paiva**  
Diretor Administrativo e Financeiro

**Ana Maria Nogales Vasconcelos**  
Diretora de Estudos e Pesquisas Socioeconômicas (respondendo)

**Ana Maria Nogales Vasconcelos**  
Diretora de Estudos e Políticas Sociais

**Aldo Paviani**  
Diretor de Estudos Urbanos e Ambientais

## **EQUIPE RESPONSÁVEL**

### **Companhia de Planejamento do Distrito Federal - Codeplan**

- **Larissa Maria Nocko**  
Chefe de Gabinete da DIPOS/Codeplan
- **Luiz Rubens Câmara de Araújo**  
Assistente da Gerência de Pesquisas Socioeconômicas - GEREPS/DIEPS/Codeplan

### **Secretaria de Fazenda do Distrito Federal - SEF/DF**

- **Evandro Manzano dos Santos**  
Gerente da Gerência de Gestão do Malha Fiscal
- **Gunther Siqueira Lemos Gomes**  
Chefe do Núcleo de Mineração de Dados

---

#### **Revisão e copidesque**

Eliane Menezes

#### **Editoração Eletrônica**

Maurício Suda

# SUMÁRIO

1. INTRODUÇÃO.....	5
2. AS NOTAS FISCAIS ELETRÔNICAS.....	6
3. CÓDIGOS FISCAIS DE OPERAÇÕES E PRESTAÇÕES (CFOP).....	8
3.1. Por que as operações devem ser zeradas ou negativadas?.....	8
4. SIGILO FISCAL: O MÉTODO DE PENEIRAS SUCESSIVAS .....	12
4.1. Estrutura e Composição da CNAE.....	12
4.2. Estrutura e Composição da NCM.....	12
4.3. Descrição do procedimento de agrupamento para manutenção do sigilo fiscal .....	13
4.4. Avaliação das informações geradas .....	17
4.4.1. Avaliação do volume de compras realizadas pelo Distrito Federal.....	18
4.4.2. Avaliação do volume de vendas realizadas pelo Distrito Federal .....	20
5. CONSIDERAÇÕES FINAIS .....	22

## 1. INTRODUÇÃO

O objetivo deste trabalho é apresentar uma proposta de método de tratamento de dados tributários sigilosos do Distrito Federal (DF). Pretende-se aqui apresentar informações sobre a NFe, bem como esclarecer os procedimentos metodológicos adotados para a adequação da relação de registros fiscais para uma base de dados de comércio interestadual.

Na intenção de mapear a interdependência comercial do Distrito Federal, foi estabelecida uma parceria entre a Secretaria de Fazenda do Distrito Federal (SEF/DF) e a Companhia de Planejamento do Distrito Federal (Codeplan), na qual a primeira encarregou-se de fornecer os dados de registros fiscais das Notas Fiscais Eletrônicas (NFe) e, a segunda, de produzir estudos que contribuam para a difusão do conhecimento relativo ao perfil das trocas e do comércio dentro do Distrito Federal e com as demais unidades federativas, além de suas fragilidades e potencialidades, que serão abordados em estudos subsequentes.

No entanto, a principal preocupação envolvendo a utilização de registros fiscais como subsídio à pesquisa é a garantia do sigilo fiscal. Essa questão foi contornada com base em um algoritmo desenvolvido pela SEF/DF especificamente para essa finalidade, no qual os dados foram agregados de forma a inibir a identificação dos contribuintes. Em outras palavras, a ferramenta possibilitou que os dados da NFe - praticamente a totalidade dos registros comerciais - fossem acessados para fins de pesquisa em política pública. Assim, essa nota técnica introduz os trabalhos desenvolvidos a partir desse banco de dados.

As categorias que, devido à preservação do sigilo fiscal, foram agregadas, apresentaram um elevado grau de comprometimento para análises mais aprofundadas das cadeias produtivas do Distrito Federal.

## 2. AS NOTAS FISCAIS ELETRÔNICAS

A Nota Fiscal Eletrônica (NFe)<sup>1</sup> é um documento de existência apenas digital, ou seja, é emitido e armazenado eletronicamente com a finalidade de documentar transações comerciais entre empresas. Ela tem a sua validade jurídica garantida pela assinatura digital do emitente e sua Autorização de Uso, fornecida pela Secretaria de Fazenda do Estado (Sistema Nota Fiscal Eletrônica, 2015). Trata-se de uma iniciativa das Secretarias de Fazenda estaduais e Receita Federal para reduzir custos, simplificar os processos e agilizar as transações comerciais entre as empresas, substituindo a Nota Fiscal em papel. Como sua implementação elimina o uso de papel, contribui para a redução dos custos com material, impressão, envio e armazenagem dos documentos fiscais.

Quanto ao fluxo operacional da NFe, inicialmente, o vendedor (ou emissor do documento fiscal) gera um arquivo digital com as informações da transação comercial e o transmite à Secretaria de Fazenda ao qual está sob jurisdição. A Secretaria recebe e valida o documento, tornando-o apto a amparar a operação para a qual foi emitida. Após a autorização ou denegação, a NFe é armazenada no site da Receita Federal e da respectiva receita estadual. Uma vez que o documento eletrônico é gerado, um documento auxiliar é impresso em papel apenas para acompanhar no trânsito da mercadoria. Então, o emitente deve fazer o arquivo eletrônico chegar ao recebedor da mercadoria ou serviço, que deve aceitar e fazer a devida contabilização. Após essa etapa, as fiscalizações tributárias, de posse das informações fiscais eletrônicas, armazenam e cruzam os dados para detectar indícios de sonegação e, conseqüentemente, iniciar procedimentos como conferência física de cargas e auditorias em empresas.

Entre as informações que constam na NFe, aquelas que foram consideradas interessantes ao presente estudo foram:

- **UF:** indica o código da Unidade da Federação brasileira, fornecido pelo Instituto Brasileiro de Geografia e Estatística (IBGE), que emitiu a nota. Esse emitente pode ser tanto o remetente quanto o destinatário da mercadoria, mas, cruzando com outros elementos da nota, pode-se determinar a UF de origem e a UF de destino da mercadoria que está sendo comercializada.
- **CNAE:** trata-se da Classificação Nacional de Atividades Econômicas, instrumento oficialmente adotado pelo Sistema Estatístico Nacional e pelos órgãos federais gestores de registros administrativos. Sua padronização é derivada da quarta versão da *International Standard Industrial Classification of All Economic Activities* (ISIC 4), classificação internacional de atividades. Essa informação consta de forma imprecisa na NFe, por esse motivo, optou-se por identificar os indivíduos na Relação Anual de Informações Sociais (RAIS 2015) e capturar a CNAE declarada por eles nessa segunda base, para então trabalhar as informações registradas na NFe com a CNAE corrigida pela RAIS.
- **NCM:** a Nomenclatura Comum do Mercosul é a classificação de mercadorias utilizada pelos países desse bloco, composta por oito dígitos e utilizada nos registros comerciais do Ministério da Indústria, Comércio Exterior e Serviços. Toma como base a classificação internacional do Sistema Harmonizado (SH),

---

<sup>1</sup> Esse trabalho utilizou a NFe modelo 55.

que corresponde aos seus primeiros seis dígitos, e os dois últimos representam informações específicas dos países que integram o bloco.

- **CFOP:** consiste no Código Fiscal de Operações e Prestações, que indica a natureza da circulação de mercadoria. É por meio do código fiscal que se define se a operação deverá recolher impostos. Composto por quatro dígitos, o primeiro indica se envolve uma operação de entrada ou saída de mercadoria do estabelecimento emitente da NFe; o segundo o caracteriza como uma operação intraestadual, interestadual ou internacional; e os dois últimos dizem respeito à especificidade da operação (se é uma venda, compra, devolução etc.).

A NFe é uma realidade na legislação desde 2005, e a partir daí diversos protocolos, convênios e ajustes foram estabelecidos, regulamentando a sua obrigatoriedade e adequando particularidades técnicas inerentes ao seu planejamento. Isso fez com que houvesse um aumento no número de registros nos anos iniciais, o que indicou a migração da documentação do formato impresso para o formato eletrônico, e não a intensificação das transações econômicas, como se poderia esperar ao observar a variação do volume de dados.

Por esse motivo, neste trabalho serão utilizados os dados a partir de 2010, ano que apresentou um volume expressivo de registros. É importante destacar, no entanto, que estes não compreendem a maioria das transações realizadas entre pessoas físicas, pois a NFe modelo 55 ampara as transações realizadas com mercadorias essencialmente entre pessoas jurídicas. Ou, em outras palavras, a NFe modelo 55 não se destina a substituir o modelo para pessoa física, tais como os cupons fiscais ou a Nota Fiscal do Consumidor Eletrônica (NFCe) localmente, entretanto, o movimento interestadual de mercadorias tem como único documento previsto a NFe. Assim, uma vez que o objetivo deste trabalho é analisar o comércio intrafirma, em um primeiro momento, optou-se por descartar todas as trocas realizadas para pessoa física. No entanto, o volume dessas transações é não negligenciável. Entre as vendas do Distrito Federal para outras unidades da Federação, por exemplo, 7% correspondem a vendas para pessoa física (categoria denominada **Z** nas próximas seções). Em um segundo momento, na intenção de não descartar essa parcela de trocas, mas ao mesmo tempo utilizar informações fidedignas, optou-se por desconsiderar as transações internas para pessoa física, mas mantê-las nas transações interestaduais.

### **3. CÓDIGOS FISCAIS DE OPERAÇÕES E PRESTAÇÕES (CFOP)**

O Código Fiscal de Operações e Prestações (CFOP) é um código do sistema tributário brasileiro que indica se a nota fiscal recolhe ou não impostos, além de definir a natureza da operação. Para o presente estudo, a importância do código está na sua descrição, que permite selecionar as operações que configuram transações comerciais e desconsiderar as operações que, por exemplo, são meramente ajustes contábeis e transações não econômicas, como a circulação de mercadorias sem contrapartida direta de pagamentos para venda ou compra.

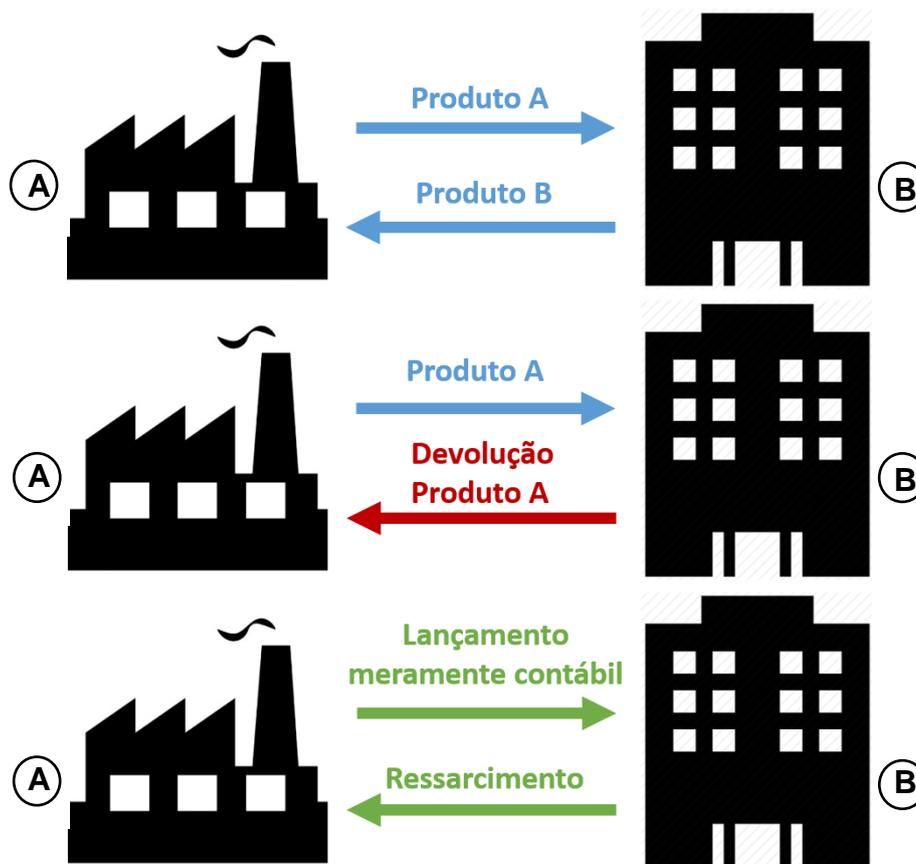
Nesse sentido, os códigos foram analisados individualmente a fim de se convencionar uma regra geral para quais operações seriam ou não levadas em consideração. A consolidação dessa regra estabeleceu três grupos: o primeiro consiste nas operações que configuram transações comerciais, de interesse dessa pesquisa; o segundo reúne os códigos que devem ser desconsiderados (zerados), como ressarcimentos e lançamentos meramente contábeis; e o terceiro trata daqueles códigos cujas operações devem ser negativadas para que a transação compense uma outra já realizada. Isso será detalhado a seguir.

#### **3.1. Por que as operações devem ser zeradas ou negativadas?**

Primeiramente, a intenção do uso dessa base de dados é traçar o perfil do comércio do Distrito Federal, portanto não faria sentido considerar operações que configuram meros registros contábeis, sem refletirem a transação de mercadorias. Por esse motivo, essas operações que não refletem a atividade comercial são anuladas. São exemplos disso operações cujo CFOP descreve um ressarcimento de Imposto Sobre Circulação de Mercadorias e Serviços (ICMS) ou operações a título de simples faturamento. Além disso, por vezes, há um registro quando a mercadoria é adquirida (compra) e um novo registro quando ela é transportada (remessa). Considerar essas duas operações significaria incorrer em dupla contagem, portanto, na presença dessas duas, optou-se por manter apenas aquelas atividades que indicam a realização da compra/venda, desconsiderando as que indicam somente a sua remessa.

De forma similar, quando uma operação interestadual de entrada é registrada no DF, por exemplo, nem sempre se trata da aquisição de uma mercadoria. Essa entrada pode indicar a devolução da mercadoria que foi adquirida anteriormente pela outra UF, o que não pode ser equivalente a uma compra interestadual realizada pelo DF, mas sim à anulação de uma operação anterior de venda. Ou, ainda, pode indicar a anulação de uma nota emitida com algum erro de preenchimento. Assim, foi convencionado aqui que CFOP que indicam operações de anulação, devolução ou retorno de mercadorias têm seus valores negativados (classificados como -1), o que corresponde a anular a operação original. A Figura 1 representa esses códigos.

Figura 1 - Operações fiscais que são utilizadas, negativas ou anuladas



A Figura 1 mostra as três situações geradas a partir do CFOP. Na primeira situação, a indústria “A” vende o produto “A” para a empresa “B” e compra o produto “B” dela. Essa operação fiscal representa a transação comercial que se pretende investigar, portanto, atribui-se à classificação 1. A segunda situação exemplifica a indústria “A” vendendo o seu produto “A” para a empresa “B” e, devido a algum problema com a mercadoria, ele é devolvido à indústria “A”. É a mesma situação que ocorre quando uma nota é emitida registrando a venda do produto “A”, mas, por exemplo, ao invés da nota ser preenchida com o valor de R\$ 100.000,00, um erro de digitação faz com que seja preenchida como R\$ 1.000.000,00. Nesse caso, a nota de anulação é emitida com o mesmo valor da nota original e no sentido contrário. Assim, se não fosse feita a negativação de uma das operações, as trocas entre a indústria “A” e a empresa “B” seriam erroneamente superestimadas. Por essa razão, essa transação é negativada, compensando a operação original. E a terceira situação mostra transações meramente contábeis entre a indústria “A” e a empresa “B”, que não refletem o comércio que se pretende mensurar, sendo, portanto, desconsideradas (anuladas). Por exemplo, no caso de simples faturamento, tem-se uma operação meramente contábil, de garantia prévia do negócio. Supondo que a empresa “A” faça um simples faturamento para a empresa “B” de um lote de centenas de toneladas de piso cerâmico. Em seguida, emita outras NFes com CFOP de “remessa por simples faturamento” em lotes menores, que se ajustem ao transporte e ao armazenamento pelo cliente “B”, mas correspondentes à primeira operação. Nesse sentido, considerar o simples faturamento seria duplicar a base de dados com transações em altos valores. Esta situação também é representada na Figura 2.

**Figura 2** - Operação fiscal de compra x operação de remessa: mesma transação econômica indicada em duas ou mais categorias de natureza da operação



Essas convenções foram formalizadas em uma sequência de 14 regras gerais com base em palavras-chave para definir os CFOP que seriam considerados na análise (multiplicados por 1), anulados (0) ou negativados (-1). Assim, as regras gerais adotadas estão apresentadas na Tabela 1.

**Tabela 1** - Sequência de regras gerais convencionadas para a análise comercial

1. Retornos, anulações e devoluções são negativos.	-1
2. Compra, entrada, venda, aquisição e transferência são positivos.	1
3. Ressarcimentos são nulos, pois sempre dizem respeito a ICMS.	0
4. Outras são positivos, pois sempre representam outra entrada ou outra saída.	1
5. Remessas são positivas.	1
6. Prestações são positivas.	1
7. Industrializações são positivas.	1
8. Lançamentos são nulos, pois correspondem a "lançamento a título de...", correspondendo a transferências meramente contábeis.	0
9. Recebimentos e utilizações são nulos.	0
10. Códigos cuja finalidade era exportação foram zerados.	0
11. Aqueles que representavam uma operação simbólica foram zerados.	0
12. Operações relativas a crédito/transferência/ressarcimento de ICMS foram anuladas.	0
13. Operações de compra e venda de consignação mercantil ou industrial foram anuladas, mantendo-se as entradas (1), remessas (1), retornos (-1) e devoluções (-1), na intenção de capturar o saldo.	1/0/-1
14. Operações de compra e venda para demonstração foram anuladas, mantendo-se as entradas (1), remessas (1), retornos (-1) e devoluções (-1).	1/0/-1

Com base na Tabela 1, os valores das operações são multiplicados por 1, 0 ou -1, o que convencionou a denominação *Valor Bruto* aos valores sem o tratamento dado ao CFOP, *Valor Retificado* ao somatório de valores que foram negativados e, por sua vez, *Valor Líquido* à diferença entre *Valor Bruto* e *Valor Retificado*. A validação dos dados com relação

a essas denominações – por parceiros, atividades e produtos - será abordada nas próximas seções.

Os critérios por trás das regras gerais foram formalizados com base na expertise dos colaboradores da Secretaria de Fazenda do Distrito Federal e seu conhecimento adquirido quanto ao preenchimento dos CFOP. A sua construção está sujeita à discricionariedade dos auditores em debate com os pesquisadores da Codeplan, o que compõe uma classificação sujeita a alterações.

## 4. SIGILO FISCAL: O MÉTODO DE PENEIRAS SUCESSIVAS

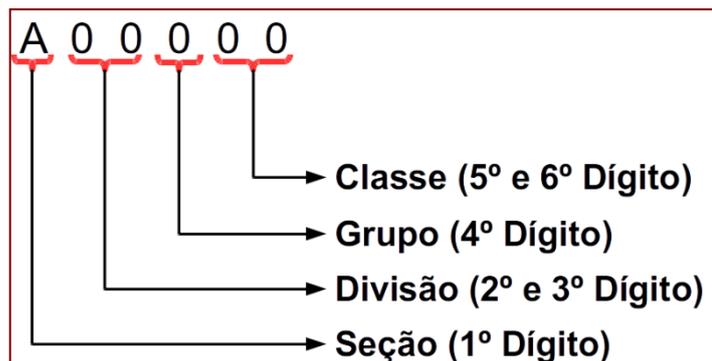
O sigilo fiscal tem como objetivo a preservação dos dados que os contribuintes entregam à autoridade tributária, uma vez que estes contêm informações econômicas, financeiras, de negócios e atividades relativas às suas vidas privadas.

Visando à disponibilização das informações contidas na NFe para este trabalho sem violar as regras de sigilo fiscal, os técnicos da SEF/DF desenvolveram uma metodologia de extração de dados, que consiste em agrupar as informações das NFe analisadas, considerando três campos: CNAE compra, CNAE venda e NCM.

### 4.1. Estrutura e Composição da CNAE

A CNAE é composta por uma estrutura de oito dígitos, mas, para a realização deste trabalho, foram considerados apenas os seis primeiros (da esquerda para a direita). Os dígitos selecionados apresentam a estrutura descrita na Figura 3.

Figura 3 - Descrição da estrutura do código da CNAE

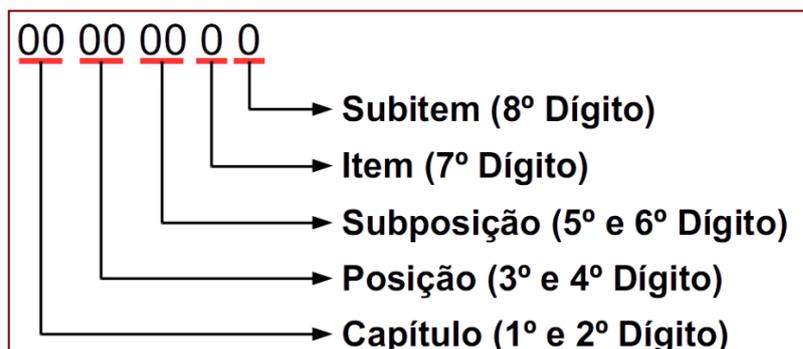


Fonte: <http://cnae.ibge.gov.br/?view=estrutura>

### 4.2. Estrutura e Composição da NCM

Assim como a CNAE, sua estrutura também é composta por uma numeração de oito dígitos. Os dígitos selecionados apresentam a estrutura descrita a seguir.

Figura 4 - Descrição da estrutura do código da NCM



Fonte: <http://www.mdic.gov.br/balanca/metodologia/NCM.txt>

### **4.3. Descrição do procedimento de agrupamento para manutenção do sigilo fiscal**

O procedimento de agrupamento consiste em agregações sucessivas do Cadastro Nacional da Pessoa Jurídica (CNPJ), levando em consideração as estruturas da CNAE e da NCM. O objetivo é manter, ao menos, cinco<sup>2</sup> CNPJ vinculados às estruturas das CNAE e NCM, desconsiderando repetições. Caso a quantidade mínima não seja atendida, os respectivos códigos das estruturas são substituídos pela letra “X”, de modo a se manter o sigilo fiscal.

O trabalho básico consiste em contar o número de CNPJ dentro de um agrupamento; caso a quantidade seja maior ou igual a cinco, a técnica considera este dado compartilhável; caso seja menor do que cinco, a técnica inclui um ou dois x na posição à direita. Inicia-se uma nova contagem no grau hierárquico superior e assim por diante até o final, onde resta apenas a classe CNAE, que por não ter grau superior assume a letra X.

Isto foi feito primeiro para cada uma das três variáveis em separado, e depois para o conjunto de todas para que em uma análise cruzada não fosse possível identificar o remetente ou o destinatário da mercadoria.

Durante a manipulação dos dados, foram encontrados casos de CNAE não identificadas. Isto ocorreu por dois motivos: o primeiro deveu-se a alguns CNPJ existentes nas NFe não terem sido identificados na Relação Anual de Informações Sociais (RAIS)<sup>3</sup>, utilizada para obtenção das CNAE. Nestes casos, o código da CNAE recebeu a letra “Y” na primeira posição, e os demais dígitos foram preenchidos com a letra “X”; o segundo motivo deveu-se a algumas NFe terem sido emitidas por pessoas físicas. Nestes casos, a CNAE recebeu a letra “Z” na primeira posição, e os demais dígitos foram preenchidos também com a letra “X”.

Os fluxogramas das figuras a seguir mostram as quatro etapas de forma esquemática.

---

<sup>2</sup> Os primeiros processamentos realizados consideraram três ou mais CNPJ. No entanto, como o risco de quebra de sigilo mostrou-se alto, optou-se por considerar cinco CNPJ.

<sup>3</sup> A RAIS é um instrumento de coleta de dados criado pelo Governo Brasileiro em 1975 com o objetivo de controlar as atividades trabalhistas das empresas brasileiras situadas no País. Ela é uma das principais fontes de informações sobre o mercado de trabalho formal brasileiro. Considerada como um censo pela sua abrangência que é de âmbito nacional e feita anualmente.

Figura 5 - Fluxograma das etapas 1 e 2, que considera a CNAE de origem e destino

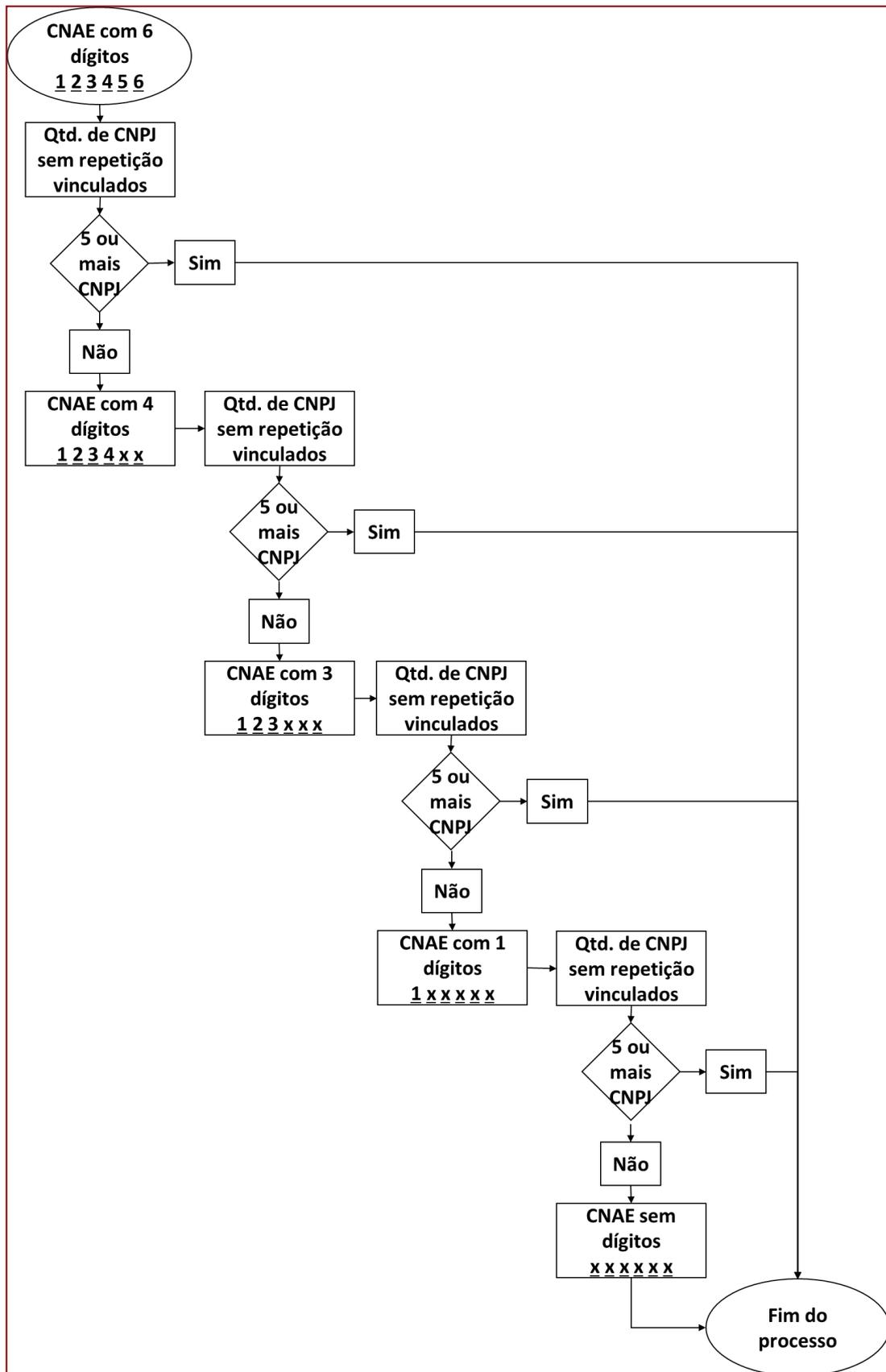


Figura 6 - Fluxograma da etapa 3 que considera a NCM

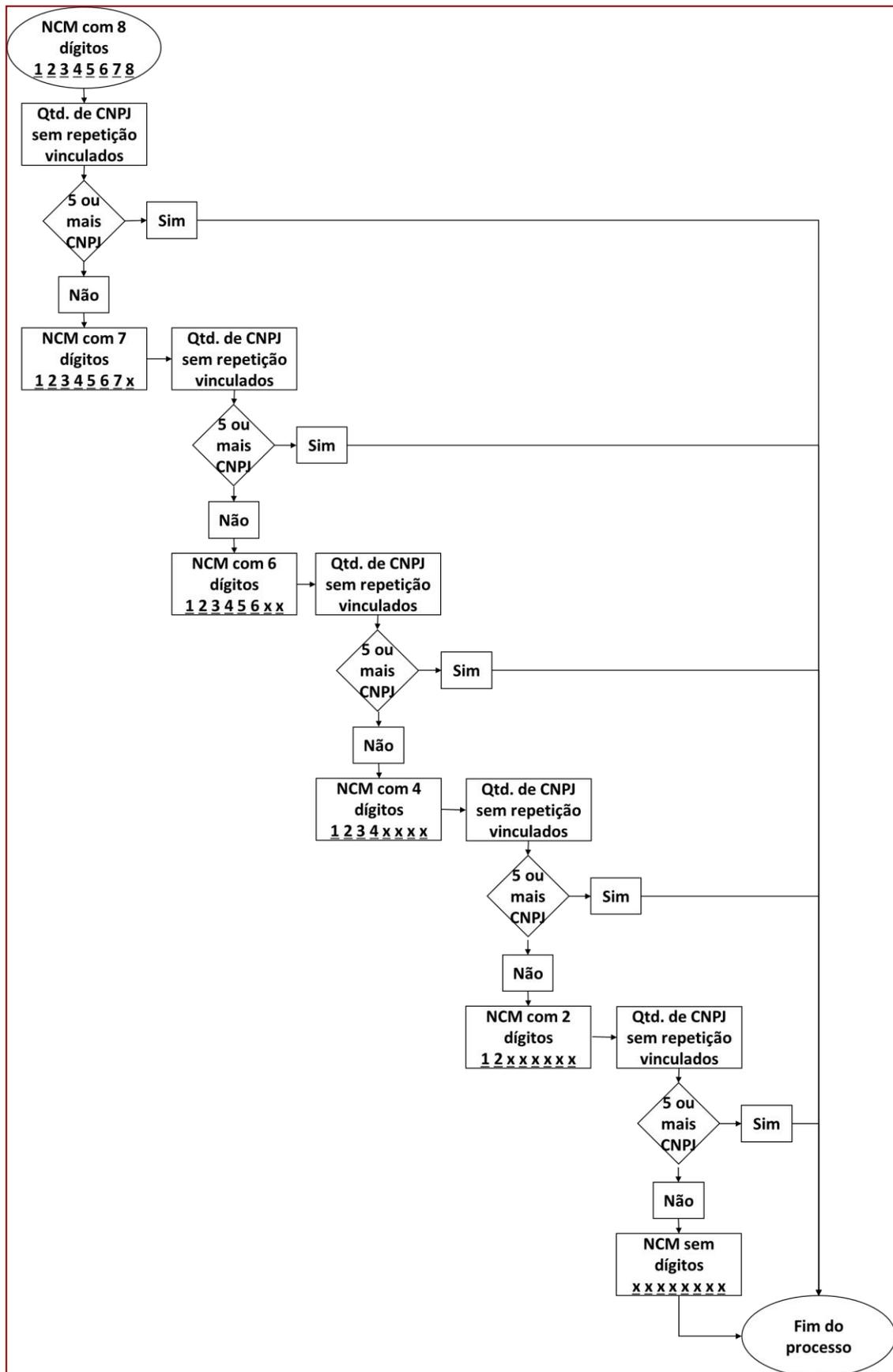
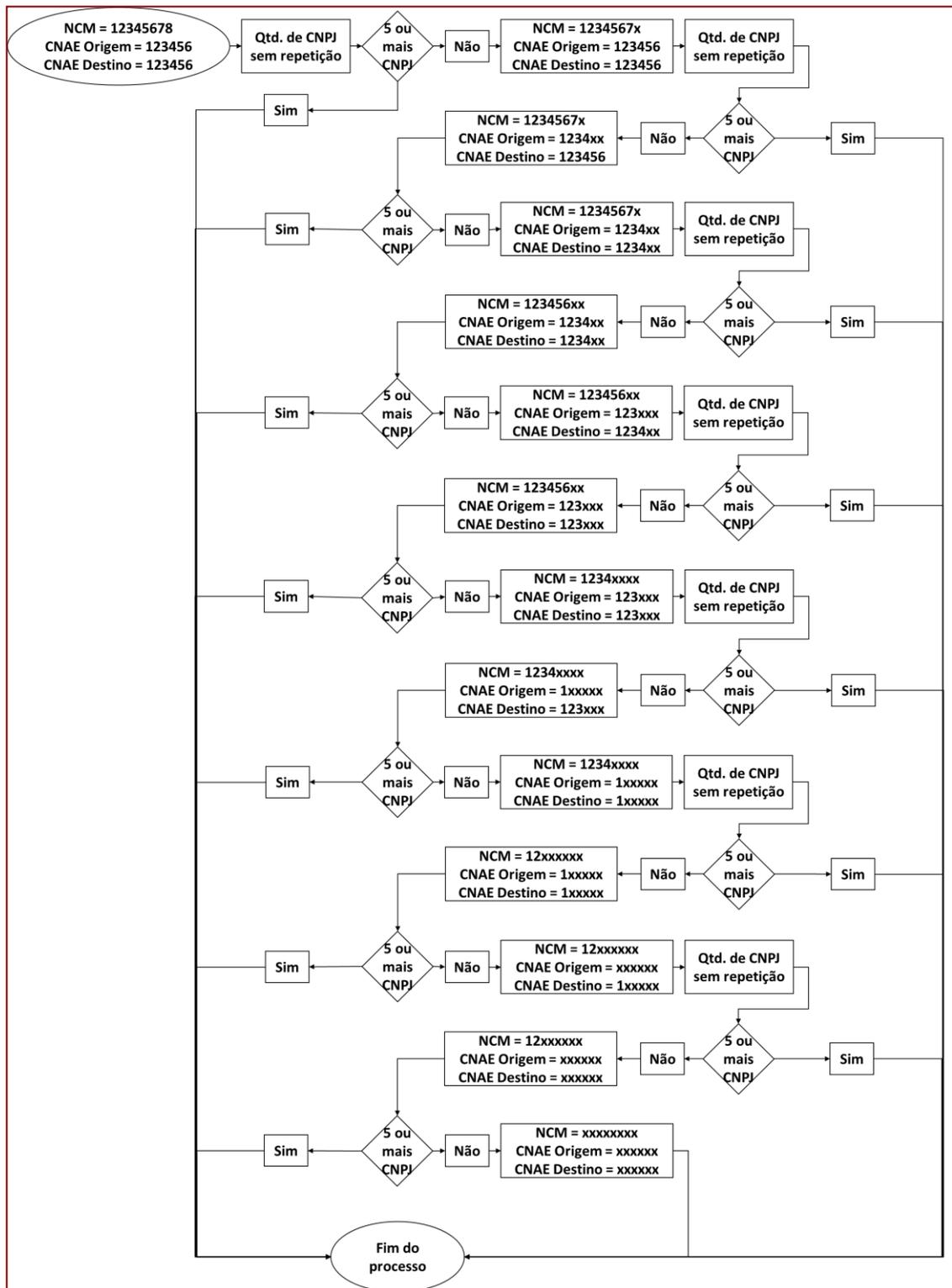


Figura 7 - Fluxograma da quarta etapa



#### 4.4. Avaliação das informações geradas

Para análise do impacto do processo de ocultação das informações, foram verificados os volumes de compra e venda do Distrito Federal, de acordo com as estruturas da CNAE e da NCM.

Para identificação do volume de compras feitas para o Distrito Federal (neste caso, o Distrito Federal foi considerado comprador), foram selecionadas todas as NFe com marcação de UF de compra igual a “Distrito Federal” e analisadas todas as informações da CNAE de venda e suas respectivas NCM.

Os registros em que a UF de venda também foi Distrito Federal e as NFe foram emitidas por pessoa física foram excluídos da análise.

Similarmente, para identificação do volume de vendas realizadas pelo Distrito Federal (neste caso o Distrito Federal, foi considerado como vendedor), foram selecionadas todas as NFe com marcação de UF de venda igual a “Distrito Federal” e analisadas todas as informações da CNAE de compra e suas respectivas NCM.

Os registros em que a UF de compra também foi Distrito Federal e as NFe foram emitidas por pessoa física também foram excluídos da análise.

As avaliações da CNAE (tanto de compra quanto de venda) e de suas respectivas NCM foram feitas a partir da atribuição de marcações, segundo regras descritas no quadro a seguir:

**Quadro 1** - Descrição das regras utilizadas para avaliação das informações geradas a partir da NFe

<b>Campo analisado</b>	<b>Posição analisada</b>	<b>Tipo de preenchimento</b>	<b>Marcação atribuída</b>
<b>CNAE</b>	1 <sup>a</sup>	Letra Z	CPF
		Letra Y	CNAE não identificada
		Letra X	X a partir da Seção
	2 <sup>a</sup> e 3 <sup>a</sup>	Letra X	X a partir da Divisão
	4 <sup>a</sup>		X a partir da Grupo
	5 <sup>a</sup>		X a partir da Classe
	Demais casos		CNAE Completa
<b>NCM</b>	1 <sup>a</sup> e 2 <sup>a</sup>	Letra X	X a partir do Capítulo
	3 <sup>a</sup> e 4 <sup>a</sup>	Letra X	X a partir do Posição
	5 <sup>a</sup> e 6 <sup>a</sup>		X a partir do Subposição
	7 <sup>a</sup>		X a partir do Item
	8 <sup>a</sup>		X a partir do Subitem
	Demais casos		NCM Completa

As informações do volume de compra e venda foram agregadas pelas marcações atribuídas e avaliadas em termos percentuais. As categorias, tanto da CNAE quanto da NCM, que apresentaram altos percentuais não podem ser avaliadas, uma vez que sua identificação fica impossibilitada devido à atribuição da letra “X”. Já as categorias que

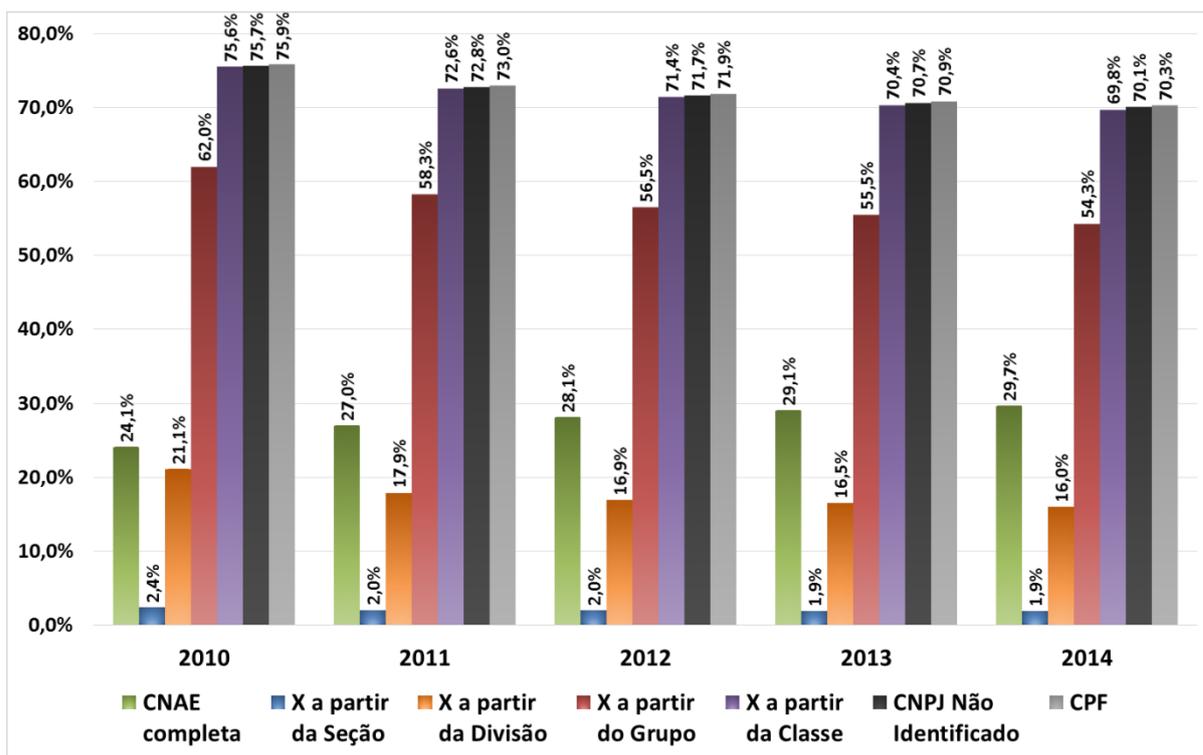
apresentaram baixos percentuais podem ser avaliadas, uma vez que a maior parte de sua identificação não fica impossibilitada devido à atribuição da letra “X”.

#### 4.4.1. Avaliação do volume de compras realizadas pelo Distrito Federal (o Distrito Federal é o comprador)

Analisando as informações relativas às compras feitas pelo Distrito Federal, é possível observar que, em média, pouco mais de 25% do volume negociado nos anos de 2010 e 2011 encontra-se vinculado à CNAE completa. Para os anos de 2012 a 2014, esta proporção chegou a quase 30% dos valores negociados. Quanto às CNAE que identificam a seção, o percentual de valores negociados vinculados à letra “X” oscilaram em torno de 2% ao longo dos anos analisados, o que mostra uma perda mínima nas informações geradas. Quanto às demais estruturas da CNAE (Divisão, Grupo e Classe), o percentual acumulado de valores negociados vinculados à letra “X” foi muito alto, tornando inviáveis análises relacionadas a essas estruturas.

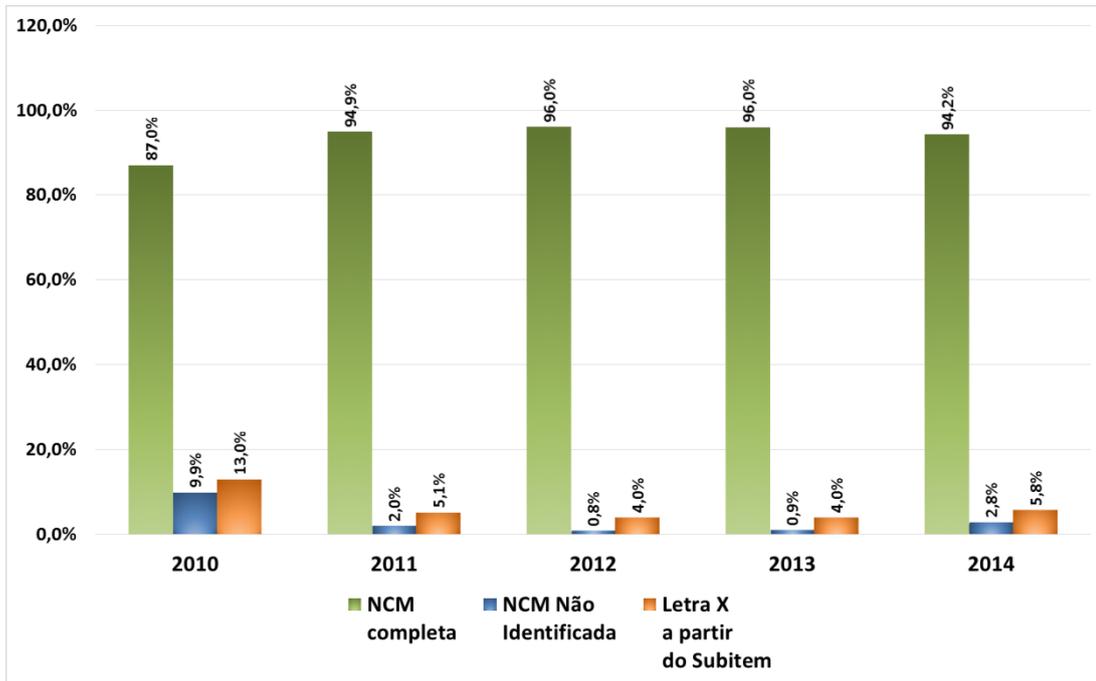
Cabe destacar o baixo percentual de valores vinculados ao CNPJ não identificado e notas emitidas por pessoas físicas, identificadas por meio do Cadastro de Pessoa Física (CPF) (ambos oscilaram em torno dos 0,2%). O Gráfico 1 mostra os percentuais descritos.

**Gráfico 1** - Avaliação do volume negociado, nas compras feitas para o Distrito Federal, segundo estrutura da CNAE



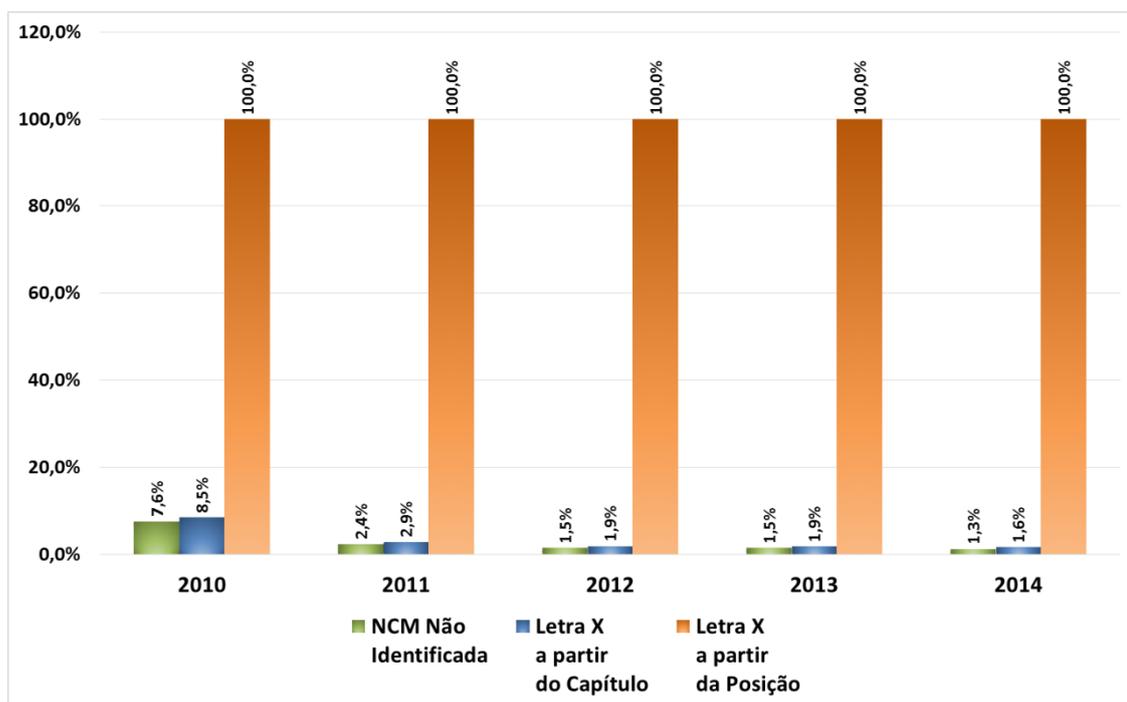
Analisando as NCM relacionadas às CNAE completas, observa-se que o percentual de volume negociado vinculado à NCM completa oscilou em torno dos 90% ao longo dos anos, o que possibilita uma análise mais detalhada. Já os percentuais acumulados de volumes negociados vinculados a NCM não identificadas e demais estruturas analisadas oscilaram entre 2% e 5% ao longo dos anos, com exceção do ano de 2010, quando esse percentual oscilou entre 9,9% e 13%. O Gráfico 2 mostra os percentuais descritos.

**Gráfico 2** - Avaliação do volume negociado nas compras feitas, para o Distrito Federal, segundo as estruturas da NCM relacionadas com a CNAE completa



Finalmente, analisando as NCM relacionadas às Seções da CNAE, observa-se que os percentuais de volume negociado, vinculados a NCM não identificadas e com a letra "X" a partir do capítulo foram baixos, oscilando em torno dos 2% ao longo dos anos, com exceção do ano de 2010, quando esse percentual ultrapassou os 7,6%. Níveis de detalhamento maiores que a Posição, ou seja, informações adicionais aos quatro primeiros dígitos do código NCM, comprometem a análise devido ao volume de dados denominados com a letra "X". O Gráfico 3 mostra os percentuais descritos.

**Gráfico 3** - Avaliação do volume negociado nas compras feitas, para o Distrito Federal, segundo as estruturas da NCM relacionadas com a estrutura de Seção CNAE

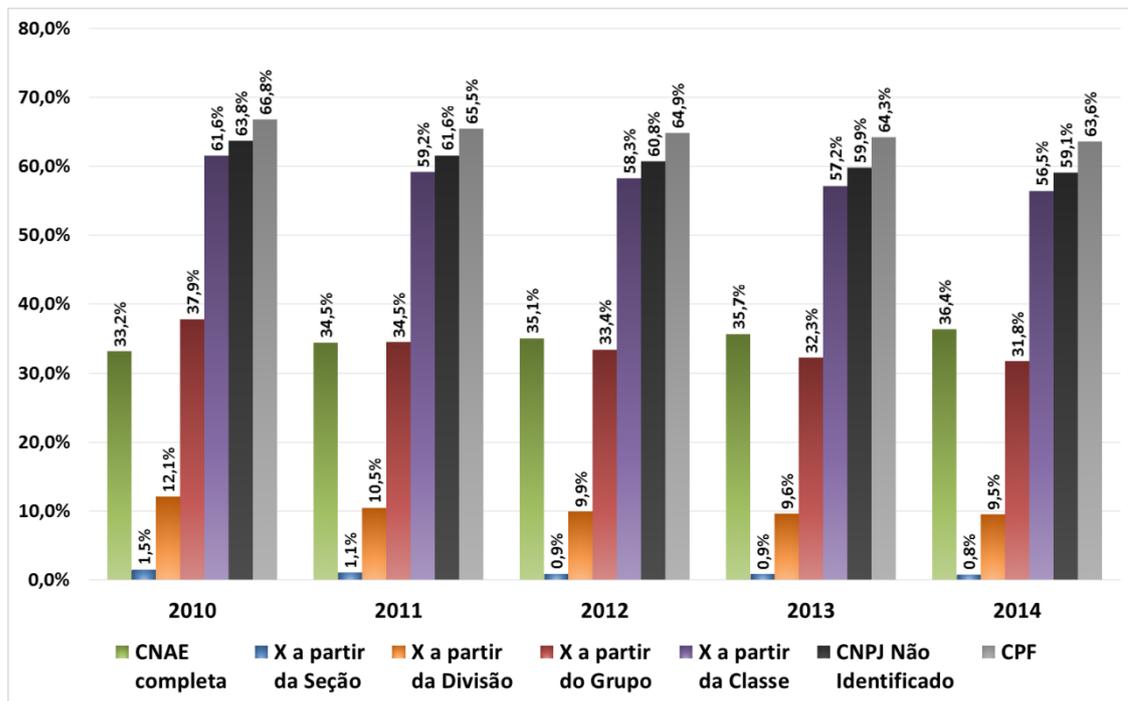


#### 4.4.2. Avaliação do volume de vendas realizadas pelo Distrito Federal (o Distrito Federal é o vendedor)

Analisando as informações relativas às vendas feitas pelo Distrito Federal, é possível observar que, em média, pouco mais de 35% do volume negociado nos anos de 2010 e 2014 encontra-se vinculado à CNAE completa. Para as CNAE que identificam a Seção, o percentual de valores negociados vinculados à letra “X” oscila em torno de 1% ao longo dos anos analisados, o que mostra uma perda mínima nas informações geradas, e para as demais estruturas da CNAE (Divisão, Grupo e Classe), o percentual acumulado de valores negociados vinculados à letra “X” foi muito alto, tornando inviáveis análises relacionadas a essas estruturas.

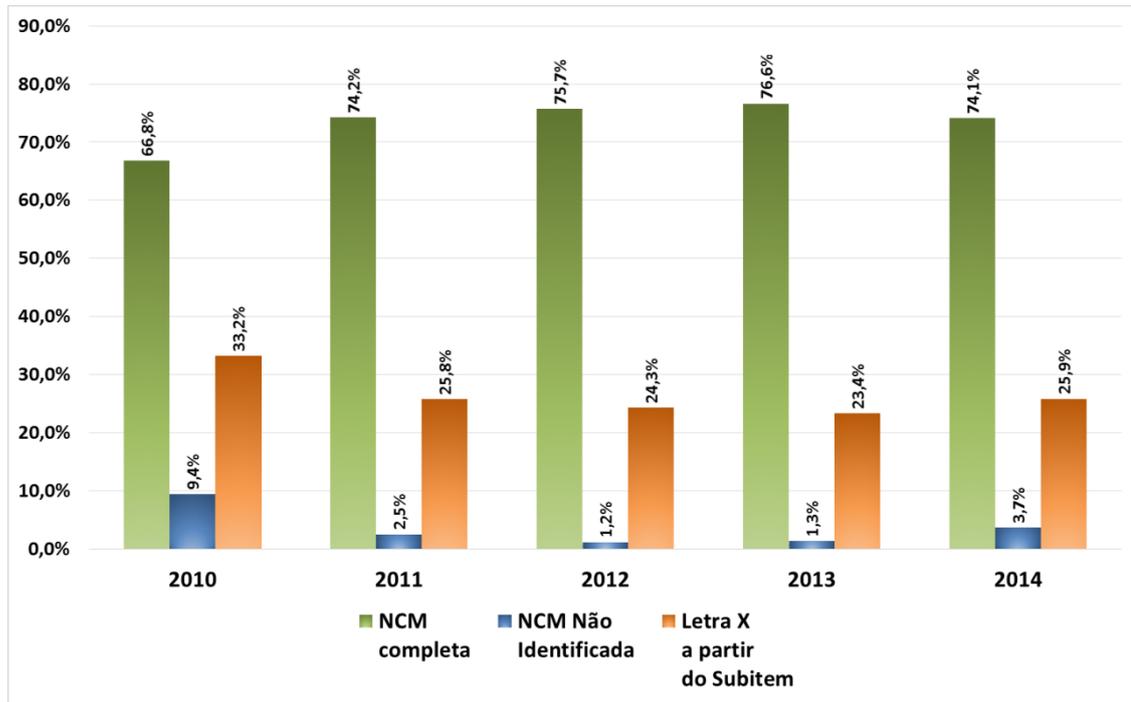
Cabe destacar o baixo percentual de valores vinculados a CNPJ não identificado, que oscilou em torno dos 2%, e notas emitidas por pessoas físicas identificadas por meio do CPF que oscilaram em torno dos 4% ao longo dos anos analisados. O Gráfico 4 mostra os percentuais descritos.

**Gráfico 4** - Avaliação do volume negociado nas vendas, feitas pelo Distrito Federal, segundo estrutura da CNAE



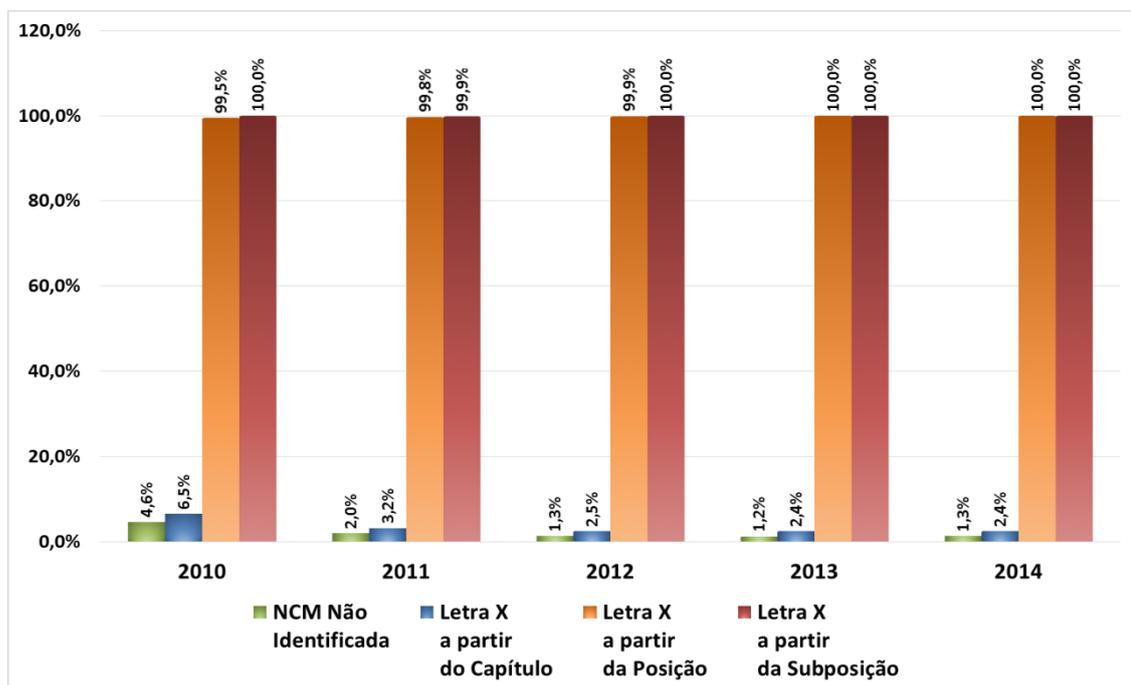
Analisando as NCM relacionadas às CNAE completas, observa-se que o percentual de volume negociado vinculado à NCM completa oscilou em torno dos 75% ao longo dos anos, o que possibilita uma análise mais detalhada. Já os percentuais de volumes negociados vinculados à NCM não identificada oscilou em torno dos 2% ao longo dos anos, com exceção do ano de 2010, no qual esse percentual chegou a quase 10%, e a única estrutura da NCM que apresentou volume negociado significativo vinculado à letra “X” foi o Subitem. O percentual acumulado apresentado oscilou em torno dos 25% ao longo dos anos analisados, comprometendo uma possível análise dessa estrutura. O Gráfico 5 mostra os percentuais descritos.

**Gráfico 5** - Avaliação do volume negociado nas vendas, feitas pelo Distrito Federal, segundo as estruturas da NCM relacionadas com a CNAE completa



Por fim, analisando as NCM relacionadas às Seções da CNAE, observa-se que os percentuais de volume negociado, vinculados à NCM não identificada ou vinculada à letra “X” a partir do capítulo foram baixos, oscilando entre 1,5% e 2,5% ao longo dos anos, com exceção do ano de 2010, quando esse percentual variou entre 4,6% e 6,5%. As análises relacionadas à Posição e Subposição da NCM não podem ser feitas, pois o percentual de perda ultrapassa os 99%. O Gráfico 6 mostra os percentuais descritos.

**Gráfico 6** - Avaliação do volume negociado nas vendas, feitas pelo Distrito Federal, segundo as estruturas da NCM relacionadas com a estrutura de Seção CNAE



## 5. CONSIDERAÇÕES FINAIS

Esta nota técnica teve como objetivo discutir uma proposta de tratamento de dados tributários sigilosos a partir do método das peneiras sucessivas. Deve-se ressaltar o avanço analítico permitido por esse método, uma vez que, na sua ausência, os resultados obtidos até aqui seriam inviáveis. A partir do algoritmo proposto, foi possível que pesquisadores externos à Secretaria de Fazenda do Distrito Federal (SEF/DF) conseguissem tecer um conjunto de conclusões, ainda que sem acesso direto aos microdados. No entanto, dada a riqueza de informações e possibilidades de subsídios para políticas públicas oriundas desses dados, torna-se fundamental a utilização de uma sala de sigilo, na qual pesquisadores possam, juntamente com a equipe da SEF/DF, trabalhar as informações de maneira mais desagregada, retornando apenas tabelas agregadas, os outputs.

Portanto, atenta-se para a necessidade de um maior detalhamento das informações concedidas que não fira o sigilo fiscal mas, ao mesmo tempo, permita o aprofundamento do estudo das relações de troca, garantindo a precisão da informação aos formuladores de política pública. Obviamente que, nessa proposta, a divulgação de resultados obtidos preservaria o sigilo dos dados. É importante recordar que o interesse dos pesquisadores, especialmente em se tratando de trabalho com dados quantitativos, está circunscrito à identificação de grandes tendências e perspectivas que permitam generalizações. Via de regra, nessa circunstância, as saídas de dados são matrizes agregadas ou operações sobre essas matrizes (por exemplo tabelas de regressão), o que elimina a possibilidade de se realizar engenharia reversa para quebra de sigilo.

**Companhia de Planejamento  
do Distrito Federal - Codeplan**

Setor de Administração Municipal  
SAM, Bloco H, Setores Complementares  
Ed. Sede Codeplan  
CEP: 70620-080 - Brasília-DF  
Fone: (0xx61) 3342-2222  
[www.codeplan.df.gov.br](http://www.codeplan.df.gov.br)  
[codeplan@codeplan.df.gov.br](mailto:codeplan@codeplan.df.gov.br)