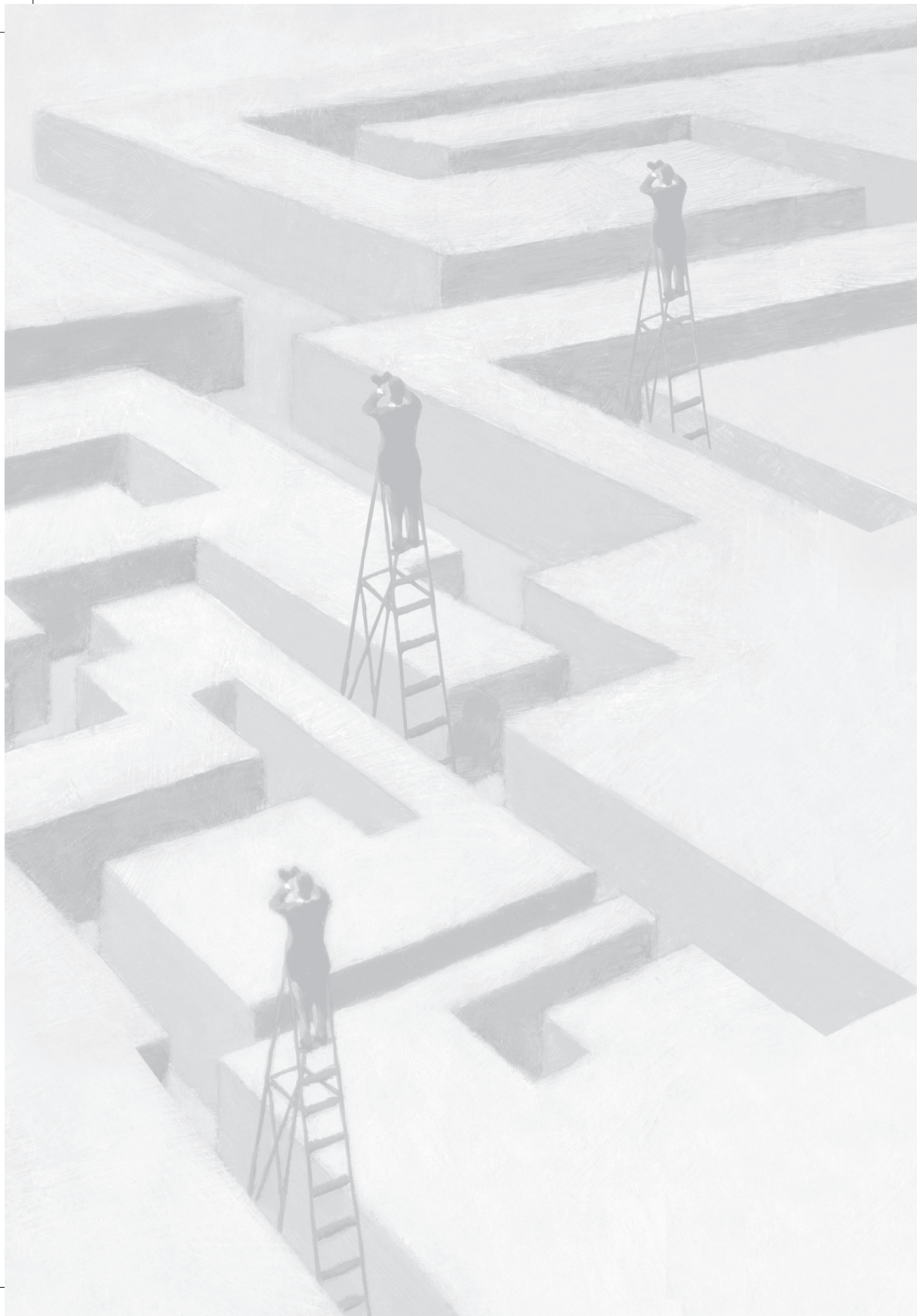


# **Método de delineamento de grupos de risco com representação dos resultados em cubo de critérios**

3º Lugar

**HENRIQUE DE CARVALHO RIGITANO\***

\* Graduado em Engenharia Civil pela Universidade de São Paulo (USP).  
Auditor-Fiscal da Receita Federal do Brasil, Delegacia da Receita Federal do  
Brasil em Porto Velho (DRF/PVO).



## **Método de delineamento de grupos de risco com representação dos resultados em cubo de critérios**

### **Resumo**

#### **A) Objetivos básicos**

O foco deste trabalho é a busca pela inovação e melhoria dos processos de trabalho da RFB. Será apresentado método inovador cujo objetivo é tornar mais justa e rápida a análise de processos. Trata-se de procedimento que permite encontrar processos em estoque com alta probabilidade de deferimento (grupos seguros) e processos com baixa probabilidade de deferimento (grupos de risco). Com essa informação, passa a ser possível realizar uma priorização inteligente dos processos.

Em se tratando de método estatístico que usa conceitos sofisticados, é proposta a visualização dos resultados em um cubo de critérios que permite a compreensão das principais ideias envolvidas sem a necessidade de conhecimento mais aprofundado em estatística.

O método foi construído pela aplicação de conceitos estatísticos na realidade da Receita Federal do Brasil (RFB). Depois de concebido, o método foi aplicado com sucesso em um dos estoques da unidade.

## **B) Metodologia utilizada**

As etapas e o desenvolvimento são sumariamente descritos, de maneira a situar o leitor sobre como realizar a implantação. Em primeiro lugar, é explicado como construir um critério de risco de deferimento que seja sólido e confiável. A técnica é encontrar pedidos similares já analisados no passado e verificar qual foi o percentual de deferimento para eles. Posteriormente, são usados conceitos estatísticos para extrapolar o comportamento no passado para os pedidos não analisados no estoque.

Em seguida, é realizada proposta de integração do critério de risco com dois outros: o de valor e o de tipo do pedido. Essa integração é visualizada por meio do cubo de critérios.

Por fim, é feita breve exposição sobre a solução das dificuldades que podem ser encontradas.

## **C) Adequação do trabalho aos critérios de julgamento**

### ***I – Criatividade e inovação***

O método é critério novo para seleção e agrupamento de processos. Ele foi desenvolvido dentro do contexto dos estoques de processos da RFB, mas pode ser utilizado em qualquer tipo de estoque. Basta haver dados históricos de deferimento.

As ideias podem ser aplicadas para melhoria de processos da RFB de forma imediata. O método foi desenvolvido para ser facilmente replicável.

### ***II – Relação custos versus benefícios***

O método pode ser utilizado em âmbito nacional. A implantação pode ser feita com facilidade, pois o custo é nulo e o número de funcionários necessários é mínimo.

Os benefícios são importantes, porque passa a haver alocação eficiente de servidores e recursos. Isso é realizado com a concentração de esforços na análise de processos com maior risco. Os resultados

podem ser mensurados pelo grau de diminuição do estoque e pelo percentual de deferimento após aplicação do método.

Ganhos não mensuráveis seriam a maior justiça na análise processual e no aumento na presença fiscal, porque os processos de alto risco serão examinados com critérios mais profundos, enquanto que o bom contribuinte não sofre fiscalizações morosas e desnecessárias. Outro ponto positivo é contribuição para evolução na cultura gerencial do órgão que passa a utilizar métodos inteligentes de análise de processos.

### ***III – Aumento de produtividade***

Como os esforços fiscalizatórios concentram-se nos processos de alto risco, os de risco pequeno podem ser tratados por ordens de serviço de rito sumário. Em alguns casos, podem, inclusive, ser tratados automaticamente. Esse tipo de ação tem forte aumento no índice de produtividade.

O sucesso do método foi comprovado com dados. Após sua aplicação, o tratamento diferenciado causou redução de 26% no estoque selecionado (de 124.913 para 91.669) em três meses.

### ***IV – Viabilidade de implementação***

A viabilidade foi comprovada pela aplicação na nossa unidade com bons resultados. Não há grandes dificuldades técnicas na implementação, já que os sistemas necessários para obtenção dos dados (ContÁgil, DW, SCC etc.) já existem e funcionam bem. Além desses sistemas, só é requerido o uso de programa de processamento de planilhas. Quanto à viabilidade política, ela fica bastante facilitada porque o método é transparente, isso permite que o gestor tenha confiança e segurança em utilizá-lo.

O público-alvo atingido é tanto o contribuinte quanto o servidor. O contribuinte recebe tratamento mais justo e os servidores que analisam o assunto podem ter modalidades de trabalho diferentes em função da probabilidade de deferimento do processo.

### ***V – Melhoria da qualidade dos serviços prestados e dos resultados estratégicos***

Analisando o comportamento da análise dos processos em função de características da solicitação e do requerente, nos aproximamos do objetivo “conhecer o perfil integral do contribuinte”. Com esse conhecimento, podemos prestar serviços de maneira melhor em vários aspectos, tanto no tempo quanto na equidade, o que tem reflexo em aprimorar e ampliar os serviços prestados à sociedade e elevar a percepção de risco e a presença fiscal. Internamente, estamos utilizando melhor os servidores e recursos disponíveis, o que influencia no objetivo de assegurar recursos e otimizar sua aplicação.

O maior impacto na cadeia de valor está na gestão do crédito tributário/gerir o crédito tributário do contribuinte, já os estoques de pedidos diversos relacionados ao crédito tributário podem ser geridos de maneira otimizada. Isso porque passa a ser possível a realização de priorização em função do risco.

#### **D) Principais lições aprendidas**

Pudemos verificar que novas ideias e ferramentas podem ser utilizadas no âmbito da RFB com grandes benefícios. Foi possível compreender o comportamento do contribuinte para, a partir dessa compreensão, traçar estratégias eficientes. Estudo técnico aprofundado e testes de cada hipótese foram fundamentais ao sucesso do método.

Todas as dificuldades enfrentadas durante a construção do método estão documentadas junto com suas soluções, o que poderá auxiliar quem eventualmente passe pelos mesmos problemas no futuro.

## Sumário

<b>1</b>	<b>Introdução .....</b>	<b>123</b>
<b>2</b>	<b>Justiça na priorização dos pedidos .....</b>	<b>124</b>
<b>3</b>	<b>Como encontrar a probabilidade de deferimento.....</b>	<b>126</b>
3.1	Passo 1: Escolher o estoque onde o método será aplicado.....	126
3.2	Passo 2: Levantar as possíveis características formadoras de grupos	127
3.3	Passo 3: Obter as informações sobre os processos .....	131
3.3.1	<i>Como gerar relatório de características no ContÁgil.....</i>	<i>132</i>
3.3.2	<i>Como gerar relatório de características no SCC.....</i>	<i>133</i>
3.3.3	<i>Como gerar relatório de características no e-processo.....</i>	<i>135</i>
3.3.4	<i>Como gerar relatório de características no DW .....</i>	<i>136</i>
3.4	Passo 4: Analisar o histórico de deferimento e indeferimento.....	137
3.5	Passo 5: Calcular a margem de erro do percentual de deferimento...	138
3.6	Passo 6: Verificar se os grupos podem ser destacados da população .....	139
3.7	Passo 7: Associar o percentual de deferimento aos processos do estoque .....	143
<b>4</b>	<b>Apresentação visual dos resultados .....</b>	<b>143</b>
4.1	Critério de risco mostrado isoladamente.....	143
4.2	Critério de risco associado com outros critérios – o cubo de critérios .....	144
<b>5</b>	<b>Dificuldades enfrentadas e suas soluções .....</b>	<b>146</b>
5.1	Limites entre grupos .....	146
5.2	Poucos indivíduos no grupo.....	148
5.3	Relevância do percentual de deferimento.....	149
5.4	Diferença entre correlação e causa .....	150
<b>6</b>	<b>Conclusão.....</b>	<b>151</b>
	<b>Referências .....</b>	<b>153</b>





## **Método de delineamento de grupos de risco com representação dos resultados em cubo de critérios**

---

### **1 Introdução**

É problema recorrente, na Receita Federal do Brasil (RFB), a grande entrada de processos sem capacidade de saída equivalente, por falta de servidores.

No momento em que se nota aumento incessante no estoque de processos sem perspectiva de diminuição, torna-se claro que temos situação insustentável e é necessário buscar novas saídas.

A solução para um problema complexo como esse exige certo grau de inovação. De acordo com Bessante Tidd (2009), os grandes saltos de desenvolvimento vieram de soluções que procuraram sair dos métodos tradicionais tratando problemas com uma visão diferente. De acordo com os autores, a inovação é a chave para o desenvolvimento e deve ser aplicada, não só em busca de lucro no setor privado, mas também no setor público, visando a benefícios para a sociedade.

Obviamente a solução inovadora não pode ser realizada sem o devido cuidado. De acordo com Vargas (2005), um dos maiores motivos de falhas em projetos é trabalhar com base somente no *feeling*, sem utilizar dados e, se possível, análises estatísticas.

A inovação que este trabalho apresenta como possível solução para diminuição de estoques de processos é um método de seleção de processos que utiliza conceitos básicos de *datamining* para encontrar, dentro do estoque, quais pedidos têm alta probabilidade de deferimento e quais têm baixa probabilidade de deferimento. Ao final da aplicação desse método, são obtidas listas de processos que provavelmente serão deferidos e listas de processos que provavelmente serão indeferidos.

Como o objetivo é encontrar no estoque processos com percentual de deferimento alto ou baixo, surge o desafio de descobrir a probabilidade de deferimento de processos que ainda não foram analisados.

Seria possível simplesmente utilizar argumentação razoável para propor critérios de que se imagine ter correlação com o deferimento, mas, no presente trabalho, a proposta é agir de maneira diferente.

A proposta é olhar para o passado e ver qual foi o percentual de deferimento de processos com características similares. Esse percentual é associado aos processos não analisados utilizando técnicas estatísticas.

Com os dados de probabilidade de deferimento dos processos em estoque, torna-se possível analisar com procedimentos diferenciados os processos com alto risco de indeferimento. Para eles, cabe realizar análise aprofundada. Já os processos com baixo risco de indeferimento podem passar por análise sumária. Assim é possível alocar a força de trabalho com maior eficiência.

Foi tomado o devido cuidado para que esse método fosse, além de conceitualmente correto, possível de ser compreendido sem a necessidade de formação específica em estatística.

O método foi aplicado com sucesso dentro do estoque de declarações em malha da unidade.

## **2 Justiça na priorização dos pedidos**

Sandel (2011) demonstra que há vários possíveis critérios de justiça. Esse pensamento pode ser aplicado no momento de elaborar um critério de tratamento justo de processos em estoque. É importante termos a visão das múltiplas possibilidades.

A regra mais comum é a priorização em função da data do protocolo.

Outra forma, amplamente utilizada na RFB, é a de diferenciar o estoque para idosos e portadores de moléstia grave. Esses casos esperam menos tempo em fila devido à sua condição.

Diferente critério é o de complexidade da solicitação. Seria injusto ter que esperar tempo igualmente longo para obter serviços simples e serviços complexos. É mais justo que cada serviço tenha sua própria fila. Esse tipo de diferenciação ocorre, por exemplo, quando pedidos simples são atendidos imediatamente nos Centros de Atendimento ao Contribuinte e pedidos complexos são protocolados via processo administrativo para posterior análise.

Também é possível utilizar como critério o valor vinculado ao processo, porque o valor é uma forma de medir o impacto da análise para o governo e para a sociedade.

Outra forma de priorizar é pelo uso do histórico de relacionamento do solicitante, ou seja, pessoas que reiteradamente realizaram pedidos justos e sempre cumpriram suas obrigações corretamente podem ter tratamento diferente daquelas que realizam pedidos fraudulentos, protelatórios e desatenciosos.

Caso não exista um histórico para um solicitante específico, a solução é comparar suas características com a de casos semelhantes que já tiveram seus pedidos analisados. Com isso, é possível prever, em algum grau, seu comportamento. Nesse caso, será associada ao processo uma probabilidade de agir conforme a previsão.

Não é necessário escolher somente um dos critérios. É perfeitamente possível realizar uma combinação de dois ou mais deles.

Esses critérios podem ser utilizados para diferenciação na análise tanto pela elaboração de regra de priorização (qual processo trabalhar), como também pela realização de análise diferenciada (como trabalhar o processo).

### 3 Como encontrar a probabilidade de deferimento

Como o objetivo é descobrir a probabilidade de indeferimento para utilizar como critério justo de diferenciação na análise, passamos a descrever como se descobre o valor dessa probabilidade.

A probabilidade de indeferimento também pode ser chamada de “risco” do pedido. O valor do risco associado a processos em estoque é encontrado pela observação de análises já concluídas de pedidos similares, ou seja, pesquisaremos se um pedido com as mesmas características costumou ser deferido ou indeferido em análises anteriores.

A reunião dos processos de um grupo com deferimento alto será um grupo seguro. Simetricamente, uma reunião de processos que tiveram o percentual de indeferimento alto será um grupo de risco.

Cada processo em estoque é comparado com os processos nos grupos seguros e nos grupos de risco. Caso suas características coincidam com as de algum grupo, podemos enquadrá-lo como membro. Com isso, é transmitida para ele a probabilidade de deferimento do respectivo grupo.

Explicaremos melhor o procedimento. Os passos do método são os seguintes:

Passo 1: Escolher o estoque onde o método será aplicado

Passo 2: Levantar as possíveis características formadoras de grupos

Passo 3: Obter as informações sobre os processos

Passo 4: Analisar o histórico de deferimento e indeferimento

Passo 5: Calcular a margem de erro do percentual de deferimento

Passo 6: Verificar se os grupos podem ser destacados da população

Passo 7: Associar o percentual de deferimento ao estoque

#### 3.1 Passo 1: Escolher o estoque onde o método será aplicado

É requisito fundamental que o estoque seja grande, pois só com uma massa grande de dados é possível aplicar conceitos estatísticos.

Além disso, é necessário que existam dados sobre aquele assunto no sistema, principalmente a informação de resultados de análises realizadas no passado.

Na busca de um estoque para aplicar o método, podemos mostrar como exemplo o estoque da malha levantado na tabela 1. Entre os vários parâmetros, os melhores candidatos seriam o MF400 e MF540.

Tabela 1: Exemplo de quantidades em estoque para escolha do assunto a ser atacado

	2014	2013	2012	2011	Total
MF 100 - Omissão Rendimento PJ Titular	14855	10037	684	870	11591
MF 101 - Omissão Rendimento PJ Titular Ações Judiciais	76	99	459	290	924
MF 102 - Omissão Rendimento PJ Titular Aluguéis	487	394	318	257	1456
MF 103 - Omissão Rendimento PJ Dependente	455	628	311	251	1645
MF 110 - Ausência de DIRF	1379	3715	4323	595	10012
MF 120 - Divergência de IRRF	209	251	430	193	1083
MF 210 - Fonte Pagadora em DIRF x DARF	541	529	327	41	1438
MF 231 - Previdência Oficial Indevida PJ	211	231	335	117	894
MF 233 - Previdência Privada	186	255	191	0	632
MF 250 - Pensão Alimentícia acima do limite	477	431	414	154	1476
MF 300 - Omissão de Rendimento DIMOB Titular	461	429	317	230	1437
MF 350 - Carnê-Leão e/ou Imposto Complementar recolhidos a menor	220	258	416	411	1305
MF 400 - Despesas Médicas acima do limite	32941	17047	8216	698	58902
MF 500 - Declaração liberada por VEL do parâmetro MF110	3316	2154	5	11	5486
MF 501 - Declaração liberada por VEL do parâmetro MF111	57	71	0	0	128
MF 510 - Declaração liberada por VEL do parâmetro MF350	648	690	7	51	1396
MF 520 - Declaração liberada por VEL do parâmetro MF210	4723	2117	3	3	6846
MF 530 - Declaração liberada por VEL do parâmetro MF340	115	68	0	0	183
MF 531 - Declaração liberada por VEL do parâmetro MF341	18	13	0	0	31
MF 540 - Declaração liberada por VEL do parâmetro MF100 ou MF103	11194	3993	7	56	15250
MF 550 - Declaração liberada por VEL do parâmetro MF120 ou MF123	725	1463	7	25	2220
MF outros	188	171	134	85	578

124913

### 3.2 Passo 2: Levantar as possíveis características formadoras de grupos

Quando falamos de grupo, estamos querendo dizer uma reunião de processos que tenham uma ou mais características em comum. Se vários processos compartilharem uma característica, eles fazem parte de um mesmo grupo. Por exemplo: “grupo de todos os processos protocolados tempestivamente”, ou “grupo de todos os processos vinculados a contribuintes que nunca foram multados” etc.

É preciso deixar isso claro: é sempre alguma característica dos processos o critério usado para delinear um grupo. Portanto, quaisquer

características que os processos tenham podem ser usadas como critério para formação de grupo.

Teoricamente, poderíamos associar a cada processo todas as suas características armazenadas em algum sistema da RFB, como as da Declaração de Imposto de Renda da Pessoa Física (DIRPF), Declaração de Imposto de Renda Retido na Fonte (Dirf), Cadastro de Pessoa Física (CPF) etc. Em seguida, todas seriam utilizadas para gerar grupos. O problema é que demoraria muito tempo para testar tantos grupos.

Nossa proposta é realizar uma reunião com espírito de *brainstorming* para imaginar características que sejam boas para formar grupos seguros ou de risco. Nessa fase, é necessário imaginar livremente possíveis características que poderiam influenciar, de alguma maneira, no deferimento do pedido, nenhuma ideia pode ser reprimida ou considerada tola. De acordo com Karlgaard (2014), os melhores resultados surgem com reuniões realizadas com três a sete participantes de diferentes formações e personalidades. Eles devem ter confiança e respeito mútuo.

Seguem como exemplo algumas características que foram levantadas na unidade como possíveis formadores de grupos:

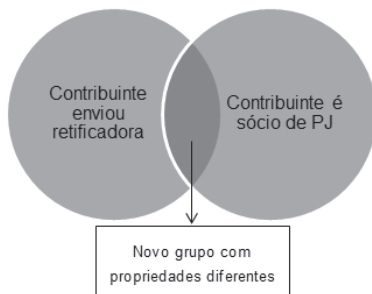
- 1) Data da transmissão
- 2) Período de apuração
- 3) Data do pagamento
- 4) Diferença entre o período de apuração e a data de transmissão
- 5) Diferença entre a data do pagamento e a data de transmissão
- 6) Diferença entre o período de apuração e a data do pagamento
- 7) Versão do programa utilizado para envio
- 8) Vinculação a processo judicial
- 9) Valores envolvidos
- 10) Situação cadastral do CPF
- 11) Sexo do contribuinte
- 12) Nacionalidade do contribuinte

- 13) Idade do contribuinte
- 14) Tipo do logradouro do contribuinte
- 15) Bairro do contribuinte
- 16) Diferença entre a data de nascimento e a data de expedição do CPF
- 17) Profissão do contribuinte
- 18) Se houve mudança recente de endereço
- 19) Se é portador de moléstia grave
- 20) Existência de declaração de despesa com instrução
- 21) Banco do contribuinte
- 22) Sistema operacional utilizado no envio
- 23) Uso de certificação digital
- 24) Envio de retificadora
- 25) Existência de ganho de capital da DIRPF
- 26) Diferença entre a data de envio e a data-limite
- 27) Valor do imposto a pagar na DIRPF
- 28) Valor do imposto a restituir na DIRPF
- 29) Existência de bloqueio de restituição da DIRPF
- 30) Estado do resgate de restituição da DIRPF
- 31) Tipo de intimação enviada em função de fiscalização na DIRPF
- 32) Tipo de resultado do trabalho realizado pela malha em função da resposta do contribuinte à intimação enviada
- 33) Valor dos bens do contribuinte
- 34) Quantidade de bens do contribuinte
- 35) Tipo de rendimento recebido pelo contribuinte
- 36) Valor do rendimento recebido pelo contribuinte
- 37) Quantidade de dependentes
- 38) Tipo de rendimento recebido pelo dependente
- 39) Valor do rendimento recebido pelo dependente

- 40) Existência de incompatibilidade entre o aumento no patrimônio e a renda
- 41) Enquadramento como contribuinte especial
- 42) Endereço da placa de rede do contribuinte
- 43) Existência de cônjuge dependente na DIRPF
- 44) Contribuinte enquadrado no Estatuto do Idoso
- 45) Contribuinte é sócio de pessoa jurídica (PJ)

Além de podermos formar grupos de processos que possuam *uma* dessas características em comum, também é possível criar grupos que tenham simultaneamente *duas ou mais* delas. Por exemplo: além de criarmos um grupo com a característica 24 (contribuinte enviou retificadora) e outro grupo separado usando a característica 47 (contribuinte é sócio de PJ), é possível gerar novo grupo concatenando as duas características. Nesse caso, o grupo seria formado por processos de contribuintes que enviaram retificadora e que, além disso, são sócios de PJ. O efeito é o de uma intersecção de dois grupos. O exemplo dado pode ser ilustrado como na figura 1.

Figura 1: Grupo novo formado pela combinação de dois outros



Quando concatenamos dois grupos, podem surgir diferenças importantes no percentual de deferimento. Podem surgir correlações que não existiam para cada característica vista isoladamente.

Os resultados mostraram que grupos formados a partir de duas ou três características concatenadas foram os que apresentaram percentual de deferimento com maior diferença em relação à média.



Depois do levantamento das características, será necessário observar o comportamento passado para verificar se elas realmente influenciam ou não no deferimento dos pedidos.

### 3.3 Passo 3: Obter as informações sobre os processos

Para que consigamos agrupar os processos com as características levantadas, é necessário associar essas características a cada processo analisado no passado. Então precisamos buscá-las em algum sistema da Receita Federal. Nesse ponto, o problema a ser enfrentado é a dificuldade de extrair dos bancos de dados as informações que desejamos. Essa dificuldade é abordada por Balaniuk (2010).

Nas nossas análises, obtivemos as informações no ContÁgil, gerenciais do SCC-Sief, gerenciais do *e-processo* e DW. O procedimento é extrair de cada um desses sistemas as informações desejadas. O resultado será uma tabela para cada sistema contendo informações sobre os processos já analisados. Com todas em mãos, será necessário realizar a tarefa de organização. Todas as informações deverão ficar em uma única tabela, como na tabela 2.

Tabela 2: Processos com suas características

Características retiradas dos diversos sistemas				
Processo	resultado da análise	Envio de DIRPF	CPF	etc
11111.111111/2015-01	deferido	enviou DIRPF	regular	
22222.222222/2015-02	indeferido	Omisso	regular	
33333.333333/2015-03	deferido	Omisso	irregular	
44444.444444/2015-04	deferido	enviou DIRPF	irregular	
55555.555555/2015-05	deferido	enviou DIRPF	irregular	
66666.666666/2015-06	indeferido	enviou DIRPF	regular	

Nessa tabela, as características de cada processo ficam nas mesmas linhas. Cada característica pode ter vindo de um sistema diferente, mas todas devem ser integradas nessa mesma planilha.

Segue breve tutorial de como gerar relatórios em cada um desses sistemas.

### 3.3.1 Como gerar relatório de características no ContÁgil

Diversas características do contribuinte podem ser extraídas por meio de funcionalidade do ContÁgil que consolida informações oriundas do Dossiê Integrado.

Primeiro é necessário importar os dados na tela da figura 2:

Caminho: Importar/importar dossiê integrado pessoa física.

Figura 2: Tela que importa dados no ContÁgil

Importar Dados do Dossiê Integrado: Bases da pessoa física

Período  
Ano: indefinido

Inclusão de contribuinte na lista

- 
- 
- 
- 
- 

Exclusão de contribuinte da lista

- 
- 

Lista de contribuintes

CPF	Nome

Seleção as bases a serem recebidas

Resumo

- DIRF Resumo
- Sinal resumo data de pagamento
- Sinal resumo data de apuracao

Bases

<input type="checkbox"/> Extrato da pessoa física	<input type="checkbox"/> Dimof	<input type="checkbox"/> Sinal/RDOC data pagamento
<input type="checkbox"/> Cadastro CPF	<input type="checkbox"/> DIRF (como beneficiário)	<input type="checkbox"/> Sinal/RDOC data de apuracao
<input type="checkbox"/> Ação fiscal	<input type="checkbox"/> DIRPF	<input type="checkbox"/> Sipade
<input type="checkbox"/> Cadin	<input type="checkbox"/> DOI	<input type="checkbox"/> Siscomex exportação
<input type="checkbox"/> CNPJ	<input type="checkbox"/> DPV	<input type="checkbox"/> Siscomex importação
<input type="checkbox"/> Compras DIPJ terceiros	<input type="checkbox"/> ITR	<input type="checkbox"/> Vendas DIPJ terceiros
<input type="checkbox"/> Comprota	<input type="checkbox"/> Perdiscomp	
<input type="checkbox"/> Conta corrente pessoa física	<input type="checkbox"/> Profisc	
<input type="checkbox"/> DBF	<input type="checkbox"/> Rendimentos DASN	
<input type="checkbox"/> DCPMF/DCPMF III	<input type="checkbox"/> Rendimentos DIPJ	
<input type="checkbox"/> Decred	<input type="checkbox"/> Rendimentos recebidos de pessoa física	
<input type="checkbox"/> Derc	<input type="checkbox"/> Siafi	
<input type="checkbox"/> Dimob	<input type="checkbox"/> Sida	

Ajuda

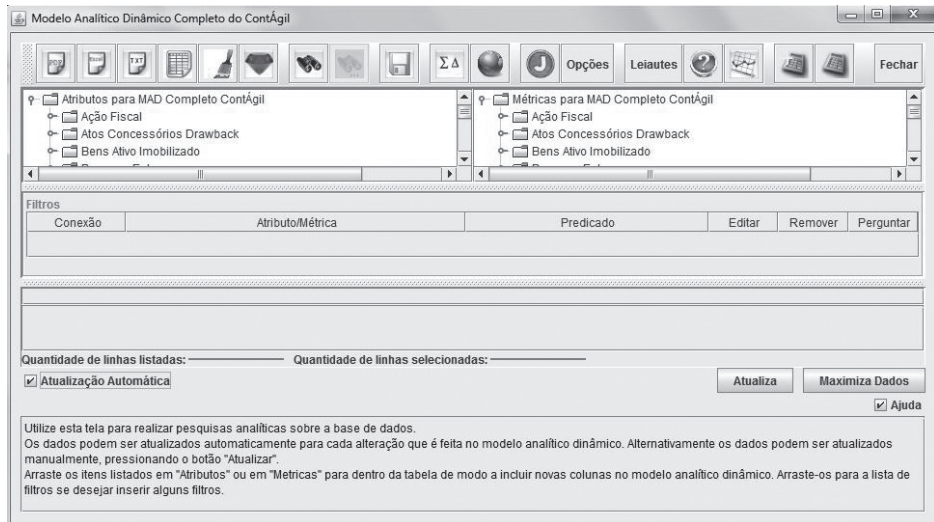
Procedimentos:

- 1) Incluir os números de CPF de interesse usando o botão “adicionar contribuintes de um arquivo txt”.
- 2) Selecionar os últimos cinco anos no botão “alterar ano”.
- 3) Selecionar com um clique os sistemas e as declarações onde estejam as características a testar.
- 4) Clicar no botão “Importar”.


Depois de importar os dados é que se gera o relatório na tela da figura 3.

Caminho: Geral/Modelo Analítico Dinâmico Completo/Bases do ContÁgil/Geral.

Figura 3: Tela de relatórios do ContÁgil



Procedimentos:

- 1) Arrastar atributos e métricas de interesse para o campo vazio cinza da janela.
- 2) Exportar resultados para planilha no botão .


### 3.3.2 Como gerar relatório de características no SCC

No caso da necessidade de obter características de uma Perdcomp, é possível obtê-las em relatório gerado na tela do Sief-SCC mostrada na figura 4.

Caminho: Sief / Perdcomp / Consulta.

Figura 4: Tela que gera relatório no SCC

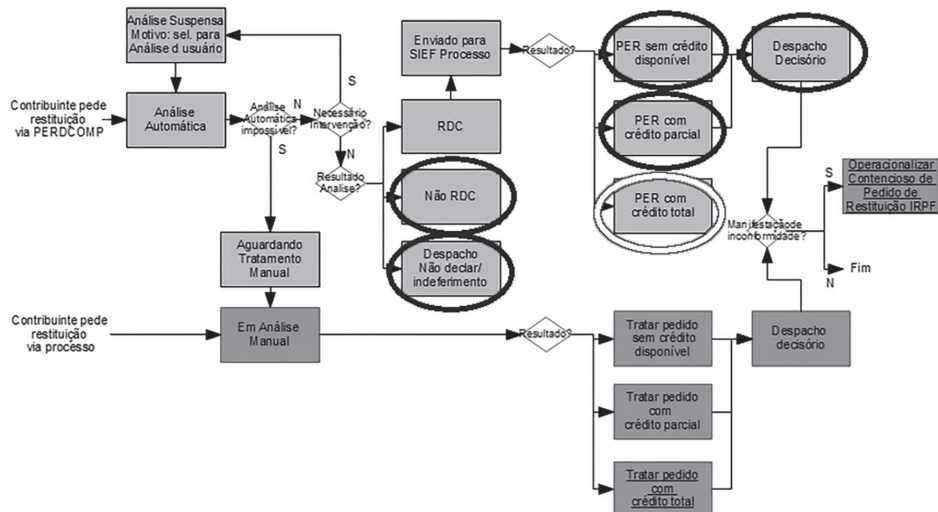
Procedimentos:

- 1) Selecionar a delegacia e a UA nos campos de mesmo nome.
- 2) Clicar no botão .

A lista gerada mostra todos os Perdcomp da delegacia com suas características.

Esse relatório não especifica claramente se o pedido foi deferido ou indeferido. Em vez disso, ele mostra qual é a fase atual do pedido. Portanto, precisamos saber quais fases significam que o Perdcomp já houve análise com deferimento e quais fases significam que já houve análise com indeferimento. A figura 5 mostra todas as fases do processamento dos Perdcomp, indicando com círculo preto as fases que significam indeferimento e com círculo branco as que significam deferimento.

Figura 5: Fluxo de análise de Perdcomp



Obs.: deferimentos parciais foram considerados indeferimentos.

### 3.3.3 Como gerar relatório de características no e-processo

No *e-processo* é possível obter informações vinculadas ao processo protocolado. A figura 6 mostra a tela onde se monta esse relatório.

Caminho: *e-processo*/gerencial/relatórios gerenciais/Visão=Estoque/Relatório=Relatório de estoque configurável/Selecionar Unidade e equipe.

Figura 6: Tela que gera relatório no e-processo

Editar Visão

---

Descrição: 31769984811 - TODOS

Hierarquia abaixo  (Só pode ser utilizado caso seja selecionada uma equipe)

Totalizador:  Quantidade de processos  
 Hora Estimada  
 Valor  
 Percentual

Filtro: Há Filtro Configurado

Palavras-Chave (Datas, Valores e Número do Processo apenas para detalhe)

<input type="checkbox"/>	ACT - Código
<input type="checkbox"/>	ACT - Código Completo
<input type="checkbox"/>	ACT - Código do Tema
<input type="checkbox"/>	ACT - Origem
<input type="checkbox"/>	ACT - Tributo
<input type="checkbox"/>	Alargações no Recurso
<input type="checkbox"/>	Ano Exercício Malha

Ordem agrupamento  
Nome Equipe Atual  
Nome Atividade Atual

Ordem Exibição  
Nome Tarifa Atual  
Número Processo  
Data Entrada Atividade  
NI Contribuinte  
Indicador se Existe Processo de Representação p  
Motivo da Prioridade  
Idade Contribuinte  
Indicador de Solicitação de Juntada Pendente

Procedimentos:

- 1) Dar nome ao relatório no campo “Descrição”.
- 2) Marcar campos “hierarquia abaixo” e totalizador por “quantidade de processos”.
- 3) Configurar filtro no botão “Editar Filtro”.
- 4) Escolher características de interesse no quadro “Palavras-Chave”. Marcar em azul permite expandir a característica como tópico no resultado do relatório, marcar em verde não permite.
- 5) Clicar no botão “Confirmar”.

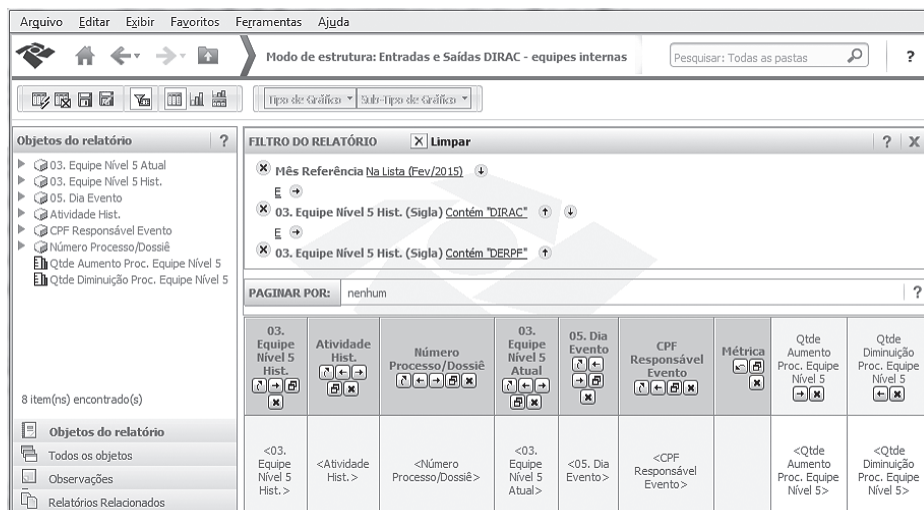
### 3.3.4 Como gerar relatório de características no DW

O DW é recurso poderoso onde se encontram inúmeras informações de vários tipos.


A figura 7 mostra a tela onde são construídos os relatórios.

Caminho: Base de publicação/ Escolher Dimensão / Relatório em branco.

Figura 7: Tela que gera relatório no DW



Procedimentos:

- 1) Clicar no botão “todos os objetos” e navegar na estrutura encontrando características de interesse.
- 2) Acrescentar as características arrastando cada uma até o campo de “filtros do relatório” ou para as colunas do relatório.
- 3) Gerar o relatório no botão .

### 3.4 Passo 4: Analisar o histórico de deferimento e indeferimento

Associadas as informações para cada processo analisado no passado, o que deve ser feito é: 1) criar os grupos de processos com uma, duas ou três características em comum; e 2) calcular a proporção de deferimento e indeferimento para cada grupo. Cada resultado deve ser listado em planilha que mostre o percentual de deferimento para cada grupo testado.

Utilizamos programação VBA em macro para realizar esse procedimento de forma rápida. O código foi construído para, inclusive, criar grupos com todas as possibilidades de combinação de uma, duas ou três características.

É perfeitamente possível pesquisar manualmente os grupos usando a funcionalidade “filtros” ou “tabela dinâmica” dos programas de processamento de planilha. A programação em VBA-Macro serve somente para acelerar a busca.

Neste estudo, pudemos verificar que o resultado dos testes é bastante surpreendente. Os melhores candidatos a apresentarem correlação com o alto percentual de deferimento não tinham correlação nenhuma. Ao mesmo tempo, alguns candidatos bastante peculiares apresentaram correlação forte com o deferimento.

A título de demonstração de ordens de grandeza, podemos dizer que, em uma das análises realizadas na nossa unidade, das 47 características listadas, apenas 12 sobreviveram ao teste. As outras não apresentaram correlação com deferimento nem com indeferimento. Esse resultado demonstra a importância dos testes. Hipóteses, por mais razoáveis que pareçam, não podem ser aplicadas sem a devida análise real de fatos e dados. Testar previamente os critérios de priorização de processo é um dos pontos forte desse método.

### **3.5 Passo 5: Calcular a margem de erro do percentual de deferimento**

Nas análises de percentual de deferimento, o que estamos observando é o que ocorreu no passado. Tentar extrapolar esses dados para o futuro envolve certo grau de incerteza.

Essa tarefa de adivinhação, aparentemente impossível, torna-se factível utilizando um simples pressuposto: de que as pessoas se comportarão no futuro próximo da mesma maneira que se comportaram no passado recente. Esse pressuposto é bastante razoável e raramente está errado.

Utilizando esse pressuposto, podemos usar uma *analogia* de que a nossa fotografia do passado funciona como uma amostra dos pedidos de restituição que receberemos ao longo de toda a linha do tempo.

Vendo a situação dessa maneira, podemos calcular a margem de erro associada a essa amostra. Nesse caso, a margem de erro serve para quantificar matematicamente o tamanho da incerteza associada ao uso de dados passados para prever o comportamento no futuro.



Trata-se de *critério razoável* para que se quantifique, de maneira mais realista, o tamanho da incerteza da extrapolação.

A fórmula para o cálculo da margem de erro da proporção é a seguinte:

$$Z_{(1-\alpha)/2} \cdot \sqrt{\frac{p \cdot q}{n}}$$

Na fórmula:

$Z_{(1-\alpha)/2}$  é uma constante oriunda da curva normal. Esta constante é escolhida para refletir o nível de confiança desejado. No nosso caso, foi escolhido o valor de 1,96, que corresponde a um nível de confiança de 95%.

$p$  é a proporção de deferimento, ou seja, a quantidade deferida dividida pela quantidade total de pedidos na amostra (a amostra são os pedidos feitos somente por certo grupo a ser estudado, como, por exemplo, os pedidos feitos por estudantes).

$q$  é a proporção de indeferimento, que é igual a  $(1-p)$ .

$n$  é a quantidade total de pedidos na amostra.

### **3.6 Passo 6: Verificar se os grupos podem ser destacados da população**

O critério que utilizamos para saber se um grupo pode ser destacado da população como um todo é calcular a margem de erro para cada uma dos grupos e, então, verificar se elas se sobrepõem à margem de erro da população. Caso não fiquem sobrepostas, podemos dizer que o grupo tem realmente comportamento diferente da população, ou seja, que o grupo tem realmente um percentual de deferimento diferente.

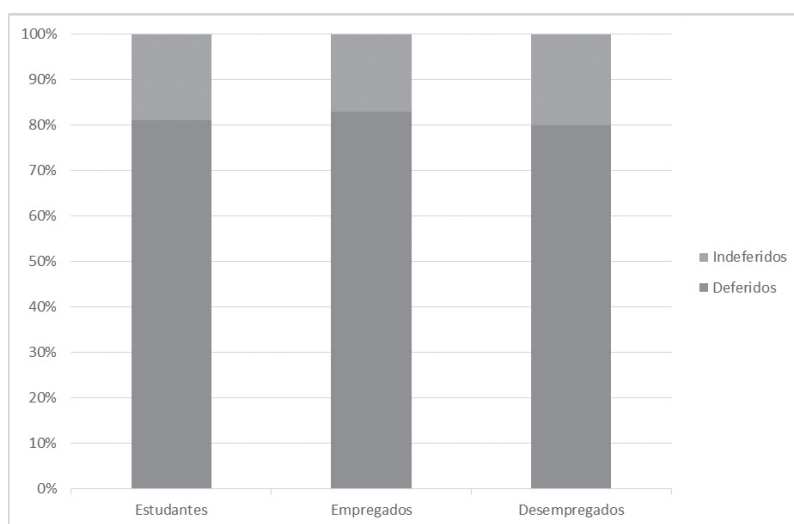
Trata-se de critério aproximado. Sabemos que o cálculo pode ser aperfeiçoado e que existem critérios ainda melhores da estatística para justificar o destacamento de um grupo. Optamos por utilizar este porque, além de razoável, pode ser utilizado por usuários apenas com conhecimento básico em estatística e planilhas. O uso de outros critérios

exigiria nível mais avançado em estatística. A ideia é abrir a possibilidade de compreensão e utilização segura para mais pessoas.

A maior simplicidade não prejudica seu aspecto objetivo. Independentemente da facilidade, o critério impede a existência de arbitrariedades na seleção.

Um exemplo hipotético: digamos que coloquemos, lado a lado, três grupos, cada um com percentual de deferimento idêntico, conforme mostrado na figura 8.

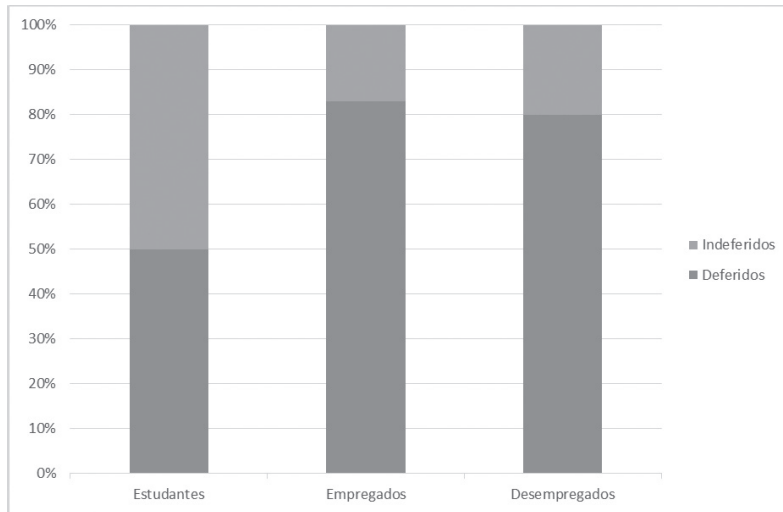
Figura 8: Grupos em teste hipotético (percentual idêntico)



Neste exemplo, seja qual for a ocupação do contribuinte, o percentual de deferimento é próximo a 80%. Portanto, não existe variação no percentual de deferimento em função da ocupação. Dessa forma, não podemos destacar nenhum grupo de risco com base nessa característica.

Situação diferente seria o caso da figura 9.

Figura 9: Grupos em teste hipotético (percentual diferente)

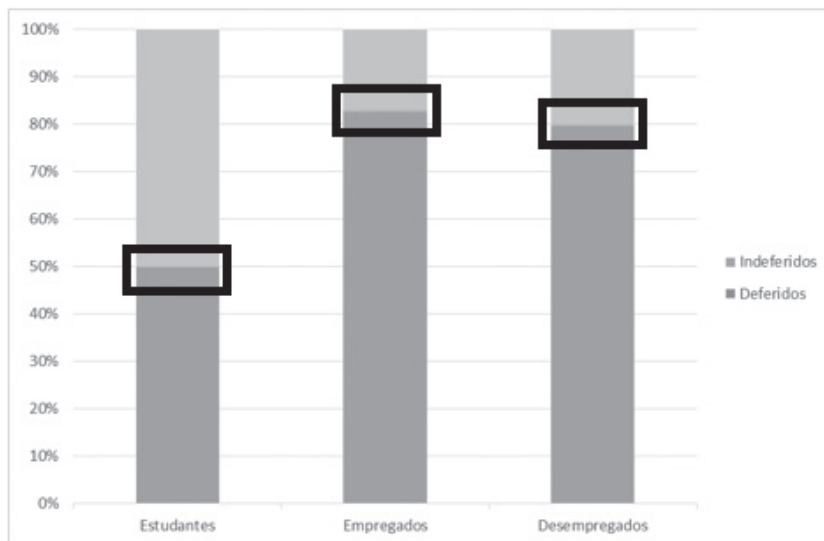


Nesse segundo caso, existe uma característica que afeta o percentual de deferimento. Podemos notar que o percentual de deferimento dos estudantes é bem menor. Esse provavelmente seria um grupo de risco a ser destacado

Porém uma decisão desse tipo não pode ser tomada com base somente no olhar. Precisamos de um critério objetivo que impeça a existência de arbitrariedades. Esse critério é o da superposição das margens de erro.

Nesse ponto, entra o papel da margem de erro. Podemos ilustrar o cálculo da margem de erro nos resultados da maneira apresentada na figura 10.

Figura 10: Grupos em teste hipotético (percentual diferente com margem de erro calculada)



Na figura 10, a margem de erro está sendo mostrada como um quadrado dentro do qual o limite entre os deferimentos vai estar. Para a fórmula utilizada, podemos dizer, com 95% de confiança, que o percentual, no futuro, ficará dentro desses limites.

Portanto, no exemplo, é fácil visualizar que, mesmo com as incertezas vinculadas a uma extrapolação dos dados para o futuro, os percentuais de deferimento entre os estudantes e os demais não se sobrepõem. Então é razoável definirmos esse grupo como diferente dos demais.

No caso da comparação entre empregados e desempregados, existe grande sobreposição. Isso significa que não podemos dizer que seus percentuais de deferimento são realmente diferentes, porque a sobreposição diz que há maiores chances de que eles sejam idênticos. Por causa disso, não podemos destacá-los como grupos separados entre si.

Ao final da verificação de todos os grupos, excluídos aqueles que não podem ser destacados, o resultado é uma planilha como a exemplificada na tabela 3.

Tabela 3: Resultado da pesquisa por grupos seguros e de risco

Grupos encontrados	Tipo	Percentual de deferimento	Margem de erro
Contribuinte com característica x	Grupo Seguro	95%	3%
Contribuinte com característica y	Grupo Seguro	91%	2%
Contribuinte com característica z	Grupo Seguro	89%	4%
Contribuinte com característica w	Grupo de Risco	9%	3%

### 3.7 Passo 7: Associar o percentual de deferimento aos processos do estoque

Depois de finalizado o passo 6, é necessário verificar se há, ou não, processos não analisados (em estoque) que se enquadrem nos dos grupos encontrados.

Isso é feito verificando, para cada processo em estoque, se ele possui característica que tenha sido utilizada para formar algum grupo. Se possuir, pertence ao grupo correspondente e podemos associar a ele o percentual esperado de deferimento daquele grupo.

Nem todos os processos são enquadrados. Além disso, alguns processos podem se enquadrar em dois grupos. Caso ocorra essa segunda hipótese, sugerimos agir a favor da segurança e utilizar o grupo com percentual de deferimento mais baixo.

O resultado desse enquadramento está exemplificado na tabela 4.

Tabela 4: Processos enquadrados em grupo e com probabilidade associada

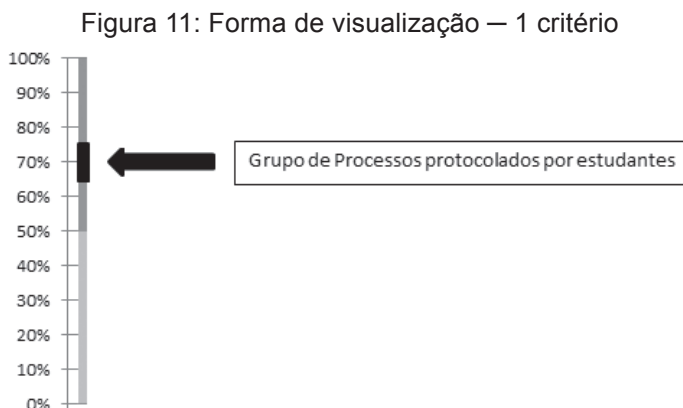
Processo em estoque	Grupos encontrados	Tipo	Percentual de deferimento	Margem de erro
10220.111.111/2015-01	Contribuinte com característica x	Grupo Seguro	95%	3%
10220.222.222/2015-02	Contribuinte com característica y	Grupo Seguro	91%	2%
10220.333.333/2015-03	Contribuinte com característica y	Grupo Seguro	91%	2%
10220.444.444/2015-04	Contribuinte com característica y	Grupo Seguro	91%	2%
10220.555.555/2015-05	não enquadrado			
10220.666.666/2015-06	não enquadrado			
10220.777.777/2015-07	Contribuinte com característica w	Grupo de Risco	9%	3%

## 4 Apresentação visual dos resultados

### 4.1 Critério de risco mostrado isoladamente

Com o conhecimento de quais são os grupos de risco, podemos mostrar os resultados não apenas em uma planilha, como na tabela 3, mas também de uma maneira visual.

Isso pode ser feito localizando cada grupo em uma linha, de acordo com seu percentual esperado de deferimento. Na figura 11, está o enquadramento de um grupo fictício de “processos protocolados por estudantes” que teria probabilidade de indeferimento de 70% e margem de  $\pm 5\%$ .



O interessante da presente proposta é poder ir além do enquadramento qualitativo e mostrar uma hierarquia quantitativa e contínua entre os grupos. Isso significa que o grupo não é meramente “de risco” ou “seguro” de uma maneira binária. É um grupo com grau determinado de segurança associado a uma margem de erro, ou seja, o grupo é, por exemplo, “seguro porque a probabilidade de deferimento é  $x\%$  com margem de erro de  $\pm y\%$ ”.

#### 4.2 Critério de risco associado com outros critérios – o cubo de critérios

O método de enquadramento de pedidos para tratamento diferenciado poderia até terminar só com o enquadramento em grupos de risco, mas é possível melhorar ainda mais a inteligência na análise dos processos se agregarmos outros critérios, além do risco.

Podemos, por exemplo, enquadrar os pedidos também em função do seu valor. Essa visão é relevante pela diferença de impacto que há, por exemplo, entre um pedido de restituição de R\$10,00 e outro de R\$1.000.000,00.

O valor associado ao pedido é uma medida do impacto da análise na sociedade, algo que é independente do risco. Por isso, é interessante visualizarmos essa variável de forma separada.

Visualmente, a melhor forma de mostrar dois critérios independentes é montar uma matriz com um dos critérios no eixo x e o outro no eixo y, como na figura 12.

Figura 12: Forma de visualização — 2 critérios

Alto Risco		■
Baixo Risco	■	
	Baixo Valor	Alto Valor

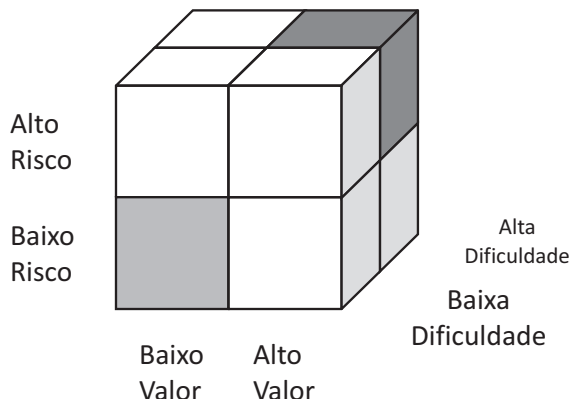
Com isso, temos separados em quatro quadrantes para localizar os grupos encontrados de acordo com combinações de risco e valor. O quadrante vermelho seria o escolhido para análise aprofundada. O verde para análise sumária. Os demais poderiam ter análise comum.

Cabe a observação de que seria até possível que o critério valor fosse o determinante de algum grupo de risco. Nessa situação, ele teria papel duplo. Apareceria como um grupo de risco e também como um eixo isolado.

Além dos critérios de risco e valor, o tipo do pedido também precisa ser levado em consideração de forma separada. A justificativa é que o tipo de pedido é variável totalmente independente das anteriores e é importante para priorização, já que está ligado à dificuldade de análise e qual será a legislação aplicável.

Para conseguirmos visualizar essa nova variável de forma separada, podemos apelar para um eixo z, tornando o quadro anterior um cubo tridimensional, como na figura 13.

Figura 13: Forma de visualização — 3 critérios — cubo de critérios



Esse é o cubo de critérios, com critérios que consideramos bons para justificar tratamento diferenciado entre grupos de processos. Para isso, é necessário verificar em qual quadrante o grupo do processo está.

São oito quadrantes. O quadrante vermelho é aquele onde estão os pedidos mais difíceis, de maior valor e de maior risco. Certamente merecem tratamento aprofundado. No quadrante verde, estão os pedidos simples, de baixo valor e de baixo risco, ideal para tratamento sumário.

Estando os grupos de processos localizados dentro do espaço desse cubo, temos o produto final do presente trabalho. É possível realizar uma alocação de esforços de análise muito mais inteligente.

## 5 Dificuldades enfrentadas e suas soluções

### 5.1 Limites entre grupos

Ao formar grupos para cada característica, conforme explicado no item 3.2, na maioria das vezes, o enquadramento é direto. Por exemplo, para a característica “envio de retificadora”, o enquadramento é binário, se o contribuinte enviou, está enquadrado. Não há dificuldades. Para algumas características, pode até haver algum desdobramento, por exemplo, na “versão do PGD utilizado para envio da declaração”, cada versão do PGD forma um grupo diferente. Isso também não é problema.



No entanto existem situações em que a formação de grupo leva a discussões mais complexas. São os casos em que as características não são rótulos claros.

O caso mais importante é o das características que são valores contínuos, como, por exemplo, a renda. Se a pessoa tem renda de R\$5.000,00 por mês, em que grupo ela será enquadrada? No grupo dos que ganham exatamente R\$5.000,00? Nesse caso, teríamos que criar outro grupo diferente para quem ganha R\$5.000,01, outro com R\$5.000,02 e assim por diante. Se fizéssemos isso, nenhum desses grupos teria uma quantidade de membros suficientemente grande para podermos aplicar inferências estatísticas.

A resposta é aumentar a faixa. No entanto isso leva a outras questões, porque teremos que escolher qual será o tamanho da faixa e onde estará seu limite. O problema é que, dependendo do limite escolhido, podem surgir ou não correlação com o deferimento.

A nossa recomendação é, para a maioria das características com valores contínuos, substituir o uso de faixas de mesmo tamanho por faixas em escala logarítmicas. O motivo disso é que diversos estudos apontam que o comportamento humano geralmente é representado melhor nessa escala (BELLOS, 2015; GOLDSTEIN, 2009; MATTHEWS, 2000).

Isso pode ser demonstrado em uma série de exemplos. Um deles é que a maioria das pessoas concorda ser uma boa decisão largar um emprego com rendimento mensal de R\$1.000,00 para buscar outro de R\$2.000,00. Mas a mesma maioria não considera válido largar um emprego com renda de R\$10.000,00 para buscar um de R\$11.000,00. Trata-se de um raciocínio proporcional. Isso faz com que a variação do comportamento, em função do valor, não seja bem representada em escala linear, mas sim na logarítmica.

No caso específico, falar em escala logarítmica é outra forma de dizer que as faixas *mudam sucessivamente, mas na mesma proporção*, como, por exemplo, nas faixas: R\$500,00 a R\$1.000,00; R\$1.000,00 a R\$2.000,00; R\$2.000,00 a R\$ 4.000,00. As faixas do exemplo vão dobrando sucessivamente de tamanho, ao invés de manter o mesmo tamanho inicial de R\$500,00.

Na prática, foi possível observar que esse fenômeno é verdadeiro. Na maioria dos grupos criados com base em valor, não foram encontradas correlações fortes para grupos formados com faixas lineares, mas foi encontrada correlação quando as mesmas características tiveram seus grupos com forma dos com faixas logarítmicas. No nosso caso, o tamanho de cada faixa foi dimensionado para que fosse sempre o dobro da faixa anterior, mas, quanto a essa proporção de aumento, cabe testar e verificar os melhores resultados para cada situação.

Apesar da grande prevalência desse comportamento, isso não significa que todas as características numéricas têm sempre comportamento logarítmico. É preciso pensar caso a caso. Para a idade, por exemplo, é mais recomendado separar por faixas de 10 anos. O motivo é que a influência da idade no comportamento tem, geralmente, aspecto mais linear, em função da faixa etária.

Outro caso interessante é o do número do CEP, utilizado para caracterizar localização geográfica. A decisão a se tomar é qual será o tamanho da faixa geográfica de cada grupo. Na prática, isso significa definir quais dígitos do CEP serão utilizados. Utilizar só o primeiro dígito englobaria uma cidade inteira. Usar todos os dígitos restringiria o grupo para o tamanho de uma vizinhança.

Nas nossas análises, os melhores resultados vieram na situação intermediária. Usar somente os três primeiros dígitos do CEP, em geral, traz melhores resultados. Estes dígitos delimitam o equivalente a um bairro grande.

## **5.2 Poucos indivíduos no grupo**

No item 3.5, mostramos como é feito o cálculo da incerteza associada à extrapolação de dados. É importante destacar que um dos fatores que influencia na incerteza da extrapolação é a quantidade de membros do grupo amostra.

A margem de erro diminui à medida que a amostra se torna maior, porque, com dados robustos do que aconteceu no passado, a certeza sobre o futuro é maior. Com menos dados, a certeza diminui.

Ellenberg (2015) alerta para esse problema e relata caso real em que escolas pequenas que passavam por determinado exame avaliativo ficavam ora entre as melhores e ora entre as piores. Uma variabilidade que não fazia sentido, ainda mais a curto prazo. Esse resultado intrigou os envolvidos até que ficou comprovado tratar-se de mera consequência do fato que a menor quantidade de alunos das escolas menores tem o comportamento de uma amostra pequena, que tem alta variabilidade. As escolas grandes obtinham sempre avaliações estáveis.

Grupos menores estão mais suscetíveis a fatores aleatórios, ou seja, à sorte. Essa influência só começa a sumir com uma massa de dados grande, pois só nela os casos de sorte acabam contrabalanceados por casos de azar. É por esse motivo que o tamanho da amostra faz parte da fórmula da margem de erro. Amostras com poucos membros apresentam margem de erro elevada, ou seja, maior incerteza em relação ao percentual de deferimento.

Para evitar incertezas excessivas, a solução é realizar corte entre os grupos encontrados para eliminar todos que contenham um número de membros menor que um valor mínimo preestabelecido.

Sugerimos definir esse valor utilizando a fórmula de cálculo da margem de erro mostrada no item 3.5. O procedimento é definir, *a priori*, um valor para a margem de erro aceitável. Depois, basta fazer o cálculo para encontrar o tamanho da amostra que resulta na margem definida.

### **5.3 Relevância do percentual de deferimento**

Lembremos que o método descrito neste trabalho mostrou como encontrar um grupo com comportamento diferenciado em relação à população como um todo.

Porém cabe deixar o alerta para o caso da existência de grupo com percentual de deferimento, muitas vezes, menor que o percentual de deferimento da população.

Uma diferença relativa grande pode saltar aos olhos e induzir a pensarmos que esse é um grupo para ser tratado de maneira diferenciada, mas *dependendo do caso, o grupo não seria passível de receber tratamento diferenciado*.

Ellenberg (2015) explica que a atenção tem que se voltar também para o percentual absoluto de deferimento. Não podemos tomar atitudes com base na diferença relativa entre o percentual do grupo e o da população. Não é correto dar o resultado nos termos “é x vezes menor/maior”.

Isso porque, se o percentual de deferimento da população como um todo já for baixo, um grupo possuir percentual “x vezes menor” não o deixa tão separado da média da população em termos absolutos. Isso é porque ambos serão igualmente pequenos. Exemplo: se a população tem percentual de deferimento de 0,001 e o do grupo é de 0,0001. Um é 10 vezes menor. A diferença relativa é muito grande, mas a diferença absoluta é irrisória: estão separados por um pequeno valor igual a 0,0009. Nesses casos, não faz sentido que o grupo seja tratado de maneira diferente dos demais.

A solução para não cair nessa armadilha é sempre adotar um critério de valor absoluto para que se trate um grupo de maneira diferenciada.

O critério deve ser discutido em cada caso, mas a sugestão é mantê-lo o mais alto possível para tratamento benéfico (sugerido de  $\geq 70\%$  de deferimento) e o mais baixo possível para tratamento que crie algum procedimento adicional (sugerido de  $\leq 30\%$  de deferimento).

Essa trava elimina a possibilidade de tratamentos diferenciados indevidos.

#### **5.4 Diferença entre correlação e causa**

Ellenberg (2015) também alerta para os perigos de confundir métodos de correlação com causa. O que o método descrito neste trabalho se propõe a fazer é encontrar correlações, não causas. Isso significa que ele não explica o porquê das correlações que encontra. É importante esclarecer isso.

Ocorre que é normal cometer o erro de atribuir aquela correlação a algum motivo totalmente imaginado e provavelmente falso. Frases como “pessoas desse endereço não têm o cuidado devido ao enviar o pedido” ou “pessoas dessa profissão têm mais chance de ter comportamento fraudulento” ou “quem tem empresa nesse bairro costuma sonegar” etc. Pinker (2013) destaca que tentativas de explicação sem

os testes apropriados só servem para confundir, gerando uma série de mal entendidos.

Um exemplo ilustrativo: se for realizado um estudo sobre correlação entre consumo de chimarrão e sotaque gaúcho, certamente a correlação seria encontrada. Nesse caso, é fácil ver que a correlação não tem nada a ver com causa. Apesar de haver correlação entre chimarrão e sotaque gaúcho, não é o chimarrão que causa o sotaque gaúcho.

É importante que exista clareza sobre a diferença entre correlação e causa para evitar erros na interpretação dos resultados. O uso incorreto do conceito pode causar problemas.

## 6 Conclusão

Com a finalidade de melhorar a eficiência das análises pela concentração de esforços em processos de alto risco e tratamento sumário em processos de baixo risco, foi descrito o método que calcula, *a priori*, qual a probabilidade dos pedidos serem deferidos ou indeferidos. A grande força do método é que o cálculo é feito com base em informações sobre comportamento real de contribuintes, não em suposições.

A previsão sobre o deferimento de um processo é informação valiosa que permite alocar esforços no lugar certo.

Essas técnicas têm aplicação imediata da RFB, já que tudo é realizado com informações já disponíveis nos bancos de dados da instituição. Durante a construção do método, a pesquisa intensa foi realizada. Agora, passos estão traçados. Ferramenta de aceleração, utilizando programação em VBA-Macro, está pronta. A replicação em outras unidades é simples.

A vantagem dessa metodologia é que o critério de seleção de processos fica à prova de arbitrariedades e suposições que podem estar erradas, pois o método não supõe nada, ele verifica o comportamento real dos processos já analisados.

A forma de visualização dos resultados é outro ponto relevante. O cubo de critérios permite visualizar, em uma só figura, uma série de informações.

Há um balanceamento entre critérios estatísticos sólidos e a possibilidade de compreender a metodologia e seus resultados. Com isso, o gestor pode se sentir seguro sobre qual o critério que está sendo usado como justificativa para tratamentos diferenciados dos processos. Portanto, apesar de já apresentar resultados por si só, o método também pode ser encarado como uma porta de entrada para que se fortaleça uma cultura de priorização inteligente nas unidades descentralizadas. O uso da presente ferramenta pode levar a uma progressão para ferramentas mais avançadas.

Como o método permite aplicação correta de tratamentos diferenciados para os processos, há tanto ganho de produtividade quanto de presença fiscal. A produtividade é obtida pelo tratamento sumário de processos no quadrante verde do cubo de critérios (baixo risco, baixo valor e baixa dificuldade). A presença fiscal decorre da análise aprofundada dos processos na região vermelha do cubo de critérios (alto risco, alto valor e alta dificuldade).

Na prática, bons contribuintes não são incomodados com diligências infrutíferas. Maus contribuintes são analisados com mais profundidade.

Em relação ao ganho de produtividade, podemos relatar que o estoque da malha em que o método foi aplicado junto com aplicação de tratamento diferenciado diminuiu 26% (de 124.913 declarações para 91.669). Isso ocorreu no tempo recorde de três meses. Esse resultado não seria possível sem a aplicação desse método inovador.

Esse tipo de resultado entrega para a sociedade algo que ela precisa – a prestação competente de serviço público, tanto na qualidade quanto na celeridade.

## Referências

BALANIUK, Remis. A mineração de dados como apoio ao controle externo. **Revista do Tribunal de Contas da União**, v. 42, n. 117, p. 77-84, jan./abr. 2010. Disponível em: <<http://www.fiec.org.br/artigos/administracao/tcu/Miner.pdf>>. Acesso em: 10 ago.2015.

BELLOS, Alex. **Alex através do espelho**: como a vida reflete os números e os números refletem a vida. 1. ed. São Paulo: Companhia das Letras, 2015.

BESSANT, John; TIDD, Joe. **Inovação empreendedorismo**. 1. ed. Porto Alegre: Editora Bookman, 2009.

GOLDSTEIN, E. Bruce. **Encyclopedia of Perception**. 1. ed. Thousand Oaks, California, US: SAGE Publications, 2009.

ELLENBERG, Jordan. **O poder do pensamento matemático**: a ciência de como não estar errado. 1. ed. Rio de Janeiro: Zahar, 2015.

KARLGAARD, Rich. **The Soft Edge**: Where great companies find lasting success.1. San Francisco: Ed. Jossey-Bass, 2014.

MATTHEWS, Gerald. **Human performance**: cognition, stress, and individual differences.1. ed. Hove, United Kingdom: Psychology Press, 2000.

OLIVEIRA, Giovani Glaucio de. **Curso de Estatística Básica**: teoria e prática. 2. ed. São Paulo: Editora Atlas, 2015.

PINKER, Steven. **Os anjos bons da nossa natureza: por que a violência diminuiu**. 1. ed. São Paulo: Companhia das Letras, 2013.

ROCHA, Sergio. **Estatística geral e aplicada para cursos de Engenharia**. 1. ed. São Paulo: Editora Atlas, 2014.

SANDEL, Michael. **Justiça**: o que é fazer a coisa certa. 1. ed. São Paulo: Editora Civilização Brasileira, 2011.

VARGAS, Ricardo Viana. **Gerenciamento de Projetos**. 6. ed. São Paulo: Brasport, 2005.

