

UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

MÁRCIO TEODORO DIAS

**LEI DE COTAS: UM ESTUDO DA RESERVA DE VAGAS EM UMA INSTITUIÇÃO
FEDERAL DE ENSINO ATRAVÉS DA DESCOBERTA DE CONHECIMENTO EM
BASES DE DADOS**

Belo Horizonte

2014

MÁRCIO TEODORO DIAS

**LEI DE COTAS: UM ESTUDO DA RESERVA DE VAGAS EM UMA INSTITUIÇÃO
FEDERAL DE ENSINO ATRAVÉS DA DESCOBERTA DE CONHECIMENTO EM
BASES DE DADOS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Escola de Ciência da Informação da Universidade Federal de Minas Gerais para obtenção do grau de Mestre em Ciência da Informação.

Linha de Pesquisa: Gestão da Informação e do Conhecimento

Orientador: Prof. Dr. Marcello Peixoto Bax

BELO HORIZONTE

2014

D5411 Dias, Márcio Teodoro.

Lei de cotas [manuscrito] : um estudo da reserva de vagas em uma instituição federal de ensino através da descoberta de conhecimento em bases de dados / Márcio Teodoro Dias. – 2014.
84 f. : il., enc.

Orientador: Marcello Peixoto Bax.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Referências: f. 81-84.

1. Ciência da informação – Teses. 2. Gestão do conhecimento – Teses. 3. Universidades e faculdades públicas – Brasil – Teses. 4. Políticas públicas – Teses. 5. Mineração de dados (Computação) – Teses. 6. Processo decisório – Teses. I. Título. II. Bax, Marcello Peixoto. III. Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

CDU: 659.2:35



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

FOLHA DE APROVAÇÃO

"LEI DE COTAS: UM ESTUDO DA RESERVA DE VAGAS EM UMA INSTITUIÇÃO FEDERAL DE ENSINO ATRAVÉS DA DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS"

Márcio Teodoro Dias

Dissertação submetida à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais, como parte dos requisitos à obtenção do título de "**Mestre em Ciência da Informação**", linha de pesquisa "**Gestão da Informação e do Conhecimento**".

Dissertação aprovada em: 26 de setembro de 2014.

Por:

Prof. Dr. Marcello Peixoto Bax - ECI/UFMG (Orientador)

Prof. Dr. Carlos Alberto Gonçalves - FACE/UFMG
Profa. Dra. Beatriz Valadares Cendón - ECI/UFMG
Profa. Dra. Renata Maria Abrantes Baracho Porto - ECI/UFMG

Aprovada pelo Colegiado do PPGCI

Profa. Renata Maria Abrantes Baracho Porto
Coordenadora

Versão final Aprovada por

Prof. Marcello Peixoto Bax
Orientador



UFMG

Universidade Federal de Minas Gerais
Escola de Ciência da Informação
Programa de Pós-Graduação em Ciência da Informação

ATA DA DEFESA DE DISSERTAÇÃO DE **MÁRCIO TEODORO DIAS**, matrícula:
2012736054

Às 14:30 horas do dia 26 de setembro de 2014, reuniu-se na Escola de Ciência da Informação da UFMG a Comissão Examinadora aprovada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação em 01/09/2014, para julgar, em exame final, o trabalho intitulado **Lei de cotas: um estudo da reserva de vagas em uma instituição federal de ensino através da descoberta de conhecimento em bases de dados**, requisito final para obtenção do Grau de MESTRE em CIÊNCIA DA INFORMAÇÃO, área de concentração: Produção, Organização e Utilização da Informação, Linha de Pesquisa: Gestão da Informação e do Conhecimento. Abrindo a sessão, o Presidente da Comissão, Prof. Dr. Marcello Peixoto Bax, após dar conhecimento aos presentes do teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa do candidato. Logo após, a Comissão se reuniu sem a presença do candidato e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações:

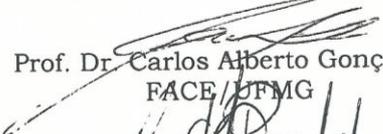
Prof. Dr. Marcello Peixoto Bax - Orientador	APROVADO
Prof. Dr. Carlos Alberto Gonçalves	APROVADO
Profa. Dra. Beatriz Valadares Cendón	APROVADO
Profa. Dra. Renata Maria Abrantes Baracho Porto	APROVADO

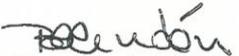
Pelas indicações, o candidato foi considerado APROVADO.

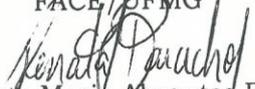
O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a sessão, da qual foi lavrada a presente ATA que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 26 de setembro de 2014


Prof. Dr. Marcello Peixoto Bax
ECI/UFMG


Prof. Dr. Carlos Alberto Gonçalves
FACE/UFMG


Profa. Dra. Beatriz Valadares Cendón
ECI/UFMG


Profa. Dra. Renata Maria Abrantes Baracho Porto
ECI/UFMG


Prof(a). Renata Maria Abrantes Baracho Porto
Coordenadora do Programa de Pós-Graduação
em Ciência da Informação

Obs: Este documento não terá validade sem a assinatura e carimbo da Coordenadora.

Dedico este trabalho à minha esposa Cristina

AGRADECIMENTOS

Aos meus pais João e Maria, por todos os ensinamentos.

À minha esposa Cristina, por todo o apoio, incentivo e compreensão.

A todos os colegas das disciplinas do Mestrado, pela convivência e pelo aprendizado, principalmente ao amigo Amarildo Martins de Magalhães.

A todos os professores com que tive contato, nas aulas, seminários e no exame de qualificação, especialmente ao orientador, o prof. Marcello Bax.

A todos que contribuíram de alguma forma para a realização deste trabalho.

Saber claramente quais são os nossos sonhos nos dá a segurança de saber onde queremos chegar, a tranquilidade de saber quando chegamos e a coragem para não pararmos no meio do caminho.

Bernardo Stamateas

RESUMO

Este trabalho relata um estudo de caso que analisou a efetividade da reserva de vagas em instituições federais de ensino, criada pela Lei de Cotas. O estudo contextualizou a lei e sua implementação no Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG), e buscou avaliar a real ocupação das vagas reservadas; descobrir se os candidatos que possuem os requisitos estão fazendo uso do direito de se inscreverem nas vagas reservadas; descobrir se há candidatos concorrendo às vagas sem ter o direito; e descobrir se há diferenças significativas no perfil de candidatos que chega a ser matriculado em cotas em relação aos candidatos matriculados na ampla concorrência. Fundamentado em conceitos da área Descoberta do Conhecimento em Bases de Dados, o estudo realizou a extração e preparação dos dados e aplicou ferramenta de mineração nos dados extraídos. Os dados se referem a inscrições nos processos seletivos e matrículas no IFMG no período de pouco mais de um ano. Em seguida os resultados foram analisados, mostrando que uma parcela significativa das vagas reservadas não foi ocupada; que muitos candidatos que poderiam se inscrever nas vagas reservadas não o fizeram; que muitos candidatos concorrem às vagas reservadas sem possuir os requisitos exigidos por lei; que não houve diferenças significativas no ingresso de alunos com os requisitos pelo sistema de cotas ou pela ampla concorrência, o que leva a crer que o sistema de cotas seria desnecessário para a instituição estudada. Finalmente, os resultados sugeriram a necessidade de maior divulgação da política de cotas, alertaram para a importância da verificação das informações prestadas pelos candidatos que concorreram às vagas e chamaram a atenção para a necessidade de estudos futuros em outras instituições federais de ensino sobre a efetividade da política de cotas, visando contribuir para o debate sobre a relevância dessa política para o país.

Palavras-chave: Gestão da Informação e do Conhecimento. Tomada de decisão. Lei de Cotas. Mineração de dados. Descoberta de Conhecimento em Bases de Dados.

ABSTRACT

The work reports a case study to analyze the effectiveness of the places reserved in federal educational institutions by the Brazilian Quota Law. The study contextualizes the law and its implementation at the Federal Institute of Education, Science and Technology of Minas Gerais (IFMG), and sought to evaluate the actual occupancy of the reserved places; to verify the extent to which people who have used their right to the reserve; find out if there are people registered to places without the right to do so; and find out if there are significant differences in the profiles of candidates as to be enrolled in quotas in relation to candidates enrolled in wide competition. Based on concepts of Knowledge Discovery in Databases, the study describes data extraction and preparation and applied a data-mining tool to the extracted data. Data refer to entries in the selection process and enrollment in IFMG for a period of just over a year. The results are analyzed, showing that a significant proportion of reserved places were not occupied; that many candidates who could enroll in reserved spaces did not; that many candidates competed for the reserved places without owning the requirements prescribed by law; and there were no significant differences in the admission of students to the requirements by the quota system or the wide competition, which suggests that the quota system would be unnecessary for the institution studied. Finally, the results suggest the need for greater advertisement of the quota policy, call attention to the importance of verifying the data provided by the candidates who ran for the places and pointed to the need for future studies in other federal educational institutions on the effectiveness of the quota policy, to contribute to the debate on the relevance of this policy for the country.

Keywords: Information and Knowledge Management. Decision making. Quota Law. Data Mining. Knowledge Discovery in Databases

LISTA DE FIGURAS

Figura 1 – As etapas do Processo de KDD	26
Figura 2 – Esforço por etapas de KDD.....	27
Figura 3 – Tipos de cotas	33

LISTA DE TABELAS

Tabela 1 – Percentual de vagas reservadas para cada grupo	35
Tabela 2 – Percentual de similaridade entre campos informados em diferentes locais do sistema	40
Tabela 3 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública	41
Tabela 4 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública, agrupados pela renda familiar <i>per capita</i>	42
Tabela 5 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública, agrupados pela autodeclaração de cor	43
Tabela 6 – Inscrições e matrículas nas vagas reservadas para cada tipo de cota ...	44
Tabela 7 – Inscrições e matrículas nas vagas reservadas em cursos superiores para alunos que estudaram em escola pública	45
Tabela 8 – Inscrições e matrículas nas vagas reservadas para alunos de cursos superiores que estudaram em escola pública, agrupados pela renda familiar <i>per capita</i>	46
Tabela 9 – Inscrições e matrículas nas vagas reservadas para alunos de cursos superiores que estudaram em escola pública, agrupados pela autodeclaração de cor	46
Tabela 10 – Inscrições e matrículas nas vagas reservadas para cursos superiores para cada tipo de cota	47
Tabela 11 – Inscrições e matrículas nas vagas reservadas em cursos técnicos subsequentes para alunos que estudaram em escola pública	48
Tabela 12 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos subsequentes que estudaram em escola pública, agrupados pela renda familiar <i>per capita</i>	49
Tabela 13 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos subsequentes que estudaram em escola pública, agrupados pela autodeclaração de cor	50
Tabela 14 – Inscrições e matrículas nas vagas reservadas para cursos técnicos subsequentes para cada tipo de cota	51
Tabela 15 – Inscrições e matrículas nas vagas reservadas em cursos técnicos integrados para alunos que estudaram em escola pública	52

Tabela 16 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos integrados que estudaram em escola pública, agrupados pela renda familiar <i>per capita</i>	53
Tabela 17 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos integrados que estudaram em escola pública, agrupados pela autodeclaração de cor	53
Tabela 18 – Inscrições e matrículas nas vagas reservadas para cursos técnicos integrados para cada tipo de cota	54
Tabela 19 – Candidatos que poderiam ter se inscrito em alguma opção de cota	55
Tabela 20 – Candidatos que poderiam ter se inscrito em cotas com critérios raciais e econômicos	56
Tabela 21 – Candidatos que poderiam ter se inscrito em cada grupo de cotas	56
Tabela 22 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em alguma opção de cota	57
Tabela 23 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cotas com critérios raciais e econômicos	58
Tabela 24 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cada grupo de cotas	58
Tabela 25 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em alguma opção de cota	59
Tabela 26 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cotas com critérios raciais e econômicos	59
Tabela 27 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cada grupo de cotas	60
Tabela 28 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em alguma opção de cota	61
Tabela 29 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cotas com critérios raciais e econômicos	61
Tabela 30 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cada grupo de cotas	62
Tabela 31 – Candidatos inscritos com os requisitos	63
Tabela 32 – Candidatos inscritos com os requisitos agrupados pela cor e pela renda familiar	63
Tabela 33 – Candidatos inscritos com os requisitos em cada grupo de cotas	64

Tabela 34 – Candidatos a vagas no ensino superior que se inscreveram em cotas possuindo os requisitos	64
Tabela 35 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar	65
Tabela 36 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cada grupo de cotas	65
Tabela 37 – Candidatos a vagas no ensino técnico subsequente que se inscreveram em cotas possuindo os requisitos	66
Tabela 38 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar	66
Tabela 39 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cada grupo de cotas	66
Tabela 40 – Candidatos a vagas no ensino técnico integrado que se inscreveram em cotas possuindo os requisitos	67
Tabela 41 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar	67
Tabela 42 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cada grupo de cotas	68
Tabela 43 – Candidatos que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência	69
Tabela 44 – Candidatos a vagas no ensino superior que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência	70
Tabela 45 – Candidatos a vagas no ensino técnico subsequente que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência	71
Tabela 46 – Candidatos a vagas no ensino técnico integrado que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência	72
Tabela 47 – Relação de inscritos por vaga agrupados pela autodeclaração de cor	74
Tabela 48 – Ocupação das vagas reservadas por cota e por tipo de curso	75

LISTA DE QUADROS

Quadro 1 – Campos extraídos do banco de dados	36
Quadro 2 – Novos campos gerados a partir dos campos extraídos do banco de dados	38

LISTA DE ABREVIATURAS

CSV	Comma-separated values
ERP	Enterprise Resource Planning
IFMG	Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais
KDD	Knowledge Discovery in Databases
SQL	Structured Query Language

SUMÁRIO

1	INTRODUÇÃO	18
1.1	Problema	18
1.2	Objetivos.....	19
1.2.1	Objetivo geral.....	19
1.2.2	Objetivos específicos.....	19
1.3	Justificativa.....	20
1.4	Estrutura da dissertação.....	20
2	CONCEITOS GERAIS E REVISÃO DA LITERATURA.....	22
2.1	Gestão do Conhecimento.....	22
2.2	Descoberta de Conhecimento em Banco de Dados.....	24
2.3	Trabalhos correlatos	30
2.4	Lei de Cotas	32
2.5	Instituto Federal de Minas Gerais	34
3	METODOLOGIA	36
3.1	Preparação dos dados	36
3.2	Mineração dos dados	39
3.3	Validação dos dados do questionário socioeconômico	39
4	APRESENTAÇÃO DOS RESULTADOS	41
4.1	Ocupação das vagas reservadas para cotas	41
4.1.1	Todos os cursos	41
4.1.2	Cursos Superiores	45
4.1.3	Cursos Técnicos Subsequentes	48
4.1.4	Cursos Técnicos Integrados	52
4.2	Candidatos que poderiam concorrer às vagas, mas não o fizeram	55
4.2.1	Todos os cursos	55
4.2.2	Cursos Superiores	57

4.2.3	Cursos Técnicos Subsequentes	59
4.2.4	Cursos Técnicos Integrados	60
4.3	Candidatos inscritos com os requisitos necessários	62
4.3.1	Todos os cursos	62
4.3.2	Cursos Superiores	64
4.3.3	Cursos Técnicos Subsequentes	65
4.3.4	Cursos Técnicos Integrados	67
4.4	Estudantes matriculados nas cotas e na ampla concorrência com os requisitos necessários para as cotas	68
4.4.1	Todos os cursos	68
4.4.2	Cursos Superiores	70
4.4.3	Cursos Técnicos Subsequentes	71
4.4.4	Cursos Técnicos Integrados	72
5	ANÁLISE DOS RESULTADOS	73
5.1	Ocupação das vagas reservadas para cotas	73
5.2	Candidatos que poderiam concorrer às vagas, mas não o fizeram	76
5.3	Inscrições que não cumprem os requisitos necessários às cotas	77
5.4	Matriculados com os requisitos nas cotas e na ampla concorrência	77
6	CONCLUSÕES	80
	REFERÊNCIAS	84

1 Introdução

Não há dúvida de que o acesso da população à educação de qualidade é um fator essencial para o desenvolvimento econômico de qualquer nação.

Com o objetivo de ampliar o acesso ao ensino técnico e superior às populações com maiores dificuldades de ingressar no ensino público federal, em 2012 a legislação brasileira passou a determinar a reserva de pelo menos 50% das vagas em Instituições Federais de Ensino para estudantes oriundos de escola pública, sendo que parte dessas vagas reservadas é destinada a pessoas de baixa renda e/ou que se considerem pretas, pardas ou indígenas.

Este trabalho tem o objetivo geral de fornecer subsídios para a tomada de decisão sobre a aplicação da política de reserva de vagas, com o conhecimento obtido por meio de um estudo de caso.

Foram estudados os registros de inscrições e matrículas de uma instituição federal de ensino, o Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG) em um período de pouco mais de um ano, bem como os dados de um questionário socioeconômico preenchido pelo candidato durante a inscrição.

Considerando que existe uma grande diferença entre o público que concorre a vagas no ensino superior, no ensino técnico subsequente e no ensino técnico integrado, a pesquisa avaliou separadamente cada um desses tipos de curso, além de apresentar também os resultados com todos os cursos.

Técnicas de Descoberta de Conhecimento em Bases de Dados foram utilizadas para obter informações úteis a partir dos dados brutos extraídos de um sistema informatizado. Pretendeu-se que, essas informações, depois de serem interpretadas, pudessem revelar conhecimentos sobre a efetividade da política de reserva de vagas, e que estes conhecimentos apoiassem a tomada de decisões a respeito dessa política.

1.1 Problema

Desde novembro de 2012, em todos os processos seletivos vestibulares do IFMG, são reservadas 50% das vagas de acordo com as regras definidas na legislação federal.

Saber se essa política de reserva de vagas tem sido efetiva no IFMG, fazendo com que o público alvo ingresse nos cursos, é um importante insumo para a tomada de

decisão na própria instituição, uma vez que podem ser necessárias ações para melhorar a divulgação das vagas ou o controle das informações prestadas pelos candidatos.

O debate sobre a necessidade da reserva de vagas levanta algumas questões, que não se referem apenas ao IFMG, mas a todas as instituições federais de ensino. Os candidatos oriundos de escola pública não conseguiriam ocupar 50% das vagas sem essa reserva? Candidatos de baixa renda e que se consideram pretos, pardos ou indígenas precisam das cotas para ocupar o percentual de vagas definidos a eles nos processos seletivos?

Responder a essas perguntas contribuirá para o entendimento da efetividade dessa política, o que é fundamental para a tomada de decisão a respeito de sua continuidade, ou da necessidade de aperfeiçoamentos na mesma.

1.2 Objetivos

Esta pesquisa busca atender aos objetivos descritos a seguir.

1.2.1 Objetivo geral

Fornecer subsídios para a tomada de decisão em relação à política de cotas para o acesso aos cursos do Instituto Federal de Minas Gerais, por meio do estudo dos dados dos últimos processos seletivos, com o apoio de técnicas de Descoberta de Conhecimento em Bases de Dados.

1.2.2 Objetivos específicos

Com este trabalho, pretendeu-se compreender a efetividade da política de cotas no Instituto Federal de Minas Gerais. Para chegar a esta compreensão, foram definidos os seguintes objetivos específicos:

- Avaliar a real ocupação das vagas reservadas;
- Descobrir se os candidatos que possuem os requisitos estão fazendo uso do direito de se inscreverem nas vagas reservadas;
- Descobrir se há candidatos concorrendo às vagas sem ter o direito;
- Descobrir se há diferenças significativas no perfil de candidatos que chega a ser matriculado em cotas em relação aos candidatos matriculados na ampla concorrência;

– Perceber as diferenças entre o público que se candidata a vagas de nível superior, e nas modalidades técnico integrado e técnico subsequente.

1.3 Justificativa

A política de cotas é utilizada em toda a rede federal de ensino, composta por 59 Universidades Federais e 38 Institutos Federais de Educação, Ciência e Tecnologia. É uma política que tem gerado grande debate sobre a sua viabilidade, e avaliar a sua aplicação pode trazer grande contribuição para o entendimento de sua relevância para o país, e a partir desse conhecimento gerado, apoiar a tomada de decisões sobre possíveis modificações.

Na pesquisa bibliográfica, não foi encontrado nenhum estudo com o mesmo objetivo, ou seja, este trabalho traz uma contribuição original e chama a atenção para a importância da realização de outros estudos similares.

A realidade encontrada no IFMG pode se repetir em outras instituições federais de ensino. Os passos seguidos nesta pesquisa podem ser aplicados por qualquer outra instituição federal de ensino, seja um Instituto Federal ou uma Universidade Federal, interessada em estudar essa mesma política de reserva de vagas.

Este trabalho busca extrair informações úteis a partir dos dados armazenados, considerando como Borko (1968, p. 3) que “A Ciência da Informação está preocupada com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação” e como Souza (2007), que a Ciência da Informação combina o entendimento da tecnologia da informação com o estudo científico do comportamento humano em busca de informação, assim como a forma de utilização do potencial de armazenamento, organização e manipulação de dados que o computador proporciona.

1.4 Estrutura da dissertação

Esta dissertação está estruturada em seis capítulos.

O Capítulo 1 traz a introdução do trabalho, contextualizando a pesquisa, bem como apresentando o problema, a justificativa e os objetivos.

O Capítulo 2 é composto por uma revisão de literatura que traz conceitos sobre Gestão do Conhecimento e Descoberta de Conhecimento em Bases de Dados, seguido de informações sobre trabalhos correlatos; por fim explica o funcionamento da reserva de

vagas criadas pela Lei de Cotas e apresenta informações sobre o Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais, bem como sobre a aplicação da Lei de Cotas nessa instituição.

No Capítulo 3 são apresentados os procedimentos metodológicos utilizados no desenvolvimento do trabalho, descrevendo os passos para preparação e mineração dos dados.

O Capítulo 4 apresenta os resultados obtidos para cada teste realizado, cuja análise é feita no Capítulo 5.

No Capítulo 6 são expostas as conclusões, com sugestões para a tomada de decisões específicas do IFMG, bem como para o debate sobre a política de reserva de vagas. Também são apresentadas propostas de trabalhos futuros.

2 Conceitos gerais e revisão da literatura

Neste capítulo serão apresentados os conceitos que embasam a pesquisa, bem como serão apresentados trabalhos correlatos, que utilizaram a Descoberta de Conhecimento em Bases de Dados, para gerar conhecimento útil para a tomada de decisões.

2.1 Gestão do Conhecimento

Neste trabalho, serão apresentados conceitos de dados, informação e conhecimento, como forma de auxiliar na compreensão do tema estudado, mas a visão proposta é que não se faça uma divisão rígida desses conceitos. O que é considerado informação em um determinado contexto pode ser chamado de conhecimento em outro contexto, por exemplo.

Davenport e Prusak (1998) conceituam o conhecimento como uma mistura de experiência, valores, informação contextual e insight, com origem e aplicação na mente dos colaboradores. Os autores também afirmam que, nas organizações, o conhecimento está embutido em documentos, repositórios, rotinas, processos, práticas e normas organizacionais.

Muitas vezes o conhecimento em uma organização está disperso, restrito apenas na mente dos seus colaboradores. A gestão do conhecimento busca identificar, de certa forma se apropriar do conhecimento pessoal e compartilhar esse conhecimento, tornando-o organizacional. Choo (2006, p.30) observa que “o conhecimento reside na mente dos indivíduos, e esse conhecimento pessoal precisa ser convertido em conhecimento que possa ser partilhado e transformado em inovação.” Nesse sentido, Rossatto (2002) considera a Gestão do Conhecimento como um processo estratégico contínuo e dinâmico para gerir o capital intangível da empresa e todos os pontos estratégicos a ele relacionados. Para Turban, McLean e Wetherbe (2004, p. 326)

A Gestão do Conhecimento (KM) é um processo que ajuda as empresas a identificar, selecionar, organizar, distribuir e transferir informação e conhecimento especializado que fazem parte da memória da empresa e que normalmente existem dentro dela de forma não-estruturada. O conhecimento sobre como os problemas são resolvidos pode ser capturado, de sorte que a Gestão do Conhecimento promove o aprendizado organizacional, levando à criação de mais conhecimento.

Muitos autores, como Drucker (1994) e Castells (1999) consideram que a sociedade passou por grandes transformações desde a década de 1990, e que vivemos em uma época em que o conhecimento é o principal recurso das organizações. Esta

importância do conhecimento faz com que a sua gestão se torne cada vez mais necessária. De acordo com Mussi e Angeloni (2000), grande parte do conhecimento necessário à uma organização já existe nela, mas está inacessível ou indisponível, sendo que a gestão do conhecimento tem como objetivo proporcionar um ambiente para identificar, criar e compartilhar conhecimento que agregue valor à instituição.

Segundo Miranda (1999, p. 286), dados são “um conjunto de registros qualitativos ou quantitativos conhecido que, organizado, agrupado, categorizado e padronizado adequadamente, transforma-se em informação”.

Percebe-se que a cada dia cresce o volume de dados disponíveis nos sistemas de informação das instituições. Porém este crescimento dos dados não significa em si que as organizações estejam melhor informadas e adquirindo mais conhecimentos. Lyman e Varian (2003, p. 20) observam que “muito tem se falado a respeito da explosão informacional e, frequentemente, alardeia-se que a informação criada diariamente vem aumentando à medida que novas tecnologias são implementadas e muito pouco destes dados poderão ser percebidos por humanos”.

O uso cada vez mais disseminado das novas tecnologias provoca o aumento dos dados disponíveis em uma quantidade que vai além da compreensão humana. De acordo com Frawley, Piatetsky-Shapiro e Matheus (1992), o advento da digitalização de documentos e o desenvolvimento das redes, fizeram com que o volume de informação aumentasse além da capacidade humana de apreensão e, dessa forma, existe um lapso crescente entre a criação de dados e a compreensão deles.

Choo (2006) afirma que como há limitações de tempo, de recursos e, acima de tudo, de energia intelectual para identificar as alternativas, prever as consequências e esclarecer as preferências, a atenção torna-se um recurso escasso, que influencia a decisão, assim como a quantidade e a qualidade da informação na qual ela se baseia.

Toda essa massa de dados que não pode ser compreendida por pessoas, e, portanto, não é útil para a organização que a mantém, ainda pode gerar uma falsa sensação de informação. Cardoso e Machado (2008, p. 496) observam que:

Ao longo do tempo, percebeu-se que a velocidade de coleta de informações era muito maior do que a velocidade de processamento ou análise das mesmas, o que gera um problema e uma contradição, pois as organizações, por possuírem uma grande quantidade de dados, possuem uma falsa sensação de que estão bem informadas; porém essas informações de nada servem se não forem analisadas de forma correta e em tempo hábil.

Portanto, percebe-se que possuir grandes volumes de dados pode não trazer nenhuma vantagem para a organização que os possui. Moresi (2001) diz que o uso de

tecnologias é necessário para controlar o conhecimento contido em uma organização, pois a mesma precisa de políticas, procedimentos e tecnologias capazes de coletar, distribuir e utilizar efetivamente o conhecimento. Goldschmidt e Passos (2005) afirmam que a análise de grandes quantidades de dados pelo homem é inviável sem o auxílio de uma ferramenta computacional apropriada.

Com o crescimento cada vez mais rápido do volume de dados e informações das organizações, as tecnologias de informação se tornam elementos indispensáveis para as iniciativas de gestão do conhecimento. Entre as tecnologias capazes de auxiliar na gestão do conhecimento, destaca-se neste trabalho a Descoberta de Conhecimento em Bases de Dados como forma de tentar solucionar os problemas gerados pelo excesso de dados, observando como Cardoso e Machado (2008, p. 495) que

A gestão do conhecimento abrange toda a forma de gerar, armazenar, distribuir e utilizar o conhecimento, tornando necessária a utilização de tecnologias de informação para facilitar esse processo, devido ao grande aumento no volume de dados. A descoberta de conhecimento em banco de dados é uma metodologia que tenta solucionar esse problema.

2.2 Descoberta de Conhecimento em Bases de Dados

No mundo atual, os avanços tecnológicos acontecem de forma cada vez mais acelerada. A Tecnologia da Informação (TI) vem se tornando a cada dia mais estratégica e capaz de proporcionar oportunidades de aumentar a eficácia nas organizações. Os autores Takeuchi e Nonaka (2008, p.271) descrevem da seguinte forma a importância da TI nos últimos anos:

O efeito da TI no mundo dos negócios tem sido avassalador. Ela reduziu, por exemplo, o tempo exigido para acessar qualquer informação, enquanto derrubou, simultaneamente, os custos das transações. Em consequência, novos produtos e serviços emergiram e muitas novas empresas explodiram em cena, deixando em dificuldades as empresas sem conhecimento da TI.

Para Turban, McLean e Wetherbe (2004), a Tecnologia da Informação proporciona vantagens estratégicas às organizações, facilitando a solução de problemas, aumentando a produtividade e a qualidade, além de contribuir para a melhoria da comunicação e da colaboração.

Weill e Ross (2006, p. 01) afirmam que “a informação, e conseqüentemente a TI, é um elemento cada vez mais importante dos produtos e serviços organizacionais e da base dos processos empresariais.”

Porém a tecnologia por si só não traz resultados. O conhecimento, que é o grande diferencial das organizações modernas, só pode ser obtido por meio de pessoas. Essas pessoas, que compõem as organizações, precisam compreender as possibilidades que a TI oferece e utilizá-las para obter as informações que auxiliem na tomada de decisões. Sobre este ponto, Takeuchi e Nonaka (2008, p. 271) ponderam:

A administração necessita considerar uma maneira de abordar a TI de forma que ela se torne um meio de diferenciação entre uma empresa e sua concorrência. É muito fácil investir em TI, mas é difícil ter lucro com esse investimento. Sem uma intenção estratégica nítida, os investimentos em TI não pagarão dividendos. A chave para obter e sustentar a vantagem competitiva através da TI é o profundo entendimento de seu impacto nos negócios e na estratégia da empresa.

A expressão em inglês *Knowledge Discovery in Databases* (KDD), geralmente é traduzida em português como Descoberta de Conhecimento em Bases de Dados ou como Descoberta de Conhecimento em Banco de Dados ou como Extração de Conhecimento. Este é um conceito utilizado para definir “o processo não trivial de identificar padrões válidos, novos, potencialmente úteis e, finalmente, compreensíveis em dados” (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996, p. 40-41).

Segundo Amaral (2001), KDD é um processo não trivial de descoberta de novos conhecimentos, sejam padrões, tendências, associações, probabilidades ou fatos, que não são óbvios ou de fácil identificação. O autor destaca que esses padrões podem ser utilizados para a geração de ações concretas, transformando-se em vantagem competitiva.

Este processo, que é fortemente amparado pela tecnologia da informação, surgiu pela necessidade e pela dificuldade de se explorar grandes volumes de dados disponíveis nas organizações. Segundo Berry e Linoff (1997) o objetivo da KDD é encontrar padrões interessantes ocultos em grandes quantidades de dados, potencializando a habilidade de otimizar a tomada de decisão, pelo uso de métodos automáticos para aprender com ações passadas.

O processo se desenvolve a partir dos dados brutos que, em si, não tem sentido. A transformação desses dados em formas mais compactas é que permite a descoberta de conhecimento que possa ser útil à organização. Segundo Fayyad, Piatetsky-Shapiro e Smyth (1996, p.39):

O campo de pesquisa de KDD preocupa-se com o desenvolvimento de métodos e técnicas que buscam trazer sentido aos dados. Seu processo básico é traduzir a informação do seu nível mais elementar, o dado, geralmente armazenado em grandes volumes, em formas mais compactas, mais resumidas e mais úteis.

Segundo Frawley, Piatetsky-Shapiro e Matheus (1992, p. 58-59), KDD apresenta 4 características:

1 – Linguagem de alto-nível – necessariamente não precisa ser utilizada por humanos de maneira direta, mas que seja compreensível para eles.

2 – Precisão – representação das descobertas descrevem, inclusive, as imperfeições da base de dados em estudo e, portanto, retratar o seu grau de confiança é fundamental. Essa confiança envolve fatores como integridade dos dados, tamanho da amostra utilizada, entre outros. Nesse sentido, a representação dessas imperfeições é expressa por medidas de graus de confiança objetivando ter segurança necessária para justificar o conhecimento descoberto.

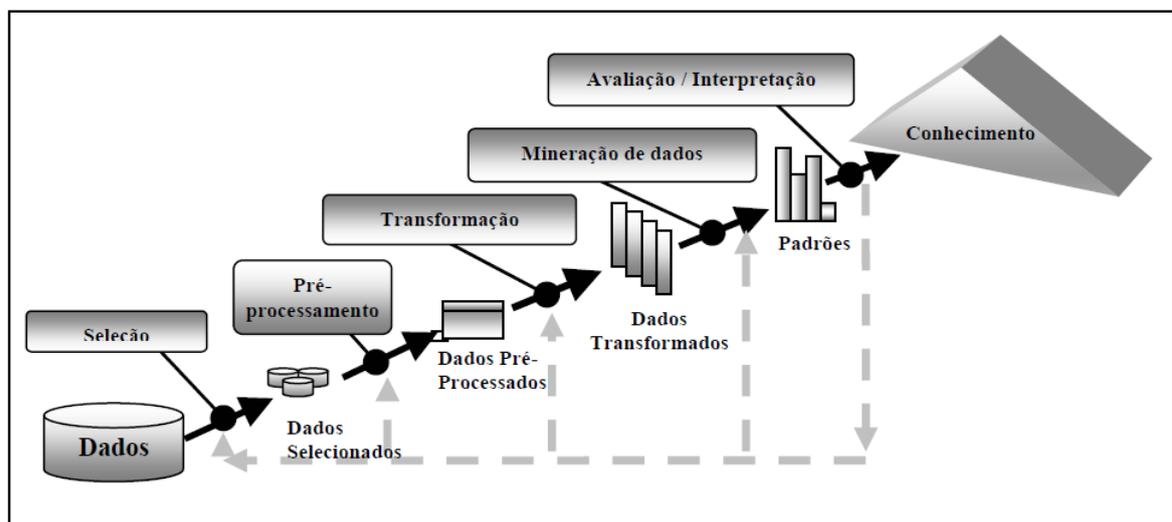
3 – Resultado interessante – os conhecimentos devem estar intrinsecamente orientados ao objetivo do usuário.

4 – Eficiência – do ponto de vista computacional, a execução das tarefas é factível, isto é, previsível e aceitável.

Muitos autores dividem o processo de KDD em etapas, com algumas diferenças entre as definições. De uma maneira geral, há um consenso de que a mineração de dados – etapa executada por um *software* – é apenas uma das fases do processo de KDD, e que existem etapas executadas por pessoas antes e depois da mineração.

De acordo com Fayyad, Piatetsky-Shapiro e Smyth (1996), existem cinco etapas, mostradas na figura 1:

FIGURA 1 – As etapas do Processo de KDD



Fonte: Adaptado de Fayyad, Piatetsky-Shapiro e Smyth (1996)

Na seleção são escolhidos e agrupados os dados que serão usados na pesquisa. Esta escolha é feita de acordo com os objetivos que se deseja alcançar. Fayyad, Piatetsky-Shapiro e Smyth (1996), afirmam que diante de cada cenário, o analista utiliza sua experiência anterior, seus conhecimentos e sua intuição para interpretar e combinar os fatos

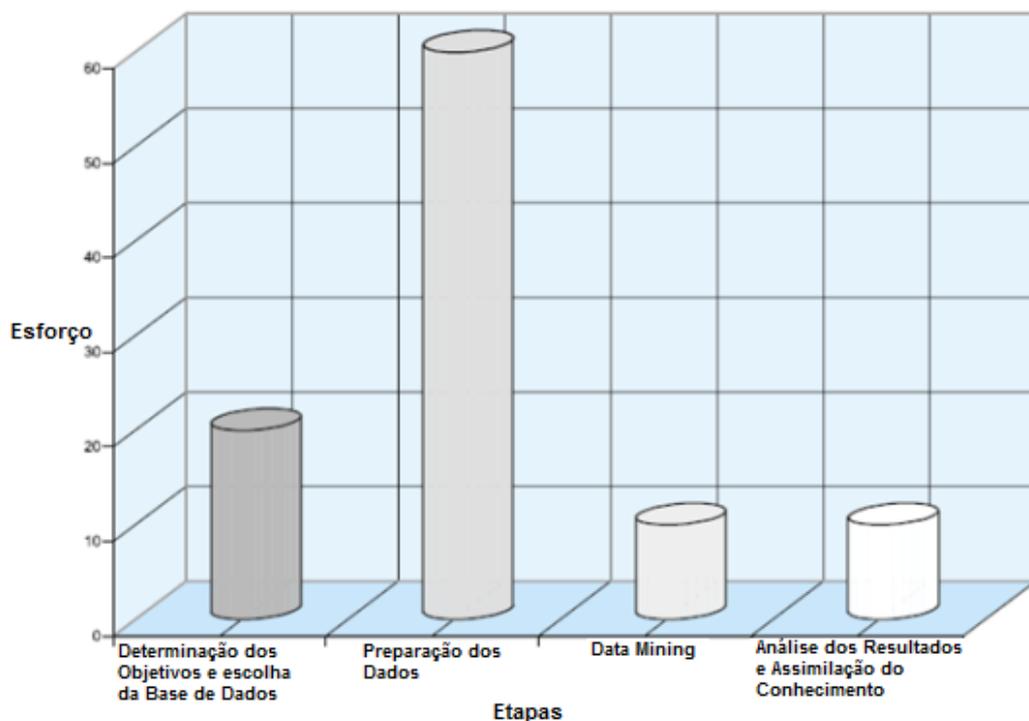
para decidir a estratégia a ser adotada. Para Boente, Goldschmidt e Estrela (2008, p. 4), esta etapa compreende “a identificação de quais informações, dentre as bases de dados existentes, devem ser efetivamente consideradas durante o processo de KDD”.

O pré-processamento é a etapa de limpeza dos dados, visando melhorar a qualidade dos mesmos. Informações ausentes, errôneas ou inconsistentes nas bases de dados devem ser corrigidas de forma a não comprometer a qualidade dos modelos de conhecimento a serem extraídos ao final do processo de KDD (BOENTE; GOLDSCHMIDT; ESTRELA, 2008).

Na transformação os dados serão codificados em formatos legíveis pelo *software* de mineração, para que possam ser utilizados como entrada dos algoritmos disponíveis. Nessa fase também pode ser obtida mais informação que agregue valor aos registros existentes, vinda de uma consulta a uma base de dados externa, por exemplo (BOENTE; GOLDSCHMIDT; ESTRELA, 2008).

Para alguns autores, como Quoniam *et al* (2001), existe uma etapa chamada de preparação, que engloba as fases definidas por Fayyad, Piatetsky-Shapiro e Smyth (1996) como pré-processamento e transformação. De uma maneira geral, os autores concordam que o maior esforço no processo de KDD acontece antes da mineração dos dados. Na figura 2, observa-se que 60% do esforço dispendido está na etapa de preparação.

FIGURA 2 – Esforço por etapas de KDD



Fonte: QUONIAM *et al*, 2001

A mineração de dados, ou *data mining*, é a etapa executada por um *software* específico para este fim. Turban, McLean e Wetherbe (2004) citam que a mineração de dados recebeu esse nome devido às semelhanças entre a pesquisa de informação de valor em grandes bancos de dados e a garimpagem em busca de minérios preciosos, pois em ambos os processos a quantidade de material a ser garimpado ou sondado é imensa. Segundo Sulaiman (2000) é a “tarefa de estabelecer novos padrões de ‘conhecimento’, geralmente imprevistos, partindo-se de uma massa de dados previamente coletada e preparada para este fim”. Destaca-se nessa definição a importância do trabalho anterior à mineração, de coleta e preparação dos dados.

Santos (2006, p. 32) define *data mining* como “um processo de suporte à decisão no qual se procura por padrões escondidos em grandes volumes de dados”. Ou seja, busca-se padrões que não seriam vistos por um ser humano, e com objetivos definidos, uma vez que a finalidade é o suporte à tomada de decisões.

Prati (2010) conceitua a mineração de dados como o processo de explorar grandes quantidades de dados à procura de padrões consistentes, concluindo que a mesma provê um extenso leque de opções para auxiliar nos processos de tomada de decisão e gestão de conhecimento.

De acordo com Fayyad, Piatetsky-Shapiro e Smyth (1996, p. 39), *data mining* é “um passo no processo KDD que consiste em aplicar análises de dados e algoritmos de descoberta que, sob limitações aceitáveis de eficiência computacional produz uma enumeração particular de padrões sobre os dados”. É interessante ressaltar nesta definição que o uso da tecnologia é apenas um passo no KDD, ou seja, a tecnologia por si só não é capaz de transformar os dados em conhecimento. Existem etapas executadas por pessoas, antes e depois do *data mining*. Da mesma forma, Bath (2004) considera a mineração de dados como parte de um processo através do qual a informação é extraída de dados e usada na tomada de decisões em diferentes contextos.

Chen e Liu (2004) afirmam que *data mining* abriu novos caminhos para a Ciência da Informação, possibilitando a pesquisa por informações valiosas em grandes volumes de dados, assim apoiando as organizações na tomada de decisões mais rápidas e com maior grau de confiança.

Por último, há a etapa de avaliação / interpretação, realizada por analistas, que utilizarão o conhecimento adquirido na tomada de decisão. Será mais efetivo de acordo com o conhecimento do analista. Segundo Santos (2006, p. 88), “O conhecimento descoberto será mais bem utilizado por usuários familiarizados com o conhecimento do domínio

trabalhado, pois estes usuários possuem o conhecimento tácito necessário para avaliar as respostas do sistema”.

Ou seja, o processo de KDD pode permitir a transformação de dados sem sentido em si mesmos em conhecimento relevante para a organização. Prati (2010) avalia que este processo é uma ferramenta essencial para uma efetiva gestão do conhecimento, uma vez que permite extrair informações valiosas dos dados armazenados, como tendências e padrões que podem ser usados para melhorar as decisões de negócios.

De acordo com Goldschmidt e Passos (2005) a complexidade do processo de KDD está na dificuldade de perceber e interpretar corretamente os fatos observados e na dificuldade de julgar tais interpretações para decidir quais ações devem ser realizadas. Quanto mais os analistas conhecerem os dados e os processos do negócio, maiores são as chances de encontrarem o conhecimento que auxilie a tomada de decisões. De acordo com Santos (2006, p. 29),

“O sistema pretende apoiar a tomada de decisões pelo analista, ou seja, pelo usuário. A resposta se relaciona com o conhecimento de domínio específico de uma certa comunidade profissional. Desta forma, o receptor do sistema idealmente deve compartilhar uma base de conhecimento de domínio para que possa interpretar e usar a resposta do sistema construtivamente, ou até rejeitá-la, se esta lhe parecer fora do contexto. Ou seja, o usuário deve ter a capacidade de tomar decisões informadas sobre a relevância do conhecimento descoberto.”

Este processo adquire sentido na medida em que permite à organização adquirir novos conhecimentos e utilizá-los em sua rotina. Afinal, segundo Davenport e Prusak, (1998), “a única vantagem sustentável que uma empresa possui, é aquilo que ela coletivamente sabe, a eficiência com que ela usa o que sabe e a prontidão com que ela adquire e usa novos conhecimentos.”

Turban, McLean e Wetherbe (2004) e Amoroso (2009) citam os seguintes exemplos de algumas das áreas de aplicação de KDD:

- Marketing: verificação de preferências e comportamento de clientes para a utilização de marketing direto;
- Distribuição e vendas: previsão de vendas, determinação de níveis adequados de estoque e cronogramas de distribuição entre lojas;
- Medicina: identificação de terapias de sucesso através do comportamento de pacientes e pesquisa por padrões de doenças novas, estabelecimento de relações entre dados demográficos de pacientes e doenças críticas;

– Instituições Financeiras: identificação de oportunidades de vendas de produtos através de análise de informações, descoberta de possíveis comportamentos em fraudes, verificação de clientes preferenciais e/ou inadimplentes, correlações entre indicadores financeiros;

– Seguradoras: previsão de volumes de indenizações, verificação de padrões de incidência de sinistros, identificação de fraudes em procedimentos médicos, verificação de distorções em procedimentos médicos, identificação de padrões de comportamento de clientes perigosos e de comportamentos fraudulentos;

– Transportadoras: identificação de padrões de comportamento de carga e rotas mais utilizadas em determinados períodos;

– Polícia: rastreamento de padrões e locais de crimes, identificação de atributos para auxiliar na solução de casos.

Prati (2010) afirma que uso de KDD permite o aprimoramento da gestão do relacionamento com o cliente, pois pode ser utilizado para que o relacionamento entre empresa e clientes seja conduzido de uma maneira mais inteligente, evitando, por exemplo, a oferta de produtos nos quais o consumidor não tenha interesse.

2.3 Trabalhos correlatos

Foram encontrados diversos trabalhos acadêmicos que tratam da utilização de KDD para apoio à tomada de decisão, incluindo teses, dissertações e artigos. Os seguintes critérios foram utilizados para a definição de trabalhos correlatos a este: pesquisas nas áreas de Ciência da Informação, Ciência da Computação ou Administração; descrições do uso de KDD para extrair conhecimento de bases de dados de instituições do governo federal ou de instituições educacionais.

Santos (2006) defendeu a dissertação com o título “Aplicação de um modelo preditivo de mineração de dados para apoio à decisão de crédito” na Escola de Ciência da Informação da Universidade Federal de Minas Gerais para obtenção do título de Mestre em Ciência da Informação. O trabalho apresentou a proposta de criação de um modelo preditivo, usando mineração de dados, capaz de identificar informações sobre o risco de crédito em instituições financeiras, visando oferecer suporte à decisão de concessão de crédito. Observa-se a correlação com este trabalho, que também pretende propor suporte a uma decisão a partir do conhecimento obtido com a mineração de dados.

Schiessl (2007) defendeu a dissertação com o título “Descoberta de Conhecimento em Texto aplicada a um sistema de atendimento ao consumidor” no Departamento de Ciência da Informação e Documentação da Universidade Federal de Brasília para obtenção do título de Mestre em Ciência da Informação. Seu trabalho apresentou a Descoberta de Conhecimento em Textos como uma extensão da Descoberta de Conhecimento em Bases de Dados e analisou uma aplicação dessas técnicas em uma base de Serviço de Atendimento ao Cliente com o objetivo de criar agrupamentos de documentos automaticamente para posterior criação de um modelo categorizador automático dos novos documentos recebidos. Assim como neste estudo, as técnicas de KDD foram usadas para a criação de um modelo, baseando-se nos dados passados, para aumentar a eficiência nas decisões.

Kisahleitner (2008) defendeu a dissertação com o título “Análise de técnicas de Data Mining na aquisição de clientes de cartão de crédito não correntistas” na Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, para obtenção do título de Mestre em Administração de Empresas. Sua pesquisa procurou compreender se a aplicação de técnicas de *data mining* em processos de aquisição de clientes de cartão de crédito, que não possuem uma conta corrente em banco poderiam trazer resultados positivos para as empresas que contam com processos de conquista de clientes. Os resultados indicaram oportunidades para o emprego da mineração de dados nesses processos, possibilitando a racionalização da operação do ponto de vista de custos de aquisição. A hipótese levantada, de que o uso de *data mining* poderia revelar conhecimento sobre um problema prático de forma a lidar com este problema de forma mais eficiente, é correlata à hipótese deste trabalho.

Filho (2009) defendeu a dissertação com o título “Mineração de regras de associação utilizando KDD e KDT: uma aplicação em segurança pública” no Instituto de Ciências Exatas e Naturais da Universidade Federal do Pará, para obtenção do título de Mestre em Ciência da Computação. As técnicas de descoberta de conhecimento foram utilizadas para analisar os dados dos delitos de estupro no Estado do Pará, nos anos de 2006 e 2007. Os resultados mostraram padrões desses delitos onde foram utilizadas armas de fogo. O objetivo apontado de propiciar um conhecimento mais aprofundado do problema que auxiliasse os tomadores de decisão a proporem medidas concretas, tem grande correlação com o objetivo deste estudo.

Cardoso e Machado (2008) publicaram o artigo “Gestão do conhecimento usando *data mining*: estudo de caso na Universidade Federal de Lavras”, descrevendo a aplicação de uma ferramenta de *data mining* para extrair conhecimento referente à produção

científica das pessoas envolvidas com pesquisa na Universidade Federal de Lavras, através da análise dos dados inseridos na plataforma Lattes. As autoras relataram uma experiência de uso de KDD para buscar um resultado que pudesse auxiliar na elaboração de novas políticas, ou seja, assim como neste trabalho, o foco do trabalho é uso do conhecimento obtido pelo processo de KDD como apoio à tomada de decisão.

Silva e Ralha (2010) publicaram o artigo “Detecção de cartéis em licitações públicas com agentes de mineração de dados”, onde descrevem um trabalho de aplicação de mineração de dados para a detecção de cartéis em licitações de compras públicas. Foram analisados os dados de pregões eletrônicos acontecidos entre os anos de 2005 e 2008, tendo como resultado a descoberta de dez regras de associação que apresentaram forte indício de cartelização. Pode-se perceber a correlação com este trabalho, na proposta de procura de padrões nos dados históricos que sirvam de embasamento para o planejamento de ações futuras.

Todos os trabalhos citados acima tem uma correlação com este trabalho no sentido em que, a partir de um problema observado na realidade, propõem o uso de técnicas de KDD para obter conhecimento sobre o problema, descrevem uma aplicação prática dessas técnicas, e tem o seu foco na utilização do conhecimento obtido para a tomada de decisão, e não no aperfeiçoamento de técnicas ou de ferramentas computacionais que são utilizadas no processo. Os trabalhos mencionados, assim como este, partem do princípio de que o conhecimento referente ao passado obtido a partir dos dados, possibilitará a descoberta de tendências em relação ao futuro, que poderão contribuir para a tomada de decisões mais eficazes.

2.4 Lei de Cotas

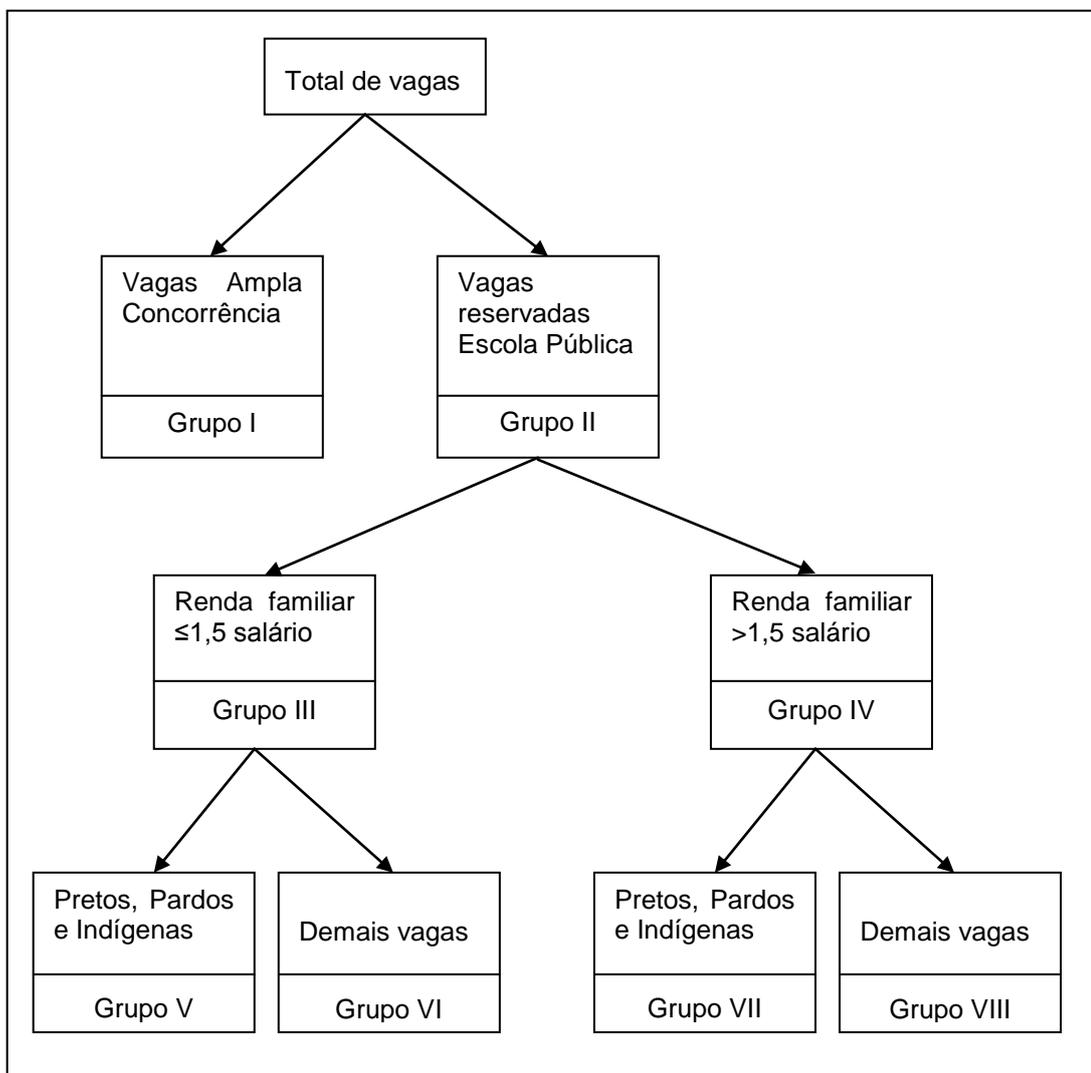
A reserva de vagas para o ingresso nas instituições federais de ensino passou a acontecer após a publicação da Lei nº 12.711 de 29/08/2012, conhecida como Lei de Cotas e de suas regulamentações, feitas no Decreto nº 7.824 de 11/10/2012 e na Portaria Normativa MEC nº 18 de 11/10/2012, sendo que está prevista a revisão desse programa no prazo de dez anos a partir da publicação da Lei 12.711/2012. A reserva de vagas funciona da seguinte forma:

De todas as vagas ofertadas em instituições federais de ensino, no mínimo 50% devem ser reservadas para estudantes que tenham cursado integralmente o nível anterior em escolas públicas. Ou seja, para as vagas de graduação, a exigência se refere ao Ensino Médio, e para as vagas de nível Técnico ao Ensino Fundamental.

No preenchimento dessas vagas, no mínimo 50% deverão ser reservadas aos estudantes com renda familiar igual ou inferior a 1,5 salário-mínimo *per capita* e as demais para estudantes com a renda familiar superior a 1,5 salário-mínimo *per capita*.

Além disso, devem ser reservadas vagas para os candidatos autodeclarados pretos, pardos ou indígenas em proporção no mínimo igual à de pretos, pardos e indígenas na população da unidade da Federação onde está instalada a instituição, segundo o último censo do Instituto Brasileiro de Geografia e Estatística (IBGE). A Figura 3 mostra essa divisão pelos tipos de cotas:

FIGURA 3 – Tipos de cotas



Fonte: Elaborada pelo autor

O número de vagas reservadas por curso deve ser informado em Edital do Processo Seletivo de cada Instituição, que tem autonomia para aumentar os percentuais previstos ou mesmo para criar outra modalidade de reserva. Num primeiro momento, todas

as instituições federais de ensino tiveram que implementar a reserva de no mínimo 25% das vagas, sendo que o prazo para que todas as instituições cumpram integralmente os critérios definidos na legislação é 30 de agosto de 2016.

O único critério que deve ser comprovado obrigatoriamente é o econômico. Quanto ao estudo em escola pública, cada instituição pode definir se exige ou não essa comprovação. Em relação à cor do candidato, só é necessária a autodeclaração. Caso após a matrícula seja apurado que o estudante prestou informação falsa, para se beneficiar de uma reserva de vaga a que não tinha direito, essa matrícula deve ser cancelada.

Sempre que o cálculo do número de vagas gerar um valor decimal, será adotado o número inteiro imediatamente superior.

No caso de não preenchimento de vagas reservadas a um determinado grupo, essas devem ser oferecidas a outro grupo de cotas, sendo levadas para a ampla concorrência como última opção.

A participação no sistema de cotas é sempre opcional. Sendo assim, mesmo os candidatos que possuem todos os requisitos para concorrer às vagas reservadas, podem decidir concorrer às vagas destinadas à ampla concorrência, se assim desejarem.

2.5 Instituto Federal de Minas Gerais

O Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais (IFMG) é uma autarquia federal, criada pela Lei nº 11.892, sancionada em 29 de dezembro de 2008 pelo então presidente Luiz Inácio Lula da Silva. No artigo 2º desta Lei, os Institutos Federais são definidos como:

Instituições de educação superior, básica e profissional, pluricurriculares e multicampi, especializados na oferta de educação profissional e tecnológica nas diferentes modalidades de ensino, com base na conjugação de conhecimentos técnicos e tecnológicos com as suas práticas pedagógicas, nos termos desta Lei.

O IFMG é hoje formado por onze unidades em diferentes cidades do estado de Minas Gerais, sendo que essas unidades são vinculadas a uma Reitoria, em Belo Horizonte. Nessas unidades são oferecidos cursos técnicos e superiores.

O Processo Seletivo Vestibular, forma utilizada para admissão nos cursos, é realizado de forma unificada ao menos duas vezes por ano. Um módulo do sistema (ERP¹) implantado na Instituição auxilia a gestão do Processo.

¹ Enterprise Resource Planning ou Sistema Integrado de Gestão

No momento da inscrição, além de informar os dados pessoais básicos, e os dados do curso a que deseja concorrer, o candidato deve informar se deseja concorrer a alguma das opções de cota e preencher um questionário socioeconômico, que traz perguntas sobre renda familiar, trabalho, estudos anteriores, entre outras.

O IFMG decidiu implementar, desde o fim de 2012 a reserva de vagas com os percentuais previstos na legislação. Portanto, em todos os processos seletivos acontecidos desde então, foram reservadas 50% das vagas para os alunos oriundos de escolas públicas. De todas as 5.231 vagas ofertadas, 2.629 foram reservadas para cotas. Considerando que, de acordo com o Censo 2010, o percentual somado de pretos, pardos e indígenas em Minas Gerais é de 53,6% (IBGE, 2010), a divisão das vagas em cada um dos 4 grupos foi feita como na Tabela 1. A variação do percentual real de vagas ofertadas em cada grupo se explica pela necessidade de arredondamento das vagas ofertadas para cada curso.

TABELA 1 – Percentual de vagas reservadas para cada grupo

Candidatos que estudaram em escola pública e têm renda familiar	inferior a 1,5 salário-mínimo <i>per capita</i> e se autodeclararam pretos, pardos ou indígenas (Grupo V)	inferior a 1,5 salário-mínimo <i>per capita</i> e não se autodeclararam pretos, pardos ou indígenas (Grupo VI)	superior a 1,5 salário-mínimo <i>per capita</i> e se autodeclararam pretos, pardos ou indígenas (Grupo VII)	superior a 1,5 salário-mínimo <i>per capita</i> e não se autodeclararam pretos, pardos ou indígenas (Grupo VIII)	Total
Número de vagas reservadas	830	512	792	495	2.629
Percentual previsto (sem arredondamento)	13,3%	11,7%	13,3%	11,7%	50%
Percentual real (após arredondamento em cada curso)	15,8%	9,8%	15,1%	9,5%	50,2%

Fonte: Elaborada pelo autor

3 Metodologia

Em relação aos seus objetivos, este estudo pode ser caracterizado como uma pesquisa descritiva, que segundo Gil (2002, p.42) “tem como objetivo primordial a descrição das características de determinada população ou fenômeno, ou então, o estabelecimento de relações entre variáveis”. Pode também ser considerada uma pesquisa quantitativa, que apresenta dados numéricos para, a partir deles, chegar a conclusões sobre o fenômeno estudado, ou seja a efetividade da aplicação da reserva de vagas no IFMG no período avaliado.

Neste capítulo estão descritos todos os procedimentos de preparação e de mineração dos dados, bem como é apresentada a ferramenta de *data mining* utilizada.

3.1 Preparação dos dados

De acordo com o objetivo da pesquisa, foi definido que seria necessário selecionar os dados referentes a todos os processos seletivos onde havia reserva de vagas para cotas e que já possuíam as informações referentes à matrícula de todos os candidatos. Quando esta pesquisa foi realizada, havia 6 Processos Seletivos com estas características no banco de dados do ERP do IFMG. Considerando que um deles continha vagas especificamente para EAD (Educação à Distância) e outro para a modalidade PROEJA (Educação de Jovens e Adultos), resolveu-se desconsiderar estes dois Processos Seletivos pelo fato de estarem direcionados a um público diferente dos demais. Sendo assim, foi preparada a consulta SQL² para extrair em formato CSV³ os dados de 4 processos seletivos, com inscrições que aconteceram entre novembro de 2012 e janeiro de 2014.

Os campos extraídos do banco de dados estão descritos no Quadro 1:

QUADRO 1 – Campos extraídos do banco de dados

Nome	Descrição
unid_curso	Nome do curso precedido pelo nome da unidade do IFMG onde ele é ofertado
sexo	Sexo do candidato no cadastro de pessoa do sistema
cor_raca	Cor/Raça do candidato no cadastro de pessoa do sistema

² Structured Query Language

³ Comma-separated values

dt_inscricao	Data em que a inscrição foi efetuada
dt_nascimento	Data de nascimento do candidato no cadastro de pessoa do sistema
status_inscricao	O status da inscrição, com os valores Pendente, Cancelada ou Confirmada
status_opcao	A situação do candidato no Processo Seletivo, que vai sendo atualizada ao longo do mesmo.
reserva_vagas_inscrito	Se o candidato selecionou alguma opção de cota, mostra essa opção (EP-Cor-Renda, EP-Cor, EP-Renda ou EP). Se não, é preenchido com AC (Ampla Concorrência)
reserva_vagas_chamado	Campo preenchido com o valor Null no momento da inscrição. Após a chamada dos aprovados, ele é atualizado com a opção de cota (EP-Cor-Renda, EP-Cor, EP-Renda ou EP) ou com AC.
q1	1ª questão do questionário socioeconômico: “Qual o curso para o qual está se inscrevendo?”
q2	2ª questão do questionário socioeconômico: “Qual o seu sexo?”
q3	3ª questão do questionário socioeconômico: “Qual a sua cor/raça?”
q4	4ª questão do questionário socioeconômico: “Qual a sua idade?”
q10	10ª questão do questionário socioeconômico: “Quantos membros de sua família moram em sua casa (incluindo você)?”
q15	15ª questão do questionário socioeconômico: “Qual a renda total de sua família? (Em salário-mínimo)”
q19	19ª questão do questionário socioeconômico: “Antes de se matricular no IFMG, até o 5º ano (4ª série), você estudou?”
q20	20ª questão do questionário socioeconômico: “Antes de se matricular no IFMG, até o 9º ano (8ª série), você estudou?”
q21	21ª questão do questionário socioeconômico: “Antes de se matricular no IFMG, até o 3º ano ensino médio, você estudou?”

Fonte: Elaborado pelo autor

Observou-se que alguns campos estavam duplicados no cadastro de pessoa, gerado pelo formulário de inscrição e no questionário socioeconômico. Os dados mais confiáveis são os do cadastro de pessoa, pois estes dados poderiam ser conferidos pelo candidato durante o Processo Seletivo, através do Comprovante de Inscrição, e no caso dos alunos que chegaram a se matricular, todos os dados foram conferidos pelo setor de Registro e Controle Acadêmico no momento da matrícula. Os dados duplicados do

questionário socioeconômico foram também extraídos, com o objetivo de avaliar a sua similaridade com os dados do cadastro de pessoa, como forma de validação do preenchimento do questionário socioeconômico.

Foram extraídos 36.542 registros. Após a exclusão das inscrições que não foram confirmadas e de inscrições que estavam duplicadas, chegou-se a um total de 19.368 registros a serem pesquisados.

A partir dos dados extraídos, foram gerados novos campos, conforme o Quadro 2, para obter melhores resultados na mineração.

QUADRO 2 – Novos campos gerados a partir dos campos extraídos do banco de dados

Nome	Descrição
unidade	O nome da unidade do IFMG, obtido pelo campo <code>unid_curso</code> que foi extraído do banco de dados
curso	O nome da curso, obtido pelo campo <code>unid_curso</code> que foi extraído do banco de dados
modalidade	Preenchido com algum dos seguintes valores: Superior-Bacharelado, Superior-Licenciatura, Superior-Tecnologia, Técnico-Integrado ou Técnico-Subsequente
nível	Preenchido com Superior ou Técnico
idade	Calculada com os valores dos campos <code>dt_inscricao</code> e <code>dt_nascimento</code>
faixa_idade	Com o resultado do campo <code>idade</code> , criadas as mesmas faixas das respostas do questionário socioeconômico
matriculado	Armazena o valor “S” caso o candidato tenha sido matriculado, que é uma das opções do campo <code>status_opcao</code>
inscrito_em_cota	Armazena o valor “S” caso o candidato tenha sido inscrito em alguma das opções de cota, obtido através do campo <code>reserva_vagas_inscrito</code>
chamado_para_cota	Armazena o valor “S” caso o candidato tenha sido chamado a se matricular em alguma das opções de cota, obtido através do campo <code>reserva_vagas_chamado</code>
inscrito_cota_cor	Através do conteúdo do campo <code>reserva_vagas_inscrito</code> , mostra se o candidato se inscreveu para uma opção de cota que contém o critério de cor
inscrito_cota_renda	Através do conteúdo do campo <code>reserva_vagas_inscrito</code> , mostra se o candidato se inscreveu para uma opção de cota que contém o critério de renda
sempre_escola_publica_qs	Para os candidatos que se inscreveram para um curso de nível superior, calculado pelo conteúdo do campo <code>q21</code> . Para os

	candidatos que se inscreveram para um curso de nível técnico, calculado pelo conteúdo dos campos q19 e q20.
cor_ppi_qs	Através do campo q3, armazena “S” se o candidato tiver se declarado preto, pardo ou indígena
renda_menos_1.5_qs	Os valores dos campos q15 e q10 possibilitaram o cálculo da renda familiar <i>per capita</i> do candidato para verificar se ela é ou não inferior a 1,5 salário-mínimo. O campo foi preenchido com “D” onde não foi possível fazer o cálculo (quando o candidato declarou ter renda de mais de 8 salários-mínimos e mais de 5 membros da família residindo na mesma casa) e com “S” ou “N” nas demais situações.

Fonte: Elaborado pelo autor

A seguir, foi feita uma verificação de dados não preenchidos ou inconsistentes, sendo possível perceber que algumas datas de nascimento estavam preenchidas com valores improváveis. Foi definido que onde as idades calculadas fossem abaixo de 13 anos ou acima de 70 anos a faixa de idade seria preenchida com o valor “Desconhecida”.

Para a maioria dos testes, foi utilizado o arquivo com os dados dos 19.368 candidatos inscritos. Para a verificação dos candidatos matriculados em cotas, foi gerado outro arquivo, filtrando apenas os 4.579 matriculados no período avaliado.

3.2 Mineração dos dados

Para efetuar a mineração dos dados, foi utilizado o software Weka (Waikato Environment for Knowledge Analysis) na versão 3.6. O código desse sistema começou a ser desenvolvido em 1993, na Universidade de Weikato, na Nova Zelândia. Construído com a linguagem de programação Java, aceita várias extensões de arquivos como entrada, entre elas a extensão csv.

O algoritmo utilizado foi Apriori, apresentado por Agrawal e Srikant (1994). Este algoritmo de associação foi citado por Wu et al. (2008) como um dos 10 algoritmos de *data mining* mais utilizados no mundo.

3.3 Validação dos dados do questionário socioeconômico

Como a pesquisa propõe a utilização de dados do questionário socioeconômico, julgou-se adequado fazer um teste sobre a confiabilidade dos mesmos, já que estes dados são preenchidos pelo candidato no momento da inscrição e não são conferidos ou alterados posteriormente. Respostas incorretas no questionário podem indicar falta de interesse do

candidato no seu preenchimento e conseqüentemente pouca confiabilidade nesses dados. Este teste foi feito através da comparação de quatro atributos do questionário socioeconômico que estão repetidos em outro local do sistema: sexo, faixa de idade, curso e cor/raça.

No ERP usado na Instituição existe um cadastro que é preenchido quando uma nova pessoa é inserida no sistema, seja aluno, professor, funcionário ou candidato. No caso dos candidatos dos Processos Seletivos, esse cadastro é preenchido no momento da inscrição. Entre os dados exigidos nesse momento estão o sexo e a data de nascimento, que permite o cálculo da faixa de idade. O curso para o qual o candidato se inscreve é preenchido no formulário de inscrição. Portanto, para estes três atributos foi possível comparar os dados do cadastro de pessoa e do formulário de inscrição para todos os 19.368 inscritos com os dados do questionário socioeconômico.

O atributo cor/raça do cadastro de pessoa não é preenchido no momento da inscrição, mas é solicitado quando o candidato efetua a matrícula. Por isso o teste descrito para este atributo foi feito com os dados dos 4.679 candidatos que foram matriculados.

Sendo assim, os atributos que se repetem no questionário e em outro cadastro do sistema foram comparados, e os resultados são mostrados na Tabela 2.

TABELA 2 – Percentual de similaridade entre campos informados em diferentes locais do sistema

Campo comparado	Similaridade entre os campos pesquisados			
	Superior	Técnico Subsequente	Técnico Integrado	Geral
modalidade	90%	90%	93%	91%
sexo	99%	98%	99%	99%
faixa_idade	94%	94%	93%	94%
cor_raca	90%	91%	89%	90%

Fonte: Elaborada pelo autor

Pelos testes feitos, conclui-se que há um alto nível de confiabilidade nas respostas do questionário socioeconômico, já que similaridade entre estas respostas e a informação que foi conferida posteriormente durante o processo seletivo ou no momento da matrícula variou de 89% a 99%.

4 Apresentação dos Resultados

Neste capítulo serão apresentados os resultados para os testes, que foram realizados de acordo com os objetivos da pesquisa. Devido às diferenças no público que concorre a vagas no ensino superior, ensino técnico subsequente e ensino técnico integrado, optou-se por apresentar os dados separados por tipo de curso, além de apresentar também os dados com os totais, sem a distinção pelo tipo de curso.

4.1. Ocupação das vagas reservadas para cotas

Para verificar se as vagas reservadas dentro do período estudado foram efetivamente ocupadas, foi feito o teste com os atributos `reserva_vagas_inscrito` e `reserva_vagas_chamado`.

4.1.1 Todos os cursos

Os primeiros testes foram feitos com os dados de todos os 19.368 inscritos e com os 4.579 matriculados. A Tabela 3 traz o resumo do teste para verificar o 1º critério, ou seja, o estudo em escola pública no nível anterior ao pretendido no processo seletivo:

TABELA 3 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública

1º critério	Todos que estudaram em escola pública (Grupo II)
Número total de vagas ofertadas	5.231
Percentual de vagas reservadas	50,2%
Número de vagas reservadas	2.629
Percentual de inscritos para as vagas reservadas	40,6%
Número de inscritos para as vagas reservadas	7.878
Relação de inscritos por vaga reservada	3,0
Relação de inscritos por vaga – ampla concorrência	4,4
Percentual de matriculados nas vagas reservadas	45,1%
Número de matriculados nas vagas reservadas	2.065
Vagas reservadas que não foram ocupadas	564
Percentual de ocupação das vagas reservadas	78,5%

Fonte: Elaborada pelo autor

Qualquer candidato que tenha estudado integralmente em escola pública no nível anterior ao que está concorrendo agora, tem o direito de se inscrever em alguma das 4 opções de cotas. Sendo assim, a Tabela 3 mostra resultados gerais, sem especificar o tipo de cota. Por este resultado percebe-se que, de todas as vagas reservadas, 78,5% foram ocupadas. 564 vagas que inicialmente seriam reservadas para alguma opção de cota acabaram sendo ocupadas pela ampla concorrência. Outro ponto importante é que apenas 40,6% dos candidatos decidiram concorrer a alguma das vagas reservadas, bem abaixo da reserva de vagas de 50,2%.

O critério seguinte leva em consideração a renda familiar *per capita*, que deve ser igual ou inferior a 1,5 salário-mínimo. Nos testes feitos a seguir, foram agrupadas as duas categorias que levam em consideração a remuneração (EP-Renda-Cor e EP-Renda) e as outras duas que não consideram a renda familiar (EP-Cor e EP). A Tabela 4 mostra o resumo dos resultados desse teste.

TABELA 4 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública, agrupados pela renda familiar *per capita*

2º critério Todos que estudaram em escola pública	e tem renda familiar inferior a 1,5 salário- mínimo (Grupo III)	e tem renda familiar superior a 1,5 salário-mínimo (Grupo IV)
Percentual de vagas reservadas	25,6%	24,6%
Número de vagas reservadas	1.342	1.287
Percentual de inscritos para as vagas reservadas	25,4%	15,2%
Número de inscritos para as vagas reservadas	4.932	2.946
Relação de inscritos por vaga reservada	3,7	2,3
Percentual de matriculados nas vagas reservadas	24,6%	20,5%
Número de matriculados nas vagas reservadas	1.126	939
Vagas reservadas que não foram ocupadas	216	348
Percentual de ocupação das vagas reservadas	83,9%	73,0%

Fonte: Elaborada pelo autor

Pelos resultados percebe-se que nas cotas com critério econômico a ocupação das vagas reservadas chegou a 83,9%, sendo muito mais alta do que nas demais opções. O percentual de inscritos para quem estudou em escola pública e tem renda familiar acima de 1,5 salário-mínimo (Grupo IV) foi de apenas 15,2%, sendo que 24,6% das vagas estavam reservadas para este público.

O 3º critério leva em consideração a autodeclaração de cor como preto, pardo ou indígena. Na Tabela 5 estão os resultados da pesquisa com o agrupamento das cotas que levam em consideração a cor (EP-Cor-Renda e EP-Cor) e as cotas que não a levam em consideração (EP-Renda e EP).

TABELA 5 – Inscrições e matrículas nas vagas reservadas para alunos que estudaram em escola pública, agrupados pela autodeclaração de cor

3º critério	e se autodeclararam pretos, pardos ou indígenas (Grupos V e VII)	e não se autodeclararam pretos, pardos ou indígenas (Grupos VI e VIII)
Todos que estudaram em escola pública		
Percentual de vagas reservadas	31,0%	19,2%
Número de vagas reservadas	1.622	1.007
Percentual de inscritos para as vagas reservadas	28,2%	12,4%
Número de inscritos para as vagas reservadas	5.477	2.401
Relação de inscritos por vaga reservada	3,4	2,4
Percentual de matriculados nas vagas reservadas	29,5%	15,6%
Número de matriculados nas vagas reservadas	1.353	712
Vagas reservadas que não foram ocupadas	269	295
Percentual de ocupação das vagas reservadas	83,4%	70,7%

Fonte: Elaborada pelo autor

Por causa do arredondamento, o percentual de vagas para os Grupos V e VII foi de 31,0%, bem acima dos 26,6% que seriam ofertadas para estes grupos se o percentual de 53,6% de pretos, pardos e indígenas fosse aplicado no número total de vagas reservadas. De acordo com os resultados, as vagas reservadas para os critérios relacionados a cor estão mais próximos de ser ocupadas. Já em relação às cotas que não levam a cor em consideração, tanto as inscrições quanto as matrículas foram bem abaixo do que seria esperado. Enquanto 28,2% dos candidatos se inscreveram para concorrer às vagas reservadas aos pretos, pardos e indígenas, apenas 12,4% se inscreveram para as cotas que não levavam em conta a cor.

Para compreender melhor essa diferença, o teste foi realizado para cada tipo de cota. O resumo dos resultados está na Tabela 6.

TABELA 6 – Inscrições e matrículas nas vagas reservadas para cada tipo de cota

3º critério	inferior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo V)	inferior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VI)	superior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo VII)	superior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VIII)
Todos que estudaram em escola pública e têm renda familiar				
Percentual de vagas reservadas	15,8%	9,8%	15,1%	9,5%
Número de vagas reservadas	830	512	792	495
Percentual de inscritos para as vagas reservadas	18,5%	6,9%	9,7%	5,5%
Número de inscritos para as vagas reservadas	3.590	1.342	1.887	1.059
Relação de inscritos por vaga reservada	4,3	2,6	2,4	2,1
Percentual de matriculados nas vagas reservadas	16,1%	8,5%	13,5%	7,0%
Número de matriculados nas vagas reservadas	737	389	616	323
Vagas reservadas que não foram ocupadas	93	123	176	172
Percentual de ocupação das vagas reservadas	88,8%	76,0%	77,8%	65,3%

Fonte: Elaborada pelo autor

Pode-se perceber que no Grupo V, a ocupação foi quase total, tendo sobrado apenas 93 vagas, de um total de 830 ofertadas. Nos demais grupos, com menos vagas oferecidas, o número de vagas que foram não ocupadas foi bem maior. Percebe-se também que o Grupo V foi o único onde o percentual de inscrições em cotas foi superior ao percentual de vagas reservadas para o grupo. O Grupo VIII teve o menor percentual de inscrições, o menor percentual de matriculados, e conseqüentemente foi o que mais cedeu vagas para a ampla concorrência, pois das 495 vagas reservadas, 172 não foram ocupadas, o que representa 34,7%.

A seguir, os mesmos testes foram realizados para 3 subconjuntos: os candidatos a cursos superiores, técnicos subsequentes e técnicos integrados.

4.1.2 Cursos Superiores

Para esta parte da pesquisa, foi utilizado um arquivo com os dados dos 9.190 candidatos inscritos para concorrer a vagas de cursos superiores no período estudado. A Tabela 7 apresenta o resumo do teste para verificar o 1º critério:

TABELA 7 – Inscrições e matrículas nas vagas reservadas em cursos superiores para alunos que estudaram em escola pública

1º critério	Todos que estudaram em escola pública (Grupo II)
Número total de vagas ofertadas	1.311
Percentual de vagas reservadas	50,3%
Número de vagas reservadas	659
Percentual de inscritos para as vagas reservadas	42,4%
Número de inscritos para as vagas reservadas	3.896
Relação de inscritos por vaga reservada	5,9
Relação de inscritos por vaga – ampla concorrência	8,1
Percentual de matriculados nas vagas reservadas	47,7%
Número de matriculados nas vagas reservadas	572
Vagas reservadas que não foram ocupadas	87
Percentual de ocupação das vagas reservadas	86,8%

Fonte: Elaborada pelo autor

A participação no sistema de cotas nos cursos superiores foi maior do que a média geral, conclusão evidenciada pelo maior percentual de inscritos, matriculados e de ocupação das vagas reservadas.

Os testes realizados com o agrupamento dos inscritos em cotas pela renda familiar *per capita* trouxeram os resultados exibidos na Tabela 8:

TABELA 8 – Inscrições e matrículas nas vagas reservadas para alunos de cursos superiores que estudaram em escola pública, agrupados pela renda familiar *per capita*

2º critério Todos que estudaram em escola pública	e tem renda familiar inferior a 1,5 salário-mínimo (Grupo III)	e tem renda familiar superior a 1,5 salário-mínimo (Grupo IV)
Percentual de vagas reservadas	25,7%	24,6%
Número de vagas reservadas	337	322
Percentual de inscritos para as vagas reservadas	26,1%	16,3%
Número de inscritos para as vagas reservadas	2.401	1.495
Relação de inscritos por vaga reservada	7,1	4,6
Percentual de matriculados nas vagas reservadas	25,3%	22,4%
Número de matriculados nas vagas reservadas	303	269
Vagas reservadas que não foram ocupadas	34	53
Percentual de ocupação das vagas reservadas	89,9%	83,5%

Fonte: Elaborada pelo autor

Os resultados mostram que houve uma ocupação mais alta das vagas reservadas a candidatos com menor renda familiar. Assim como na média geral, as inscrições para quem estudou em escola pública e tem renda familiar acima de 1,5 salário-mínimo (Grupo IV) foram baixas, um percentual de 16,3% de inscritos para uma reserva de 24,6% das vagas.

O próximo teste trouxe o agrupamento pela autodeclaração de cor, cujos resultados são exibidos na Tabela 9:

TABELA 9 – Inscrições e matrículas nas vagas reservadas para alunos de cursos superiores que estudaram em escola pública, agrupados pela autodeclaração de cor

3º critério Todos que estudaram em escola pública	e se autodeclararam pretos, pardos ou indígenas (Grupos V e VII)	e não se autodeclararam pretos, pardos ou indígenas (Grupos VI e VIII)
Percentual de vagas reservadas	31,5%	18,8%
Número de vagas reservadas	413	246
Percentual de inscritos para as vagas reservadas	28,0%	14,4%
Número de inscritos para as vagas reservadas	2.570	1.326

Relação de inscritos por vaga reservada	6,2	5,4
Percentual de matriculados nas vagas reservadas	30,4%	17,3%
Número de matriculados nas vagas reservadas	364	208
Vagas reservadas que não foram ocupadas	49	38
Percentual de ocupação das vagas reservadas	88,1%	84,6%

Fonte: Elaborada pelo autor

Houve uma ocupação ligeiramente maior entre as vagas reservadas para os candidatos que se autodeclararam pretos, pardos ou indígenas. Nos dois grupos o percentual de inscritos e matriculados ficou próximo do percentual de vagas reservadas.

A Tabela 10 mostra o resultado dos testes para cada um dos 4 tipos de cotas.

TABELA 10 – Inscrições e matrículas nas vagas reservadas para cursos superiores para cada tipo de cota

3º critério Todos que estudaram em escola pública e têm renda familiar	inferior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo V)	inferior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VI)	superior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo VII)	superior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VIII)
Percentual de vagas reservadas	15,8%	9,9%	15,7%	8,8%
Número de vagas reservadas	207	130	206	116
Percentual de inscritos para as vagas reservadas	17,9%	8,2%	10,0%	6,2%
Número de inscritos para as vagas reservadas	1.645	756	925	570
Relação de inscritos por vaga reservada	7,9	5,8	4,5	4,9
Percentual de matriculados nas vagas reservadas	16,2%	9,1%	14,2%	8,3%
Número de matriculados nas vagas reservadas	194	109	170	99
Vagas reservadas que não foram ocupadas	13	21	36	17
Percentual de ocupação das vagas reservadas	93,7%	83,8%	82,5%	85,3%

Fonte: Elaborada pelo autor

Pode-se perceber que no Grupo V, a ocupação foi quase total, chegando a 93,7% das vagas reservadas a este grupo. Nos demais grupos, com menos vagas oferecidas, o número de vagas não ocupadas foi maior, mas todos tiveram ocupação acima de 80%. Assim como na média geral, o Grupo V foi o único onde o percentual de inscrições em cotas foi superior ao percentual de vagas reservadas para o grupo. O Grupo VIII teve o menor percentual de inscrições, porém o Grupo VII foi o que teve a menor ocupação das vagas.

4.1.3 Cursos Técnicos Subsequentes

Os cursos técnicos subsequentes tiveram a menor procura entre os 3 conjuntos estudados. Foram 3.358 os candidatos que se inscreveram para as vagas dessa modalidade. A Tabela 11 mostra os resultados do teste realizado nestes dados para verificar o 1º critério.

TABELA 11 – Inscrições e matrículas nas vagas reservadas em cursos técnicos subsequentes para alunos que estudaram em escola pública

1º critério	Todos que estudaram em escola pública (Grupo II)
Número total de vagas ofertadas	1.562
Percentual de vagas reservadas	50,3%
Número de vagas reservadas	786
Percentual de inscritos para as vagas reservadas	38,5%
Número de inscritos para as vagas reservadas	1.291
Relação de inscritos por vaga reservada	1,6
Relação de inscritos por vaga – ampla concorrência	2,7
Percentual de matriculados nas vagas reservadas	39,9%
Número de matriculados nas vagas reservadas	473
Vagas reservadas que não foram ocupadas	313
Percentual de ocupação das vagas reservadas	60,2%

Fonte: Elaborada pelo autor

Nos cursos técnicos subsequentes a inscrição nas vagas reservadas foi menor do que a média geral. Esse número se refletiu nas matrículas para essas vagas que foram de apenas 39,9%, fazendo com que apenas 60,2% das vagas reservadas fossem realmente ocupadas por candidatos cotistas.

A avaliação do 2º critério mostrou uma diferença significativa entre os candidatos dos Grupos III e IV, como mostra a Tabela 12.

TABELA 12 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos subsequentes que estudaram em escola pública, agrupados pela renda familiar *per capita*

2º critério Todos que estudaram em escola pública	e tem renda familiar inferior a 1,5 salário-mínimo (Grupo III)	e tem renda familiar superior a 1,5 salário-mínimo (Grupo IV)
Percentual de vagas reservadas	25,9%	24,5%
Número de vagas reservadas	404	382
Percentual de inscritos para as vagas reservadas	25,4%	13,0%
Número de inscritos para as vagas reservadas	854	437
Relação de inscritos por vaga reservada	2,1	1,1
Percentual de matriculados nas vagas reservadas	21,6%	18,3%
Número de matriculados nas vagas reservadas	256	217
Vagas reservadas que não foram ocupadas	148	165
Percentual de ocupação das vagas reservadas	63,3%	56,8%

Fonte: Elaborada pelo autor

Considerando apenas os candidatos a vagas no ensino técnico subsequente, o percentual de ocupação foi baixo nos dois grupos, sendo ainda menor no Grupo IV. Apenas 56,8% das vagas reservadas a candidatos com a renda familiar *per capita* acima de 1,5 salário-mínimo foram ocupadas. O resultado condiz com o baixo percentual de inscritos (13,0%) nas vagas reservadas a este grupo. No Grupo III o percentual de inscritos foi próximo ao percentual de vagas reservadas, porém o percentual de matriculados ainda esteve bem abaixo do que seria possível.

O teste seguinte agrupou os candidatos pela autodeclaração de cor, encontrando diferenças muito acentuadas entre os dois grupos:

TABELA 13 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos subsequentes que estudaram em escola pública, agrupados pela autodeclaração de cor

3º critério Todos que estudaram em escola pública	e se autodeclararam pretos, pardos ou indígenas (Grupos V e VII)	e não se autodeclararam pretos, pardos ou indígenas (Grupos VI e VIII)
Percentual de vagas reservadas	31,2%	19,0%
Número de vagas reservadas	488	298
Percentual de inscritos para as vagas reservadas	29,6%	8,8%
Número de inscritos para as vagas reservadas	994	297
Relação de inscritos por vaga reservada	2,0	0,9
Percentual de matriculados nas vagas reservadas	28,2%	11,7%
Número de matriculados nas vagas reservadas	334	139
Vagas reservadas que não foram ocupadas	154	159
Percentual de ocupação das vagas reservadas	68,4%	46,7%

Fonte: Elaborada pelo autor

Os dados da Tabela 13 mostram uma grande discrepância entre os dois grupos, com números muito maiores entre os candidatos que se autodeclararam pretos, pardos e indígenas. A baixa procura pelas vagas dos Grupos VI e VIII é algo que chama a atenção. Houve menos inscritos do que vagas reservadas para este público, além do percentual de inscritos ter sido menos da metade do percentual de vagas reservadas. Outro dado que se destaca é o percentual de ocupação de somente 46,7% das vagas, de forma que houve mais vagas cedidas para a ampla concorrência do que ocupadas pelos candidatos desses grupos.

O teste mais detalhado, com cada uma das 4 opções de cotas, está descrito na Tabela 14:

TABELA 14 – Inscrições e matrículas nas vagas reservadas para cursos técnicos subsequentes para cada tipo de cota

3º critério	inferior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo V)	inferior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VI)	superior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo VII)	superior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VIII)
Todos que estudaram em escola pública e têm renda familiar				
Percentual de vagas reservadas	16,3%	9,6%	15,0%	9,5%
Número de vagas reservadas	254	150	234	148
Percentual de inscritos para as vagas reservadas	20,1%	5,3%	9,4%	3,5%
Número de inscritos para as vagas reservadas	676	178	318	119
Relação de inscritos por vaga reservada	2,7	1,2	1,4	0,8
Percentual de matriculados nas vagas reservadas	14,9%	6,7%	13,2%	5,0%
Número de matriculados nas vagas reservadas	177	79	157	60
Vagas reservadas que não foram ocupadas	77	71	77	88
Percentual de ocupação das vagas reservadas	69,7%	52,7%	67,1%	40,5%

Fonte: Elaborada pelo autor

No Grupo VIII, onde o único critério exigido era o fato de ter estudado em escola pública, o percentual de inscrições foi extremamente baixo, tendo sido o único grupo onde o número de inscritos foi inferior ao número de vagas. Isso se refletiu no percentual muito baixo de ocupação das vagas, que foi de 40,5%.

4.1.4 Cursos Técnicos Integrados

Para avaliar os candidatos a cursos técnicos integrados, foi utilizado um arquivo com os dados dos 6.820 candidatos inscritos para concorrer a vagas nesta modalidade. A tabela 15 apresenta o resumo do teste para verificar o 1º critério:

TABELA 15 – Inscrições e matrículas nas vagas reservadas em cursos técnicos integrados para alunos que estudaram em escola pública

1º critério	Todos que estudaram em escola pública (Grupo II)
Número total de vagas ofertadas	2.358
Percentual de vagas reservadas	50,2%
Número de vagas reservadas	1.184
Percentual de inscritos para as vagas reservadas	39,5%
Número de inscritos para as vagas reservadas	2.691
Relação de inscritos por vaga reservada	2,3
Relação de inscritos por vaga – ampla concorrência	3,5
Percentual de matriculados nas vagas reservadas	46,4%
Número de matriculados nas vagas reservadas	1.020
Vagas reservadas que não foram ocupadas	164
Percentual de ocupação das vagas reservadas	86,1%

Fonte: Elaborada pelo autor

Nos cursos técnicos integrados, o percentual de inscritos e de matriculados em cotas foi superior à média geral, mas ainda abaixo das possibilidades. A ocupação das vagas reservadas foi de 86,1%, que é um percentual alto, mas que ainda pode ser aumentado.

A avaliação dos dados de acordo com o 2º critério mostrou grandes diferenças entre os dois grupos, como mostra a Tabela 16.

TABELA 16 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos integrados que estudaram em escola pública, agrupados pela renda familiar *per capita*

2º critério Todos que estudaram em escola pública	e tem renda familiar inferior a 1,5 salário-mínimo (Grupo III)	e tem renda familiar superior a 1,5 salário-mínimo (Grupo IV)
Percentual de vagas reservadas	25,5%	24,7%
Número de vagas reservadas	601	583
Percentual de inscritos para as vagas reservadas	24,6%	14,9%
Número de inscritos para as vagas reservadas	1.677	1.014
Relação de inscritos por vaga reservada	2,8	1,7
Percentual de matriculados nas vagas reservadas	25,8%	20,6%
Número de matriculados nas vagas reservadas	567	453
Vagas reservadas que não foram ocupadas	34	130
Percentual de ocupação das vagas reservadas	94,3%	77,7%

Fonte: Elaborada pelo autor

O percentual de inscritos nas vagas do Grupo III esteve bem próximo ao percentual de vagas reservadas para o grupo, o que se refletiu no alto percentual de ocupação das vagas, que chegou a 94,3%. Por outro lado, no Grupo IV, os percentuais de inscrições e de matrículas foram muito menores, fazendo com que a ocupação das vagas ficasse em 77,7%.

O teste seguinte avaliou os mesmos dados, agrupando os candidatos pela autodeclaração de cor. Os resultados que estão descritos na Tabela 17 mostram diferenças significativas entre os dois grupos.

TABELA 17 – Inscrições e matrículas nas vagas reservadas para alunos de cursos técnicos integrados que estudaram em escola pública, agrupados pela autodeclaração de cor

3º critério Todos que estudaram em escola pública	e se autodeclararam pretos, pardos ou indígenas (Grupos V e VII)	e não se autodeclararam pretos, pardos ou indígenas (Grupos VI e VIII)
Percentual de vagas reservadas	30,6%	19,6%
Número de vagas reservadas	721	463
Percentual de inscritos para as vagas reservadas	28,0%	11,4%

Número de inscritos para as vagas reservadas	1.913	778
Relação de inscritos por vaga reservada	2,7	1,7
Percentual de matriculados nas vagas reservadas	29,8%	16,6%
Número de matriculados nas vagas reservadas	655	365
Vagas reservadas que não foram ocupadas	66	98
Percentual de ocupação das vagas reservadas	90,8%	78,8%

Fonte: Elaborada pelo autor

Pelos resultados precebe-se que nos Grupos V e VII as inscrições e matrículas estiveram próximas do que seria possível, fazendo com que a ocupação das vagas chegasse a 90,8%. Nos Grupos VI e VIII a procura pelas vagas foi muito menor, assim a ocupação das vagas também foi consideravelmente menor que nos outros grupos.

O teste seguinte, relatado na Tabela 18, mostra os dados de cada tipo de cota.

TABELA 18 – Inscrições e matrículas nas vagas reservadas para cursos técnicos integrados para cada tipo de cota

3º critério	inferior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo V)	inferior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VI)	superior a 1,5 salário-mínimo e se autodeclararam pretos, pardos ou indígenas (Grupo VII)	superior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas (Grupo VIII)
Todos que estudaram em escola pública e têm renda familiar				
Percentual de vagas reservadas	15,6%	9,8%	15,0%	9,8%
Número de vagas reservadas	369	232	352	231
Percentual de inscritos para as vagas reservadas	18,6%	6,0%	9,4%	5,4%
Número de inscritos para as vagas reservadas	1.269	408	644	370
Relação de inscritos por vaga reservada	3,4	1,8	1,8	1,6
Percentual de matriculados nas vagas reservadas	16,7%	9,2%	13,2%	7,5%
Número de matriculados nas vagas reservadas	366	201	289	164

Vagas reservadas que não foram ocupadas	3	31	63	67
Percentual de ocupação das vagas reservadas	99,2%	86,6%	82,1%	71,0%

Fonte: Elaborada pelo autor

O resultado que mais se destaca na Tabela 18 é a ocupação de 99,2% das vagas reservadas ao Grupo V dos cursos técnicos integrados. Esse é o maior percentual de ocupação de todos os grupos em todos os níveis e modalidades pesquisadas. No Grupo VIII a ocupação foi muito menor que nos demais grupos, o que é coerente com o baixo percentual de inscrições nas vagas reservadas a este grupo.

4.2. Candidatos que poderiam concorrer às vagas, mas não o fizeram

Para descobrir até que ponto os candidatos fizeram uso do seu direito à inscrição em cotas, foram testados os atributos que permitem comparar as inscrições em cotas com os dados do candidato que são requisitados para cada tipo de vaga reservada. Para os testes que levavam a renda em consideração, foram excluídos os candidatos para os quais não foi possível calcular a renda familiar *per capita*.

4.2.1 Todos os cursos

Primeiramente a pesquisa foi feita sobre o arquivo com os dados de todos os 19.368 inscritos, utilizando os atributos *inscrito_em_cota* e *sempre_escola_publica_qs*. A Tabela 19 traz o resultado desse teste.

TABELA 19 – Candidatos que poderiam ter se inscrito em alguma opção de cota

Requisitos	Total de candidatos que possuíam os requisitos	Número de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos	Percentual de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos
Escola pública	15.525	7.554	48,7%

Fonte: Elaborada pelo autor

Esse resultado mostra que entre todos os inscritos, 15.525 poderiam ter se inscrito em alguma das opções de cotas, uma vez que estudaram em escola pública. Desses, apenas 7.554 se inscreveram efetivamente, o que significa apenas 48,7% do público potencial de cotistas.

Na Tabela 20 segue um resumo dos resultados agrupando os candidatos em cotas que levavam em consideração os critérios racial e econômico. Foram utilizados os seguintes atributos: `sempre_escola_publica_qs`, `renda_menos_1.5_qs`, `cor_ppi_qs`, `inscrito_em_cota`, `inscrito_cota_renda` e `inscrito_cota_cor`.

TABELA 20 – Candidatos que poderiam ter se inscrito em cotas com critérios raciais e econômicos

Requisitos	Total de candidatos que possuíam os requisitos	de que possuíam os requisitos	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Escola pública e cor	10.131		5.098	50,3%
Escola pública e renda	14.290		4.648	32,5%

Fonte: Elaborada pelo autor

Entre todos os inscritos, 10.131 tinham os critérios para se inscrever no Grupo V ou no Grupo VII, ou seja, nas vagas reservadas a alunos de escola pública que se declararam pretos, pardos ou indígenas. Desses, apenas 50,3% realmente fizeram a inscrição em algum desses grupos. Em relação às cotas para alunos de escola pública com renda familiar *per capita* de até 1,5 salário-mínimo (Grupos V e VI), dos 14.290 que tinham os requisitos, somente 32,5% concorreram a uma vaga reservada em algum desses dois grupos.

O resultado do teste para cada um dos quatro tipos de cotas está descrito na Tabela 21. Foram utilizados os seguintes atributos: `sempre_escola_publica_qs`, `renda_menos_1.5_qs`, `cor_ppi_qs` e `reserva_vagas_inscrito`.

TABELA 21 – Candidatos que poderiam ter se inscrito em cada grupo de cotas

Grupo	Total de candidatos que possuíam os requisitos esperado para cada grupo	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Grupo V	9.448	3.294	34,9%
Grupo VI	4.842	1.120	23,1%
Grupo VII	669	258	39,0%

Grupo VIII	543	154	28,3%
Total	15.502	4.826	31,1%

Fonte: Elaborada pelo autor

A Tabela 21 mostra o total de candidatos que teria se inscrito em cada grupo se todos os candidatos que possuem o direito tivessem se inscrito na opção mais correspondente com o seu perfil. Sendo assim, o total de 543 candidatos do Grupo VIII não se refere a todos os alunos que estudaram em escola pública, mas a todos os alunos que estudaram em escola pública, têm renda familiar superior a 1,5 salário-mínimo e não se autodeclararam pretos, pardos ou indígenas. A soma de inscrições nos 4 grupos, que é de 4.826, não se refere ao total de inscritos em cotas, que foi de 7.878, mas ao total de candidatos inscritos na cota mais adequada ao seu perfil, de acordo com os dados informados. Isso significa que 3.052 candidatos (38,7% dos inscritos em cotas), escolheram um grupo diferente do que se esperaria, como por exemplo, candidatos que declararam ser pretos, pardos ou indígenas no questionário socioeconômico, mas concorreram a uma opção de cota que não leva a cor em consideração.

4.2.2 Cursos Superiores

Considerando apenas os 9.190 candidatos a vagas no ensino superior, foram testados os atributos `inscrito_em_cota` e `sempre_escola_publica_qs`, como mostrado na Tabela 22.

TABELA 22 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em alguma opção de cota

Requisitos	Total de candidatos que possuíam os requisitos	Número de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos	Percentual de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos
Escola pública	7.645	3.775	49,4%

Fonte: Elaborada pelo autor

Entre os 7.645 candidatos que poderiam ter se inscrito em alguma das opções de cotas, por terem estudado em escola pública, menos de 50% fizeram uso desse direito.

Em seguida, foram avaliados os requisitos referentes a renda e a cor dos candidatos. Os atributos usados foram `sempre_escola_publica_qs`, `renda_menos_1.5_qs`,

cor_ppi_qs, inscrito_em_cota, inscrito_cota_renda e inscrito_cota_cor. Os resultados estão descritos na Tabela 23.

TABELA 23 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cotas com critérios raciais e econômicos

Requisitos	Total de candidatos que possuíam os requisitos	de que os possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Escola pública e cor	4.768	2.412	50,6%
Escola pública e renda	6.774	2.271	33,5%

Fonte: Elaborada pelo autor

Percebe-se que para os candidatos a vagas no nível superior, os percentuais de inscrição em relação aos que possuem os requisitos estão muito próximos da média geral. Aproximadamente a metade dos candidatos que possuem os requisitos de escola pública e cor, e apenas um terço dos que possuem os requisitos de escola pública e renda é que se inscreveram efetivamente.

Para avaliar as inscrições em cada um dos quatro tipos de cotas, foi feito o teste cujo resumo está na Tabela 24.

TABELA 24 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cada grupo de cotas

Grupo	Total de candidatos que possuíam os requisitos esperado para cada grupo	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Grupo V	4.288	1.506	35,1%
Grupo VI	2.486	649	26,1%
Grupo VII	470	174	37,2%
Grupo VIII	387	112	29,0%
Total	7.631	2.441	32,0%

Fonte: Elaborada pelo autor

A Tabela 24 mostra o total de candidatos a vagas no ensino superior que teria se inscrito em cada grupo se todos tivessem concorrido à opção mais correspondente com o

seu perfil. Nota-se que os percentuais de candidatos que efetivaram as inscrições nesses grupos foram baixos, sendo de 32% no total de 7.631 candidatos que teriam todos os requisitos.

4.2.3 Cursos Técnicos Subsequentes

Para avaliar as inscrições às vagas no ensino técnico subsequente, foram testados os atributos inscrito_em_cota e sempre_escola_publica_qs nos registros dos 3.358 candidatos que concorreram às vagas desta modalidade. O resultado do teste é mostrado na Tabela 25.

TABELA 25 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em alguma opção de cota

Requisitos	Total de candidatos que possuíam os requisitos	Número de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos	Percentual de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos
Escola pública	3.103	1.252	40,3%

Fonte: Elaborada pelo autor

Entre os 3.103 candidatos que poderiam ter se inscrito em alguma das opções de cotas, apenas 40,3% concorreram a alguma das vagas reservadas. Este percentual é bem menor do que a média geral, de 48,7%.

Em seguida, foram avaliados os requisitos referentes à renda e a cor dos candidatos. Os atributos usados foram sempre_escola_publica_qs, renda_menos_1.5_qs, cor_ppi_qs, inscrito_em_cota, inscrito_cota_renda e inscrito_cota_cor. Os resultados estão descritos na Tabela 26.

TABELA 26 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cotas com critérios raciais e econômicos

Requisitos	Total de candidatos que possuíam os requisitos	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Escola pública e cor	2.155	926	43,0%
Escola pública e renda	2.927	813	27,8%

Fonte: Elaborada pelo autor

Nas situações exibidas na Tabela 26, os percentuais ficaram abaixo da média geral. Assim como na média geral, entre os candidatos que possuem os requisitos de escola pública e cor, a procura pelas cotas é maior do que entre os candidatos que possuem os requisitos de escola pública e renda.

Para verificar em cada um dos quatro tipos de cotas, a relação entre ter os requisitos e se inscrever para as vagas reservadas, foi feito o teste descrito na Tabela 27:

TABELA 27 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cada grupo de cotas

Grupo	Total de candidatos que possuíam os requisitos esperado para cada grupo	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Grupo V	2.053	622	30,3%
Grupo VI	874	144	16,5%
Grupo VII	100	28	28,0%
Grupo VIII	73	7	9,6%
Total	3.100	801	25,8%

Fonte: Elaborada pelo autor

A Tabela 27 mostra o total de candidatos a vagas no ensino técnico subsequente que teria se inscrito em cada grupo se todos tivessem concorrido à opção mais correspondente com o seu perfil. Nota-se que os percentuais de candidatos que efetivaram as inscrições nesses grupos foram muito mais baixos do que a média geral, principalmente no Grupo VIII, onde menos de 10% dos candidatos que tinham o perfil chegaram a se inscrever nas vagas reservadas.

4.2.4 Cursos Técnicos Integrados

Para avaliar as inscrições às vagas no ensino técnico integrado, foram testados os atributos `inscrito_em_cota` e `sempre_escola_publica_qs` nos registros dos 6.820 candidatos que concorreram às vagas desta modalidade. O resultado do teste é mostrado na Tabela 28.

TABELA 28 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em alguma opção de cota

Requisitos	Total de candidatos que possuíam os requisitos	Número de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos	Percentual de candidatos que se inscreveram nas vagas reservadas, que possuíam os requisitos
Escola pública	4.777	2.527	52,9%

Fonte: Elaborada pelo autor

Entre os 4.777 candidatos que poderiam ter se inscrito em alguma das opções de cotas, porque estudaram em escola pública em todo o ensino fundamental, 52,9% fizeram essa opção. Esta foi a única modalidade em que mais da metade dos candidatos que possuíam os requisitos fizeram uso do direito de concorrer às vagas reservadas.

Em seguida, foram avaliados os requisitos referentes à renda e a cor dos candidatos. Os atributos usados foram sempre_escola_publica_qs, renda_menos_1.5_qs, cor_ppi_qs, inscrito_em_cota, inscrito_cota_renda e inscrito_cota_cor. Os resultados estão descritos na Tabela 29.

TABELA 29 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cotas com critérios raciais e econômicos

Requisitos	Total de candidatos que possuíam os requisitos	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Escola pública e cor	3.208	1.760	54,9%
Escola pública e renda	4.589	1.564	34,1%

Fonte: Elaborada pelo autor

Nas situações exibidas na Tabela 29, os percentuais ficaram acima da média geral, com uma procura maior pelas cotas nos grupos onde os candidatos possuem os requisitos de escola pública e cor.

Para verificar em cada um dos quatro tipos de cotas, a relação entre ter os requisitos e se inscrever para as vagas reservadas, foi feito o teste descrito na Tabela 30:

TABELA 30 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cada grupo de cotas

Grupo	Total de candidatos que possuíam os requisitos esperado para cada grupo	Entre todos os que possuíam os requisitos, número de candidatos que se inscreveram nas vagas reservadas	Entre todos os que possuíam os requisitos, percentual de candidatos que se inscreveram nas vagas reservadas
Grupo V	3.107	1.166	37,5%
Grupo VI	1.482	327	22,1%
Grupo VII	99	56	56,6%
Grupo VIII	83	35	42,2%
Total	4.771	1.584	33,2%

Fonte: Elaborada pelo autor

A Tabela 30 mostra o total de candidatos a vagas no ensino técnico integrado que teria se inscrito em cada grupo se todos tivessem concorrido à opção mais correspondente com o seu perfil. Os percentuais são maiores do que a média geral, com exceção do Grupo VI. Nos Grupos VII e VIII, nota-se que os percentuais são muito superiores à média geral, mas como o número de candidatos é muito pequeno, essa diferença não produz muito impacto no percentual de inscrições da modalidade.

4.3 Candidatos inscritos com os requisitos necessários

Foram realizados testes nos dados dos candidatos inscritos nas vagas reservadas, comparando a opção de cota feita pelo candidato e o preenchimento do questionário socioeconômico. As inconsistências encontradas fornecem indícios de inscrições efetuadas sem os requisitos necessários.

4.3.1 Todos os cursos

O algoritmo *Apriori* foi executado com os atributos *reserva_vagas_inscrito*, *sempre_escola_publica_qs*, *renda_menos_1.5_qs* e *cor_ppi_qs* nos dados dos 7.878 candidatos inscritos para algum tipo de cota. Primeiramente foi avaliada a inscrição de uma maneira geral, sem especificar o tipo de cota, conforme a Tabela 31:

TABELA 31 – Candidatos inscritos com os requisitos

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo II	7.878	7.554	95,9%

Fonte: Elaborada pelo autor

O requisito comum a todas as opções de cota é ter estudado em escola pública (Grupo II). Seria esperado que 100% dos inscritos em cotas declarassem ter estudado sempre em escola pública, mas 4,1% deles declararam no questionário socioeconômico que estudaram em escola particular no nível anterior ao nível a que estavam concorrendo.

O teste seguinte foi feito com o agrupamento das cotas de cor e renda, cujos resultados estão descritos na Tabela 32. Nos grupos de cotas que levam a renda em consideração, foram excluídos os candidatos para os quais não foi possível calcular a renda familiar *per capita*.

TABELA 32 – Candidatos inscritos com os requisitos agrupados pela cor e pela renda familiar

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupos V e VII	5.477	5.098	93,1%
Grupos V e VI	4.922	4.648	94,4%

Fonte: Elaborada pelo autor

Nas cotas reservadas para estudantes de escola pública que se declararam pretos, pardos e indígenas (Grupos V e VII), 93,1% possuíam todos os requisitos, de acordo com o questionário socioeconômico. Nas cotas reservadas a quem sempre estudou em escola pública e tem renda familiar até 1,5 salário-mínimo *per capita* (Grupos V e VI) o percentual foi de 94,2%.

A Tabela 33 mostra o resultado do mesmo teste feito para cada uma das opções de cota.

TABELA 33 – Candidatos inscritos com os requisitos em cada grupo de cotas

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo V	3.590	3.294	91,8%
Grupo VI	1.342	1.243	92,6%
Grupo VII	1.887	1.744	92,4%
Grupo VIII	1.059	1.010	95,4%
Total	7.878	7.291	92,5%

Fonte: Elaborada pelo autor

Estes resultados indicam que a grande maioria dos candidatos que se inscreve nas cotas possui os requisitos. Porém, mesmo que o percentual não chame a atenção, os 587 candidatos que se inscreveram a vagas reservadas sem ter os requisitos no período estudado representam um número considerável.

4.3.2 Cursos Superiores

Considerando apenas os 3.896 candidatos a vagas do ensino superior que concorreram às vagas reservadas, foram realizados os testes com os atributos `reserva_vagas_inscrito`, `sempre_escola_publica_qs`, `renda_menos_1.5_qs` e `cor_ppi_qs`. A avaliação geral, sem especificar o tipo de cota, está na Tabela 34:

TABELA 34 – Candidatos a vagas no ensino superior que se inscreveram em cotas possuindo os requisitos

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo II	3.896	3.775	96,9%

Fonte: Elaborada pelo autor

Seria esperado que 100% dos inscritos em cotas declarassem ter estudado sempre em escola pública, mas 3,1% deles declararam ter estudado em escola particular no questionário socioeconômico.

O teste seguinte foi feito com o agrupamento das cotas de cor e renda, cujos resultados estão descritos na Tabela 35:

TABELA 35 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupos V e VII	2.570	2.412	93,9%
Grupos V e VI	2.401	2.271	94,6%

Fonte: Elaborada pelo autor

Com percentuais um pouco superiores à média geral, nos dois agrupamentos citados na Tabela 35, aproximadamente 94% dos candidatos possuíam os dois requisitos exigidos.

A Tabela 36 mostra o resultado do mesmo teste feito para cada uma das quatro opções de cota.

TABELA 36 – Candidatos a vagas no ensino superior que poderiam ter se inscrito em cada grupo de cotas

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo V	1.645	1.506	91,6%
Grupo VI	756	713	94,3%
Grupo VII	925	867	93,7%
Grupo VIII	570	555	97,4%
Total	3.896	3.641	93,5%

Fonte: Elaborada pelo autor

Os resultados indicaram percentuais um pouco acima da média geral nos grupos VI, VII e VIII. Pode-se concluir que a grande maioria dos candidatos inscritos nas cotas possuía os requisitos.

4.3.3 Cursos Técnicos Subsequentes

A seguir, foram avaliados os dados dos 1.291 candidatos que concorreram às vagas reservadas no ensino técnico subsequente. Foram utilizados os atributos `reserva_vagas_inscrito`, `sempre_escola_publica_qs`, `renda_menos_1.5_qs` e `cor_ppi_qs`. A Tabela 37 mostra o resultado do primeiro teste, referente ao estudo em escola pública.

TABELA 37 – Candidatos a vagas no ensino técnico subsequente que se inscreveram em cotas possuindo os requisitos

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo II	1.291	1.252	97,0%

Fonte: Elaborada pelo autor

O percentual de candidatos inscritos em cotas para cursos técnicos subsequentes que declararam ter estudado em escola pública foi de 97,0%, um pouco acima da média geral.

O teste seguinte foi feito com o agrupamento das cotas de cor e renda, cujos resultados estão descritos na Tabela 38:

TABELA 38 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupos V e VII	994	926	93,2%
Grupos V e VI	854	813	95,2%

Fonte: Elaborada pelo autor

Com percentuais um pouco acima da média geral, o resultado segue a tendência no fato de haver mais inconsistências em relação à autodeclaração de cor (Grupos V e VII) do que em relação à renda familiar *per capita* (Grupos V e VI).

A Tabela 39 mostra o resultado do mesmo teste feito para cada uma das opções de cota.

TABELA 39 – Candidatos a vagas no ensino técnico subsequente que poderiam ter se inscrito em cada grupo de cotas

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo V	676	622	92,0%
Grupo VI	178	167	93,8%
Grupo VII	318	292	91,8%
Grupo VIII	119	116	97,5%
Total	1.291	1.197	92,7%

Fonte: Elaborada pelo autor

Estes resultados indicam que a grande maioria dos candidatos que se inscreve nas cotas possui os requisitos. Os percentuais encontrados foram muito próximos da média geral, sendo um pouco menor apenas no grupo VII.

4.3.4 Cursos Técnicos Integrados

Considerando apenas os 2.691 candidatos que concorreram a vagas reservadas no ensino técnico integrado, o algoritmo *Apriori* foi executado com os atributos *reserva_vagas_inscrito*, *sempre_escola_publica_qs*, *renda_menos_1.5_qs* e *cor_ppi_qs*. Primeiramente foi avaliada a inscrição de uma maneira geral, sem especificar o tipo de cota, conforme a Tabela 40:

TABELA 40 – Candidatos a vagas no ensino técnico integrado que se inscreveram em cotas possuindo os requisitos

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo II	2.691	2.527	93,9%

Fonte: Elaborada pelo autor

Entre os candidatos a vagas no ensino técnico integrado, o percentual de candidatos que declarou ter estudado em escola particular em algum momento foi de 6,1%, enquanto no ensino superior e no ensino técnico subsequente este percentual esteve em torno de 3,0%.

O teste seguinte foi feito com o agrupamento das cotas de cor e renda, cujos resultados estão descritos na Tabela 41:

TABELA 41 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cotas agrupados pela cor e pela renda familiar

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupos V e VII	1.913	1.760	92,0%
Grupos V e VI	1.677	1.564	93,3%

Fonte: Elaborada pelo autor

Assim como na média geral, foram encontradas mais inconsistências em relação à cor (grupos V e VII) do que em relação à renda (grupos V e VI). Os percentuais de candidatos que possuíam os requisitos foram um pouco menores do que na média geral.

A Tabela 42 mostra o resultado do mesmo teste feito para cada uma das opções de cota.

TABELA 42 – Candidatos a vagas no ensino técnico integrado que poderiam ter se inscrito em cada grupo de cotas

Grupo	Inscritos nas vagas reservadas, cotas	Número de candidatos que possuíam os requisitos	Percentual de candidatos que possuíam os requisitos
Grupo V	1.269	1.166	91,9%
Grupo VI	408	363	89,0%
Grupo VII	644	585	90,8%
Grupo VIII	370	339	91,6%
Total	2.691	2.453	91,2%

Fonte: Elaborada pelo autor

Assim como na média geral, os resultados indicam que a grande maioria dos candidatos que se inscreve nas cotas possui os requisitos. Porém, nos cursos técnicos integrados, os percentuais de candidatos com os requisitos foram menores do que a média, com exceção do grupo V. No Grupo VI percebe-se a única situação entre todas as modalidades em que o percentual de candidatos inscritos sem os requisitos foi acima de 10%.

4.4 Estudantes matriculados nas cotas e na ampla concorrência com os requisitos necessários para as cotas

Com o objetivo de verificar a efetividade da política de cotas para o ingresso de candidatos oriundos de escola pública, de baixa renda e autodeclarados pretos, pardos ou indígenas durante o período pesquisado, foi feita uma comparação dos dados dos alunos matriculados nas cotas e dos alunos matriculados na ampla concorrência, tendo o perfil necessário para concorrer às cotas. Quanto maior a diferença entre estes grupos, mais se justificaria a reserva de vagas.

Para esses testes, foram utilizados os atributos chamado_para_cota, cor_ppi_qs, sempre_escola_publica_qs e renda_menos_1.5_qs.

4.4.1 Todos os cursos

Primeiramente foram avaliados os dados de todos os 4.579 matriculados. Os resultados desse teste estão descritos na tabela 43.

TABELA 43 – Candidatos que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência

	Escola pública, renda e cor	Escola pública e renda	Escola pública e cor	Escola pública
Grupo onde o candidato poderia concorrer	V, VI, VII ou VIII	VI ou VIII	VII ou VIII	VIII
Percentual de matriculados em cotas	60,9%	56,2%	60,7%	56,0%
Percentual de matriculados na ampla concorrência	39,1%	43,8%	39,3%	44,0%
Número de matriculados em cotas	1.250	1.840	1.351	1.997
Número de matriculados na ampla concorrência	803	1.433	873	1.569
Número total de matriculados	2.053	3.273	2.224	3.566
Percentual entre os 2.514 matriculados da ampla concorrência	31,9%	57,0%	34,7%	62,4%

Fonte: Elaborada pelo autor

Estes resultados indicam uma diferença entre os grupos muito menor do que se esperaria. Em todos os casos, observa-se altos percentuais de candidatos que possuem os requisitos (entre 39,1% e 44,0%) mas que conseguiram se matricular na ampla concorrência, ou seja, sem utilizar o sistema de cotas.

Outro dado interessante é o percentual de alunos matriculados na ampla concorrência que possui os requisitos para concorrer às vagas reservadas. Poderia-se esperar um valor baixo, já que muitos candidatos nessa situação se matricularam nas vagas reservadas. Quanto mais baixo esse percentual, mais se justificaria a necessidade da reserva de vagas. Em relação aos requisitos de escola pública e renda, o que observa é que, de acordo com as respostas do questionário socioeconômico, mesmo entre os matriculados na ampla concorrência, existem mais candidatos com os requisitos do que sem os requisitos. Em relação a cor, nota-se que cerca de um terço dos candidatos matriculados na ampla concorrência possuem esse requisito.

4.4.2 Cursos Superiores

Considerando-se apenas os 1.199 matriculados no ensino superior, chegou-se aos dados apresentados na Tabela 44.

TABELA 44 – Candidatos a vagas no ensino superior que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência

	Escola pública, renda e cor	Escola pública e renda	Escola pública e cor	Escola pública
Grupo onde o candidato poderia concorrer	V, VI, VII ou VIII	VI ou VIII	VII ou VIII	VIII
Percentual de matriculados em cotas	66,5%	59,5%	64,5%	57,8%
Percentual de matriculados na ampla concorrência	33,5%	40,5%	35,5%	42,2%
Número de matriculados em cotas	328	490	373	559
Número de matriculados na ampla concorrência	165	333	205	408
Número total de matriculados	493	823	578	967
Percentual entre os 627 matriculados da ampla concorrência	26,3%	53,1%	32,7%	65,1%

Fonte: Elaborada pelo autor

No ensino superior nota-se que a diferença entre os grupos é maior do que na média geral para os requisitos de cor e renda familiar. Mesmo assim, são altos os percentuais de candidatos que possuíam os requisitos e se matricularam na ampla concorrência, variando entre 33,5% e 42,2%. Para o requisito da escola pública, os resultados são muito próximos à média geral.

Assim como na média geral, para os requisitos de escola pública e renda familiar, nota-se que entre os matriculados na ampla concorrência há mais candidatos com os requisitos do que sem os requisitos.

4.4.3 Cursos Técnicos Subsequentes

A realização dos testes com os 1.185 matriculados no ensino técnico subsequente levou aos dados apresentados na Tabela 45.

TABELA 45 – Candidatos a vagas no ensino técnico subsequente que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência

	Escola pública, renda e cor	Escola pública e renda	Escola pública e cor	Escola pública
Grupo onde o candidato poderia concorrer	V, VI, VII ou VIII	VI ou VIII	VII ou VIII	VIII
Percentual de matriculados em cotas	46,5%	42,9%	46,8%	42,9%
Percentual de matriculados na ampla concorrência	53,5%	57,1%	53,2%	57,1%
Número de matriculados em cotas	314	435	338	469
Número de matriculados na ampla concorrência	361	580	384	624
Número total de matriculados	675	1.015	722	1.093
Percentual entre os 712 matriculados da ampla concorrência	50,7%	81,5%	53,9%	87,6%

Fonte: Elaborada pelo autor

Estes resultados indicam que menos da metade dos alunos que possuem os requisitos e se matricularam em cursos técnicos subsequentes na instituição no período estudado, fizeram uso do sistema de cotas.

O requisito que mais chama a atenção é o estudo em escola pública. Entre os matriculados na ampla concorrência, 87,6% declararam ter estudado sempre em escola pública. Mesmo se considerando a exigência dos três requisitos simultaneamente, há mais candidatos com os requisitos do que sem os mesmos, entre os matriculados na ampla concorrência.

4.4.4 Cursos Técnicos Integrados

Considerando-se apenas os 2.195 matriculados no ensino técnico integrado, chegou-se aos dados apresentados na Tabela 46.

TABELA 46 – Candidatos a vagas no ensino técnico integrado que possuíam os requisitos que foram matriculados em cotas e na ampla concorrência

	Escola pública, renda e cor	Escola pública e renda	Escola pública e cor	Escola pública
Grupo onde o candidato poderia concorrer	V, VI, VII ou VIII	VI ou VIII	VII ou VIII	VIII
Percentual de matriculados em cotas	68,7%	63,8%	69,3%	64,3%
Percentual de matriculados na ampla concorrência	31,3%	36,2%	30,7%	35,7%
Número de matriculados em cotas	608	915	640	969
Número de matriculados na ampla concorrência	277	520	284	537
Número total de matriculados	885	1.435	924	1.506
Percentual entre os 1.175 matriculados da ampla concorrência	23,6%	44,3%	24,2%	45,7%

Fonte: Elaborada pelo autor

Para os alunos matriculados no ensino técnico integrado, em todos os requisitos a diferença entre os grupos foi maior do que na média geral. De toda forma, o percentual de alunos com os requisitos que se matriculou pelo sistema de cotas não chegou a 70% para nenhum dos mesmos.

Nessa modalidade, o sistema de cotas parece fazer mais sentido do que nas demais, mas ainda assim, existem muitos candidatos que conseguem se matricular sem precisar concorrer às vagas reservadas. Nota-se que quase um quarto dos candidatos matriculados na ampla concorrência teriam todos os requisitos e poderiam se candidatar a qualquer um dos grupos de cotas, se desejassem.

5 Análise dos Resultados

Para este trabalho, foram executados testes na base de dados preparada para este fim, em busca de informações sobre 4 aspectos: a ocupação das vagas reservadas; o volume de inscrição nas cotas dos candidatos que cumprem os requisitos; o volume de inscrição nas cotas de candidatos que não cumprem os requisitos (de forma intencional ou não); e as matrículas dos candidatos que possuem os requisitos nas cotas e na ampla concorrência. Os testes foram executados com os dados de todos os cursos, e também fazendo a distinção entre cursos superiores, técnicos subsequentes e técnicos integrados, uma vez que cada um desses grupos possui um público alvo diferente. Com isso, foi possível comparar os grupos e avaliar as diferenças entre os mesmos.

5.1 Ocupação das vagas reservadas para cotas

Como mostrado na Tabela 3, houve 7.878 inscrições para as 2.629 vagas reservadas, ou seja, uma relação de 3,0 candidatos por vaga. Para a ampla concorrência, foram 11.490 inscrições para 2.602 vagas, portanto uma relação de 4,4 candidatos por vaga. Percebe-se que os candidatos que optaram por concorrer às vagas reservadas tiveram o benefício de uma concorrência menor. Esse fato se observou em todos os tipos de cursos, de acordo com as Tabelas 7, 11 e 15.

O agrupamento dos dados pela renda familiar *per capita* (Tabela 4) mostrou uma procura muito maior pelas vagas destinadas ao Grupo III (renda familiar *per capita* inferior a 1,5 salário mínimo) do que pelas vagas destinadas ao Grupo IV (renda familiar *per capita* superior a 1,5 salário mínimo). Enquanto no Grupo III a relação de candidatos por vaga foi de 3,7, no Grupo IV essa relação foi de 2,3. Essa diferença foi mais acentuada para o ensino técnico subsequente, onde a relação de inscritos por vaga foi de 2,1 no Grupo III e de 1,1 no Grupo IV, como mostra a Tabela 12.

O agrupamento dos dados pela autodeclaração de cor, descrito na Tabela 5, mostrou uma procura maior pelas vagas reservadas a candidatos que se declararam pretos, pardos e indígenas (Grupos V e VII) em relação às demais (Grupos VI e VIII). Porém essa distância entre os grupos varia por cada tipo de curso, como se vê na Tabela 47.

TABELA 47 – Relação de inscritos por vaga agrupados pela autodeclaração de cor

Ensino	Grupos V e VII	Grupos VI e VIII
Geral (todos os cursos)	3,4	2,4
Superior	6,2	5,4
Técnico Subsequente	2,0	0,9
Técnico Integrado	2,7	1,7

Fonte: Elaborada pelo autor

No ensino técnico integrado a distância entre os grupos é muito parecida com a média geral. No ensino superior, os grupos estão muito próximos, não havendo diferença significativa na procura pelas vagas reservadas a pretos, pardos e indígenas. No ensino técnico subsequente, além da distância entre os grupos ser muito grande, nota-se que houve menos candidatos do que vagas para as cotas que não levavam a cor em consideração.

Observando cada tipo de cota, pela Tabela 6, é possível verificar que no Grupo V (alunos de escola pública com renda familiar *per capita* inferior a 1,5 salário mínimo que se autodeclararam pretos, pardos ou indígenas) a procura pelas vagas reservadas foi muito maior do que nos demais grupos. Considerando todos os cursos, enquanto no Grupo V a relação de inscritos por vaga foi 4,3, nos demais grupos ela variou entre 2,1 e 2,6. Essa diferença do Grupo V para os demais se repete em todos os tipos de cursos, conforme as Tabelas 10, 14 e 18.

Como mostrado na Tabela 3, levando-se todos os cursos e tipos de cotas em consideração, foram ocupadas 78,5% das vagas reservadas. 564 vagas reservadas para cotas acabaram sendo ocupadas pela ampla concorrência.

Mais da metade dessas vagas eram do ensino técnico subsequente, onde apenas 60,2% das vagas reservadas foram ocupadas (Tabela 11). No ensino superior e no ensino técnico integrado o percentual de ocupação das vagas reservadas esteve próximo de 90% (Tabelas 7 e 15).

A avaliação dos dados agrupados pela renda familiar *per capita*, descrita na Tabela 4, mostra que houve maior ocupação das vagas reservadas a candidatos de baixa renda. Essa situação se repetiu em todos os tipos de cursos, como pode ser visto nas Tabelas 8, 12 e 16, sendo que a diferença entre os grupos foi mais acentuada no ensino técnico integrado.

O agrupamento dos dados pela cor, conforme a Tabela 5, mostrou maior ocupação das vagas reservadas a candidatos autodeclarados pretos, pardos ou indígenas. A Tabela 9 mostra pequena diferença entre os dois grupos para o ensino superior, enquanto os dados das tabelas 13 e 17 mostram que no ensino técnico a diferença foi muito significativa, principalmente na modalidade subsequente.

A Tabela 48 mostra o percentual de ocupação das vagas, com as diferenças encontradas por tipo de cota e por tipo de curso.

TABELA 48 – Ocupação das vagas reservadas por cota e por tipo de curso

Ensino	Grupo V	Grupo VI	Grupo VII	Grupo VIII	Total
Geral (todos os cursos)	88,8%	76,0%	77,8%	65,3%	78,5%
Superior	93,7%	83,8%	82,5%	85,3%	86,8%
Técnico Subsequente	69,7%	52,7%	67,1%	40,5%	60,2%
Técnico Integrado	99,2%	86,6%	82,1%	71,0%	86,1%

Fonte: Elaborada pelo autor

Avaliando-se cada tipo de cota, independente do curso, nota-se que houve uma ocupação muito alta das vagas do Grupo V, uma ocupação menor e parecida entre si para os Grupos VI e VII e uma ocupação bem mais baixa no Grupo VIII. Essa mesma relação foi vista para o ensino técnico integrado. No ensino superior, a ocupação de vagas do Grupo V se destacou das demais; nos Grupos VI, VII e VIII os percentuais foram parecidos, com uma pequena predominância do Grupo VIII. No ensino técnico subsequente, a relação foi bem diferente dos outros tipos de curso, pois os percentuais de ocupação dos Grupos V e VII foram muito próximos entre si, se destacando dos demais; No Grupo VIII apenas 40,5% das vagas reservadas para cotas foram ocupadas, sendo este o percentual de ocupação mais baixo entre todos.

De uma maneira geral, a ocupação das vagas acompanhou a tendência da procura por elas. No Grupo V houve mais candidatos inscritos por vagas e maior percentual de ocupação em todos os tipos de cursos. A exceção foi o Grupo VII no ensino técnico subsequente, que embora tenha tido uma procura muito menor do que a do Grupo V (Tabela 14) teve um percentual de ocupação muito parecido.

Para aumentar o percentual de ocupação das vagas reservadas, chegando aos 100%, seria recomendável aumentar a divulgação sobre os direitos dos candidatos, inclusive na própria interface do sistema que gerencia as inscrições, principalmente para o

ensino técnico subsequente, onde a ocupação das vagas foi muito menor do que nos demais.

5.2 Candidatos que poderiam concorrer às vagas, mas não o fizeram

Como visto no capítulo anterior, vagas não foram preenchidas, entretanto candidatos com perfil de cotas (cumprem os requisitos) se matricularam no Grupo I.

As inscrições dos candidatos e as respostas do questionário socioeconômico foram utilizadas para avaliar se o público de cotas em potencial está utilizando seu direito de concorrer às vagas que estão reservadas.

Pelos dados da Tabela 19, é possível perceber que menos de 50% dos candidatos que estudaram em escola pública no nível anterior, e portanto poderiam concorrer a alguma cota, estão fazendo uso desse direito. No ensino superior (Tabela 22) e no ensino técnico integrado (Tabela 28) o percentual está próximo de 50%; no ensino técnico subsequente (Tabela 25) apenas 40,3% dos candidatos nessas condições se inscreveram.

A Tabela 20 mostra que o uso do direito à inscrição nas vagas reservadas é maior entre os grupos que levam em consideração a cor (50,3%) do que entre os que levam em consideração a renda familiar (32,5%). Essa situação se repete em todos os tipos de cursos, como pode ser visto nas Tabelas 23, 26 e 29.

Pela Tabela 21 é possível perceber que muitos candidatos estão se inscrevendo para uma cota onde eles podem concorrer, mas que não seria a mais indicada de acordo com seu perfil, sendo esta situação recorrente em todos os tipos de curso, de acordo com as Tabelas 24, 27 e 30.

Outra situação interessante que pode ser observada na Tabela 21 é a distribuição dos candidatos pelos grupos de acordo com as respostas do questionário socioeconômico. Percebe-se que a grande maioria dos candidatos que podem concorrer a alguma cota, estariam concentrados nos Grupos V (escola pública, renda e cor) e VI (escola pública e renda) se fizessem a inscrição na opção mais adequada ao seu perfil. Isso acontece porque a grande maioria dos candidatos que responderam no questionário socioeconômico que sempre estudaram em escola pública também declararam ter renda familiar *per capita* inferior a 1,5 salário mínimo.

Sugere-se mais uma vez um trabalho de divulgação das cotas, e uma melhor exibição dessas informações no próprio formulário eletrônico de inscrição, como forma de

umentar as inscrições para as vagas reservadas e ajudar o candidato a escolher a opção mais adequada.

5.3 Inscrições que não cumprem os requisitos necessários às cotas

Os dados encontrados nesta parte da pesquisa apontam indícios de que há candidatos concorrendo a vagas reservadas sem ter os requisitos, e que não há diferença significativa entre os grupos, pois em todos os requisitos parece haver candidatos se inscrevendo sem ter o direito.

Ter estudado em escola pública é um requisito comum a todas as opções de cotas. Considerando todos os cursos, a Tabela 31 mostra que 4,1% dos candidatos se inscreveram em cotas mesmo tendo declarado no questionário socioeconômico que estudaram em escola particular. Nos cursos superiores (Tabela 34) e técnicos subsequentes (Tabela 37) esse percentual está em torno de 3%. Nos cursos técnicos integrados (Tabela 40) o percentual foi maior, de 6,1%.

Os dados da Tabela 32 mostram que há pouca diferença entre os grupos que precisam ter estudado em escola pública e possuir renda familiar *per capita* até 1,5 salários-mínimos (V e VI) e os grupos que precisam ter estudado em escola pública e se declarar pretos, pardos ou indígenas (V e VII); nos dois casos os percentuais estão apenas um pouco abaixo dos que precisam apenas ter estudado em escola pública (Tabela 31). Os dados das Tabelas 35, 38 e 41 mostram que esse cenário se mantém em todos os tipos de cursos.

Uma sugestão é que o sistema faça um cruzamento dos dados do questionário com a opção de cota selecionada, apontando a inconsistência, para que os candidatos que marcaram uma opção errada possam corrigi-la antes de finalizar a inscrição.

Outro ponto de atenção é quanto à comprovação da renda familiar e do estudo em escola pública. Pode ser que muitos dos candidatos tenham se candidatado a uma vaga sabendo que não tem o requisito exigido, mas pensando em omitir essa informação no momento da matrícula. Sugere-se um cuidado redobrado para esta comprovação quando da chamada desses candidatos para a matrícula na Instituição.

5.4 Matriculados com os requisitos nas cotas e na ampla concorrência

Nesta parte da pesquisa, foram avaliados os dados dos alunos matriculados, tanto em cotas quanto na ampla concorrência, com o objetivo de tentar entender qual foi a

efetividade do sistema de cotas no IFMG no período estudado, verificando se ele contribuiu significativamente para o ingresso de alunos com os requisitos exigidos na legislação.

A primeira avaliação foi feita, considerando os alunos que possuem os requisitos, verificando-se quantos ingressaram pelo sistema de cotas e quantos ingressaram na ampla concorrência. Partiu-se do princípio de quanto maior a diferença entre estes dois grupos, mais se justificaria o sistema de cotas. Por exemplo, se em um determinado grupo, 90% dos candidatos com os requisitos ingressassem no IFMG pelo sistema de cotas e 10% na ampla concorrência, isso apontaria para uma grande dificuldade de acesso sem a política de cotas.

Porém, os dados da Tabela 43 apontam que a diferença entre os grupos não é muito acentuada. A maior distância entre os grupos é referente aos alunos que possuem os três requisitos simultaneamente, sendo que 60,9% dos matriculados com esse perfil ingressaram na Instituição através do sistema de cotas e 39,1% conseguiram ingressar na ampla concorrência.

A segunda avaliação foi efetuada, considerando apenas os alunos que se matricularam na ampla concorrência, identificando-se quantos dentre eles possuíam os requisitos necessários para concorrer às cotas. Quanto menor esse número, maior seria a percepção de necessidade do sistema de cotas.

Pela Tabela 43, percebe-se que os percentuais variam entre 31,9% e 62,4%. São valores altos, ainda mais considerando-se que já estão descontados os alunos que ingressaram nas vagas reservadas. O que se conclui a partir desses números é que, para o IFMG no período estudado, não existiu a dificuldade de acesso aos alunos que estudaram em escola pública, que tem renda familiar *per capita* até 1,5 salário-mínimo ou que se consideram pretos, pardos ou indígenas, que justificasse a necessidade do sistema de cotas.

Pode-se dizer que no ensino superior o cenário é muito parecido com a média geral. Para estes cursos, conforme a Tabela 44, é notada na ampla concorrência uma participação um pouco maior de candidatos que estudaram em escola pública e um pouco menor de candidatos com os requisitos de cor e renda.

Pelos dados da Tabela 45 é possível notar que no ensino técnico subsequente, existem mais candidatos na ampla concorrência com os requisitos do que sem os requisitos, para todos os grupos. Os alunos de escola pública e com renda familiar *per capita* até 1,5 salário-mínimo passam de 80% entre os matriculados na ampla concorrência e mesmo quando são exigidos os três requisitos, o percentual está acima de 50%. Ou seja, não houve dificuldade alguma de acesso para candidatos com este perfil.

No ensino técnico integrado, a situação é diferente do ensino técnico subsequente. Na ampla concorrência existem mais candidatos sem os requisitos, como era de se esperar. Os percentuais de candidatos com os requisitos que se matricularam na ampla concorrência variam entre 30,7% e 36,2%. Embora estes percentuais sejam menores do que a média geral, ainda são altos, considerando-se que já estão descontados os alunos que ingressaram nas vagas reservadas. Nesta modalidade o sistema de cotas tem mais efetividade do que nas demais, e assim como na média geral, a efetividade é maior em relação à cor do que em relação à renda, mas mesmo assim, nos dados pesquisados não se observou dificuldade tão grande de acesso que justificasse a reserva de vagas.

6 Conclusões

A Lei de cotas foi criada para ampliar o acesso às Instituições Federais de Ensino a pessoas que estudaram em escolas públicas, através da reserva de 50% das vagas ofertadas por cada instituição; dentro dessas vagas reservadas existem também divisões com critérios econômicos e raciais.

Existe um grande debate político em torno da viabilidade da reserva de vagas e de seus efeitos para a sociedade brasileira, que foge ao escopo deste trabalho. Neste estudo não se procurou discutir os fundamentos da reserva de vagas para estas parcelas da população, mas sim pesquisar a aplicação da Lei de Cotas em uma instituição federal de ensino para compreender se os objetivos propostos pela legislação estão sendo alcançados nessa instituição. Partiu-se do princípio de que estudos como este podem ser feitos nas demais instituições federais de ensino, e de que os resultados encontrados podem contribuir para o debate sobre a relevância dessa política para a educação brasileira.

Acreditando que o conhecimento sobre a situação atual é fundamental para a tomada de decisão, esta pesquisa procurou avaliar os dados disponíveis e elaborar considerações sobre a política de cotas, tanto aquelas que possam servir de subsídio a decisões específicas do IFMG, quanto sugestões de melhorias na própria legislação federal.

O IFMG foi escolhido como objeto do estudo pela acessibilidade e pelo volume considerável de dados, gerados pelas inscrições para processos seletivos, para cursos superiores e técnicos entre novembro de 2012 e janeiro de 2014.

A tecnologia da informação foi fundamental para a pesquisa. O fato do preenchimento do formulário de inscrição e do questionário socioeconômico serem gerenciados por um sistema informatizado permitiu a geração de uma base de dados ampla e adequada para o estudo. O uso de técnicas de Descoberta de Conhecimento em Bases de Dados se mostrou satisfatório para ajudar a responder às questões de pesquisa levantadas, uma vez que foram reveladas situações que não correspondiam ao esperado no início do estudo. Para chegar ao resultado pretendido, o *software* Weka se mostrou uma boa ferramenta.

Foram apresentados resultados levando em consideração todos os candidatos, sem distinção de cursos, e também os resultados agrupados em três tipos de cursos: superiores, técnicos subsequentes e técnicos integrados. Essa separação se justifica pelas diferenças entre o público que concorre a esses cursos.

Para cada um dos objetivos específicos da pesquisa, os resultados conduziram às seguintes conclusões:

– Avaliar a real ocupação das vagas reservadas: em torno de 20% das vagas reservadas não foram ocupadas por candidatos cotistas, sendo transferidas para a ampla concorrência, sendo a maior parte delas no ensino técnico subsequente, onde foram ocupadas apenas 60,2% das vagas reservadas, enquanto no ensino superior e no ensino técnico integrado a ocupação esteve em torno de 86%. De uma maneira geral, quanto menor a procura pelas vagas, verificada pela relação entre inscrições e vagas reservadas, menor é o percentual de ocupação.

– Descobrir se os candidatos que possuem os requisitos estão fazendo uso do direito de se inscreverem nas vagas reservadas: menos da metade dos candidatos que poderiam se inscrever a algum tipo de cota concorreu a uma das vagas reservadas. Grande parte dos candidatos que puderam escolher entre mais de uma opção de cota, concorreu a uma opção diferente da que seria esperada pelo seu perfil. Notou-se que a grande maioria dos candidatos que sempre estudou em escola pública declarou ter renda familiar *per capita* inferior a 1,5 salário-mínimo. Não houve diferenças acentuadas entre os tipos de cursos.

– Descobrir se há candidatos concorrendo às vagas sem ter o direito: há indícios de que a grande maioria dos candidatos que concorre às cotas possui os requisitos exigidos, mas que existem candidatos que se inscrevem para concorrer às vagas reservadas sem ter o direito. Essa situação pode ser observada em todos os tipos de cursos e todos os requisitos. O maior percentual pode ser observado no ensino técnico integrado, em torno de 6%, enquanto no ensino superior e no ensino técnico subsequente o percentual está em torno de 3%.

– Descobrir se há diferenças significativas no perfil de candidatos que chega a ser matriculado em cotas em relação aos candidatos matriculados na ampla concorrência: não foram encontradas diferenças acentuadas entre os alunos que possuem os requisitos e ingressaram no IFMG pelo sistema de cotas e os alunos que, mesmo tendo os requisitos, ingressaram no IFMG pela ampla concorrência. Entre os alunos matriculados na ampla concorrência, foi encontrado alto percentual de candidatos com os requisitos para concorrer às cotas, principalmente no ensino técnico subsequente.

Não se sabe os motivos que levaram muitos candidatos a optarem pela ampla concorrência tendo os requisitos para concorrer às vagas reservadas ou a concorrer a uma opção de cota diferente da esperada. Pode-se presumir que ao menos em parte, o desconhecimento do seu direito tenha levado a essa situação. Sendo assim, sugere-se um

trabalho maior de divulgação do sistema de cotas, por parte da instituição; nesse sentido, seria recomendável que na interface do sistema que gerencia as inscrições fosse indicado com clareza ao candidato o direito que ele possui, inclusive utilizando as respostas do questionário socioeconômico para sugerir as opções mais indicadas.

Este trabalho apontou que houve candidatos inscritos nas cotas sem possuir os requisitos. Não se sabe se estas inscrições foram feitas de maneira intencional ou por equívoco dos candidatos. Considerando-se que um aluno será eliminado do processo seletivo caso seja chamado para uma vaga reservada e não tenha os requisitos necessários no momento da matrícula, sugere-se uma melhoria no sistema de inscrições, para que faça um cruzamento dos dados socioeconômicos com a opção de cota e aponte a inconsistência no momento da inscrição.

O fato de haver candidatos inscritos em cotas sem os requisitos alerta para a necessidade da comprovação das informações referentes ao estudo em escola pública e à renda familiar *per capita*. Quanto ao estudo em escola pública, a legislação permite que a Instituição exija a comprovação. O resultado deste trabalho aponta que essa é uma prática recomendável, uma vez que há candidatos que estudaram em escola particular em algum momento, e que mesmo assim se inscreveram para concorrer às vagas reservadas.

Considerando que aparentemente existem poucos candidatos que estudaram em escola pública e possuem renda familiar *per capita* acima de 1,5 salário-mínimo, sugere-se que o IFMG estude a possibilidade de alterar o percentual de 50% das vagas reservadas a este público.

Para o IFMG, uma conclusão geral é de não haver grande efetividade da reserva de vagas, principalmente para os cursos técnicos subsequentes, uma vez que o estudo aponta que os percentuais estabelecidos pela legislação seriam alcançados mesmo sem a reserva das vagas. Talvez isso aconteça devido ao perfil do público que concorre aos cursos superiores e técnicos da instituição, onde os candidatos oriundos de escola pública, com baixa renda e que se consideram pretos, pardos ou indígenas representam um percentual de inscrições muito mais alto do que o percentual de vagas reservadas a eles.

Os resultados da pesquisa permitem também tecer algumas considerações que poderiam ser avaliadas em relação à própria política de cotas.

A legislação prevê a obrigatoriedade da comprovação da renda familiar, porém esta é uma comprovação difícil por parte de cada instituição, uma vez que os candidatos podem simplesmente omitir informações para parecer que a sua renda familiar *per capita* é inferior a 1,5 salário mínimo. Uma sugestão seria utilizar o cadastro de beneficiários de

programas sociais do próprio governo federal para verificar quem são as pessoas de baixa renda.

Em vez de permitir às instituições que exijam a comprovação do estudo em escola pública, seria interessante determinar que essa comprovação seja feita, para evitar o risco de que vagas reservadas venham a ser ocupadas por candidatos oriundos de escolas particulares dispostos a burlar o sistema.

Outra questão que merece maior debate é avaliar se uma autodeclaração de cor é um critério razoável, uma vez que qualquer pessoa pode mentir sobre a sua cor, e nesse caso, não há meios de evitar que ela se beneficie dessa mentira.

Os resultados da pesquisa no IFMG apontam indícios de que a política de cotas seria desnecessária nessa instituição. Sugere-se que estudos como este sejam feitos em todo o país, com outros institutos federais e também com as universidades federais, para que seja possível avaliar se a reserva de vagas está trazendo os resultados esperados em outras instituições.

Ao fim desta pesquisa, apontam-se também possibilidades de outros trabalhos futuros, usando os dados do IFMG ou de outra instituição federal de ensino.

A aplicação de um questionário aos alunos matriculados que poderiam concorrer às cotas, mas não o fizeram, bem como os alunos matriculados que concorreram a opções de cotas diferentes da esperada, poderia ajudar a compreender os motivos que levaram a isso. Com estas informações, seria mais fácil direcionar as ações de divulgação para ampliar as inscrições nas cotas e, conseqüentemente, a ocupação das vagas reservadas.

Considerando que houve inscrições de candidatos que não tinham os requisitos, seria viável avaliar se foram efetivadas matrículas de candidatos nessa situação.

Por último, sugere-se estudos que façam uma comparação do desempenho acadêmico de alunos que ingressaram pelo sistema de cotas com os alunos da ampla concorrência.

Referências

AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules in large databases. In: INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES, 20., 1994, Santiago, Chile. **Proceedings...**, San Francisco: Morgan Kaufmann Publishers Inc., 1994. p. 487-499.

AMARAL, F. C. N. **Data mining**: técnicas e aplicações para o marketing direto. São Paulo: Berkeley, 2001. 128p.

AMOROSO, E. D. **Um estudo de caso sobre mineração de dados como instrumento de aprendizado para o investidor do mercado de ações**. 2009. 121 f. Dissertação (Mestrado em Gestão do Conhecimento e Tecnologia da Informação). Universidade Católica de Brasília, 2009.

BATH, P. A. Data mining in health and medical information. **Annual review of information science and technology**, v. 38, n. 1, p. 331-369, 2004.

BERRY, M. J. A.; LINOFF, G. **Data mining techniques** – for marketing, sales, and customer support. John Wiley & Sons, New York, 1997.

BOENTE, A. N. P.; GOLDSCHMIDT, R. R.; ESTRELA, V. V. **Uma Metodologia de Suporte ao Processo de Descoberta de Conhecimento em Bases de Dados**. In: V Simpósio de Excelência em Gestão e Tecnologia, 2008, Resende – RJ. V SEGeT, 2008. v. 1.

BORKO, H. Information Science: what is it? **American Documentation**, Washington, v. 19, p. 3-5, 1968.

BRASIL. Decreto nº 7.824, de 11 de outubro de 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/Decreto/D7824.htm>. Acesso em: 10 jun. 2014.

BRASIL. Lei nº 11.892, de 29 de dezembro de 2008. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2008/Lei/L11892.htm>. Acesso em: 10 jun. 2014.

BRASIL. Lei nº 12.711, de 29 de agosto de 2012. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12711.htm>. Acesso em: 10 jun. 2014.

CARDOSO, O. N. P; MACHADO, R. T. M. Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. **Revista de Administração Pública**, v. 42, n. 3, p. 495-528, mai/jun, 2008.

CASTELLS, M. **A sociedade em rede: a era da informação, economia, sociedade e cultura.** São Paulo: Paz e Terra, 1999. 698p.

CHEN, S. Y.; LIU, X. The contribution of data mining to information science. **Journal of Information Science**, v. 30, n. 6, p. 550-558, 2004.

CHOO, C. W. **A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões.** São Paulo: SENAC, 2006. 425p.

DAVENPORT, T. H.; PRUSAK, L. **Conhecimento Empresarial: como as organizações gerenciam o seu capital intelectual.** Rio de Janeiro: Campus, 1998. 237p.

DRUCKER, P. **Sociedade pós-capitalista.** São Paulo: Pioneira, 1994. 186p.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data mining to Knowledge Discovery in Databases. **AI Magazine**, v. 17, n. 3, p. 37-54, 1996.

FILHO, L. A. S. **Mineração de regras de associação utilizando KDD e KDT: uma aplicação em segurança pública.** 2009. 85 f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belém, 2009.

FRAWLEY W. J.; PIATETSKY-SHAPIRO G.; MATHEUS C. J. Knowledge discovery in databases: An overview. In G. Piatetsky-Shapiro and W. J. Frawley, editors, **Knowledge Discovery in Databases**, pages 1-27. AAAI/MIT Press, 1992. 539p.

GIL, A. C. **Como elaborar projetos de pesquisa.** 4.ed. São Paulo: Atlas, 2002. 176p.

GOLDSCHMIDT, R. R.; PASSOS, E. **Data mining: Um Guia Prático.** Rio de Janeiro: Campus, 2005. 255p.

IBGE – Instituto Brasileiro de Geografia e Estatística. **Censo demográfico 2010: Características da População e dos Domicílios – Resultados do Universo.** Rio de Janeiro, 2011. Disponível em: <http://biblioteca.ibge.gov.br/d_detalhes.php?id=793>. Acesso em: 07 jul. 2014.

KISAHLEITNER, M. **Análise de técnicas de Data Mining na aquisição de clientes de cartão de crédito não correntistas.** 2008. 93 f. Dissertação (Mestrado em Administração de Empresas) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2008.

LYMAN, P.; VARIAN, H. R. **How Much Information 2003?** School of Information Management and Systems, University of California at Berkeley. 2003. Disponível em <<http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>>. Acesso em 20/08/2013.

MIRANDA, R. C. R. O uso da informação na formulação de ações estratégicas pelas empresas. **Ciência da Informação**, Brasília, v. 28, n. 3, p. 286-292, set/dez, 1999.

MORESI, E. A. D. Integração Organizacional: um referencial integrado. **Ciência da Informação**, Brasília, v. 30, n. 2, p. 35-46, mai/ago, 2001.

MUSSI, C. C.; ANGELONI, M. T. **Mapeamento das fontes do conhecimento organizacional: um suporte ao compartilhamento do conhecimento tácito.** Anais ISKM/DM 2000 – International Symposium on Knowledge Management/Document Management, Curitiba, 26-29-nov, 2000.

PRATI, R. C. Mineração de dados: uma ponte entre informação e conhecimento. **Fonte**, Belo Horizonte, v.7, n.10, p. 122-128, dez/2010.

QUONIAM, L.; TARAPANOFF, K.; JÚNIOR, R. H. A.; ALVARES, L. Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil. **Ciência da Informação**, Brasília, v. 30, n. 2, p. 20-28, mai/ago, 2001.

SANTOS, R. S. **Aplicação de um modelo preditivo de mineração de dados para apoio à decisão de crédito.** 2006. 92 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

SCHIESSL, J. M. **Descoberta de Conhecimento em Texto aplicada a um sistema de atendimento ao consumidor.** 2007. 106 f. Dissertação (Mestrado em Ciência da Informação) – Departamento de Ciência da Informação e Documentação. Universidade Federal de Brasília, Brasília, 2007.

SILVA, C. V. S.; RALHA, C. G. “Detecção de cartéis em licitações públicas com agentes de mineração de dados. **Revista Eletrônica de Sistemas de Informação**, v. 10, n. 1, p. 1-19, jan/jun, 2011.

SOUZA, M. P. N. Abordagem inter e transdisciplinar em ciência da informação. In: TOUTAIN, L. M. B. B. (Org.). **Para entender a ciência da informação.** Salvador: EDUFBA, p. 75-90, 2007.

SULAIMAN, A.; SOUZA, J. M. Panorama da mineração de dados no Brasil. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 4., 2000, Brasília. **Anais...** Brasília: UnB, 2000.

TAKEUCHI, H.; NONAKA, T. **Gestão do conhecimento**; tradução Ana Thorell. Porto Alegre: Bookman, 2008. 319p.

TURBAN, E.; MCLEAN, E.; WETHERBE, J. C. **Tecnologia da Informação para Gestão: Transformando os negócios na Economia Digital**. 3.ed. Porto Alegre: Bookman, 2004.

WEILL, P.; ROSS, J. W. **Governança de Tecnologia da Informação**. São Paulo: Makron Books, 2006. 276p.

WU, X.; KUMAR, V.; QUINLAN, J. R.; GHOSH, J.; YANG, Q.; MOTODA, H.; MCLACHLAN, G. J.; NG, A.; LIU, B.; YU, P. S.; ZHOU, Z. H.; STEINBACH, M.; HAND, D. J.; STEINBERG, D. Top 10 algorithms in data mining. **Knowledge and Information Systems**. Londres. vol. 14, n. 1, p. 1-37. Jan. 2008.