

Pontifícia Universidade Católica de Goiás
Programa de Mestrado em Engenharia de Produção e Sistemas

APLICAÇÃO DA MINERAÇÃO DE DADOS NA DESCOBERTA DE
PADRÕES DO PERFIL DE ALUNOS DO CURSO DE SI-U_nUCET-UEG

RONALDO DE CASTRO DEL-FIACO

Março 2012

Pontifícia Universidade Católica de Goiás
Programa de Mestrado em Engenharia de Produção e Sistemas

APLICAÇÃO DA MINERAÇÃO DE DADOS NA DESCOBERTA DE
PADRÕES DO PERFIL DE ALUNOS DO CURSO DE SI-U_nUCET-UEG

RONALDO DE CASTRO DEL-FIACO

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Orientador: Sibelius Lellis Vieira,
Doutor

Goiânia
Março 2012

D331a Del-Fiaco, Ronaldo de Castro.
Aplicação da mineração de dados na descoberta de
padrões do perfil de alunos do curso de SI-UNUCETUEG [manuscrito] /
Ronaldo de Castro Del-Fiaco. – 2012.
112 f. : il.

Bibliografia: f. 79-82
Dissertação (mestrado) – Pontifícia Universidade Católica de Goiás,
Programa de Pós-Graduação em Engenharia de Produção e Sistemas, 2012.
Orientador: Doutor Sibelius Lellis Vieira.
Inclui lista de figuras, tabelas e abreviaturas.
Inclui Anexo

1. Sistemas de informação – curso – perfil de alunos- mineração de dados –
utilização – UnCET-UEG – Goiânia (GO). 2. Mineração de dados. 3. WEKA
(*Waikato Environment for Knowledge Analysis*). I. Título.
CDU: 658.5:378-051(817.3)(043.3)

APLICAÇÃO DA MINERAÇÃO DE DADOS NA DESCOBERTA DE PADRÕES
DO PERFIL DE ALUNOS DO CURSO DE SI-UnUCET-UEG

RONALDO DE CASTRO DEL-FIACO

Dissertação julgada adequada para obtenção do título de Mestre em Engenharia de Produção e Sistemas, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás em Fevereiro de 2012.

Prof. Ricardo Luiz Machado, Dr.
Coordenador do Programa de Pós-Graduação em
Engenharia de Produção e Sistemas

Banca Examinadora:

Prof. Sibelius Lellis Vieira, Dr.
Orientador

Prof^a. Solange da Silva, Dr^a.

Prof. Wellington Santos Martins, Ph.D.

Prof^a. Maria José Pereira Dantas, Dr^a.

Agradecimentos

Agradeço a Deus pelo Seu infinito Dom de Amor por mim, ao colocar pessoas especiais em meu caminho, como os professores, colegas e a secretária acadêmica do Curso de Mestrado em Engenharia de Produção e Sistemas. Principalmente, ao professor Dr. Sibelius, pela paciência, dedicação, orientação, sugestões, disponibilidade, amizade, confiança e crédito à minha pessoa. Espero não tê-lo decepcionado.

Aos professores do Curso de Sistemas de Informação da UnUCET-UEG.

À Pró-Reitoria de Graduação da UEG, por liberar os dados utilizados neste trabalho.

Ao Núcleo de Seleção da UEG.

Aos colaboradores de TI da UEG.

À minha mãe Raimunda de Castro Del-Fiaco e ao meu pai Eurico Del-Fiaco (*in memoriam*), pelos seus exemplos de conquistas e determinações.

Carinhosamente, à minha esposa Idelma, pelas suas incansáveis revisões, apoio e paciência. Também aos meus filhos Nathália, João Paulo e Maria Gabriela, pelo amor e compreensão neste período de estudo, durante o qual estive ausente em muitos momentos de suas vidas.

Resumo da Dissertação apresentada ao MEPROS/ PUC Goiás, como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia de Produção e Sistemas (M.Sc.)

APLICAÇÃO DA MINERAÇÃO DE DADOS NA DESCOBERTA DE PADRÕES DO PERFIL DE ALUNOS DO CURSO DE SI-UNUCET-UEG

RONALDO DE CASTRO DEL-FIACO

Março de 2012

Orientador: Sibelius Lellis Vieira, Doutor

A mineração de dados (*Data Mining* - DM) é uma parte do processo para descobrir conhecimento em base de dados. Sua realização exige conhecimentos de várias áreas tais como a computação, a estatística, as ciências administrativas e do próprio negócio. Em particular, pode ser aplicada para descobrir conhecimento que permita ao gestor educacional melhorar a qualidade do processo ensino-aprendizagem no qual esteja envolvido. Este trabalho apresenta o embasamento teórico em mineração de dados, descreve e analisa um estudo de caso, cujo principal objetivo é aplicar a metodologia *Cross-Industry Standard Process for Data Mining* - CRISP-DM na identificação do padrão do perfil do aluno que integraliza o curso de Bacharelado em Sistemas de Informação da Unidade Universitária de Ciências Exatas e Tecnológicas da Universidade de Estadual de Goiás, no tempo mínimo previsto pelo projeto pedagógico do curso. É realizada a preparação dos dados que são utilizados no processo para, posteriormente, identificar as melhores propostas de análise do estudo de caso. Como dados de entrada, recorre-se às respostas do questionário socioeconômico e cultural aplicados aos vestibulandos e ao histórico escolar dos mesmos, que são utilizados como atributos para a avaliação dos algoritmos de árvore de decisão, através da ferramenta WEKA. Com este estudo, foi possível perceber que, para realizar a mineração de dados, é necessário um profissional que domine a teoria de DM, saiba calibrar uma ferramenta computacional e tenha conhecimento aprofundado do negócio, para determinar os objetivos da DM e interpretar os resultados encontrados.

Palavras-chave: Mineração de dados, Descoberta de conhecimento, WEKA, Classificação.

Abstract

The Data Mining (DM) is a part of the process of Knowledge Discovery in Databases. Its implementation requires knowledge of various areas such as computer sciences, statistics, management sciences and the business itself. In particular, it can be applied to discover knowledge that allows the manager to improve the quality of the learning process in which he/she is involved. This work presents the theoretical background of data mining, describes and analyzes a case study where the main objective is to apply the Cross-Industry Standard Process for Data Mining - CRISP-DM methodology to identify the profile of the student that graduates in due time from the course of Bachelor of Information Systems of the State University of Goiás at Anápolis. It describes the data preparation that is used in the process and identify the best proposals for analysis of the case study. As input data, both the transcripts and the answers of the socioeconomic and cultural questionnaire applied to the students are used as attributes for the evaluation of decision tree algorithms implemented in the data mining tool known as WEKA. It was possible to realize that data mining requires a professional who mastered the theory of DM to correctly calibrate the tools and extensive knowledge of the business in order to determine the data mining goals and interpret the results.

Keywords: Data Mining, Knowledge Discovery, WEKA, Classification.

Sumário

Lista de figuras	viii
Lista de tabelas	ix
Lista de siglas e abreviaturas	x
1 Introdução	1
1.1 Justificativa	1
1.2 Objetivos	4
1.3 Organização do trabalho	5
2 Referencial teórico	6
2.1 Descoberta de conhecimento em banco de dados - DCBD	6
2.1.1 Etapas do processo de DCBD	6
2.2 <i>Cross-industry standard proces for data mining</i> - CRISP-DM	8
2.2.1 Entendimento do negócio	9
2.2.2 Entendimento dos dados	9
2.2.3 Preparação dos dados	10
2.2.4 Construindo um modelo	13
2.2.5 Teste e avaliação	13
2.2.6 Disponibilização	14
2.3 Mineração de dados	14
2.3.1 Os dados na mineração de dados	15
2.3.2 Tarefas de mineração de dados	18
2.4 Outros processos de DCBD	28
2.5 Aplicação de mineração de dados	29
2.6 Outros trabalhos de mineração de dados na educação	36
3 Material e métodos	38
3.1 Problema da pesquisa	38
3.2 Hipótese	39
3.3 Metodologias aplicadas	39
3.4 Materiais usados	40
3.4.1 Ferramenta WEKA	40

4	Descrição do estudo de caso	46
4.1	O Curso de Sistemas de Informação da UnUCET-UEG	46
4.2	Processo seletivo	52
5	Aplicação de DCBD para identificar o perfil do aluno	55
5.1	Determinação dos objetivos do negócio	55
5.2	Avaliação da situação	56
5.3	Determinação dos objetivos da mineração dos dados	57
5.4	Preparação dos dados	58
5.5	Mineração de dados	62
5.5.1	Atributos para mineração de dados	62
5.5.2	Primeiro ciclo de mineração de dados	67
5.5.3	Segundo ciclo de mineração de dados	69
5.5.4	Terceiro ciclo de mineração de dados	70
5.5.5	Quarto ciclo de mineração de dados	70
5.5.6	Quinto ciclo de mineração de dados	72
5.5.7	Resultado da mineração de dados	74
6	Conclusão.....	76
	Referências	79
	Anexos.....	83
	Anexo A – Questionários Socioeconômico e Cultural	83
	Anexo B – Alternativas categorizadas do questionário socioeconômico e cultural.....	93

Lista de figuras

Figura 1 - Etapas do processo DCBD.....	7
Figura 2 - Processo CRISP-DM	8
Figura 3 - Etapas de pré-processamento de dados.....	11
Figura 4 - Taxonomia das tarefas de mineração de dados.....	21
Figura 5 - Árvore de decisão	23
Figura 6 - Grupos encontrados pela tarefa de agrupamento	25
Figura 7 - Dados associados a uma regra	27
Figura 8 - Tela do WEKA lendo um arquivo arff	42
Figura 9 - Botão que lista os algoritmos de classificadores disponíveis.	43
Figura 10 - Parâmetros do classificador J48.....	43
Figura 11 - Dados estatísticos dos atributos do arquivo alunos-si-01	66
Figura 13 - Árvore de decisão do arquivo alunos-si-01	68
Figura 12 - Resultado da Classificação do arquivo alunos-si-01	68
Figura 14 - Resultado da classificação do arquivo alunos-si-02	69
Figura 15 - Estatística do resultado da classificação usando fator de confiança 15%....	71
Figura 16 - Árvore de decisão com fator de confiança de 15%.	72
Figura 17 - Estatística do resultado da classificação do 5º ciclo	73
Figura 18 - Árvore criada no 5º ciclo	73
Figura 19 - Gráfico comparativo dos ciclos de mineração de dados.....	74

Lista de tabelas

Tabela 1 – Matriz Curricular 2002	47
Tabela 2 - Matriz Curricular 2009	49
Tabela 3 - Alunos matriculados de 2005 à 2009	51
Tabela 4 - Indicativos do Curso de SI UnUCET-UEG	51
Tabela 5 - Concorrência do processo seletivo do período de 2001 à 2008	54
Tabela 6 - Atributos dos dados acadêmicos	58
Tabela 7 - Colunas da planilha dados dos processos seletivos	59
Tabela 8 - Layout da tabela “Questões Vestibular”	61
Tabela 9 – Alternativas agrupadas da questão Q6.....	63
Tabela 10 - Alternativas agrupadas da questão Q13	63
Tabela 11 - Alternativas agrupadas da questão Q18	63
Tabela 12 - Alternativas agrupadas da questão Q20	63
Tabela 13 - Alternativas agrupadas da questão Q22	64
Tabela 14 - Alternativas agrupadas da questão Q26	64
Tabela 15 - Alternativas agrupadas da questão Q27	65
Tabela 16 - Alternativas agrupadas da questão Q30	65
Tabela 17 - Alternativas agrupadas da questão Q31	65
Tabela 18 - Alternativas agrupadas da questão Q33	66
Tabela 19 – Comparativo entre o 1º e o 2º ciclo	70
Tabela 20 - Percurso das melhores folhas	75

Lista de siglas e abreviaturas

ARFF	<i>Attribute-Relation File Format</i>
C4.5	Algoritmo para gerar uma árvore de decisão
CART	<i>Classification and Regression Trees</i>
CRISP_DM	<i>Cross-Industry Standard Process for Data Mining</i>
CRM	<i>Customer Relationship Management</i>
CsU	Conselho Universitário
DCBD	Descoberta de Conhecimento em Banco de Dados
DM	<i>Data Mining</i>
EDM	<i>Educational Data Mining</i>
ID3	Algoritmo para gerar uma árvore de decisão
IES	Instituição de Ensino Superior
J48	Implementação do algoritmo C4.5 no WEKA
<i>KDD</i>	<i>Knowledge Discovery in Databases</i>
MF	Média final
MP	Média parcial
PF	Prova final
PrG-UEG	Pró-Reitoria de Graduação da UEG
SEMMA	Selecionar Explorar Modificar Modelar e Avaliar
SGBD	Sistemas de Gerenciamento de Banco de Dados
SI	Sistemas de informação
UEG	Universidade Estadual de Goiás
UnUCET	Unidade Universitária de Ciências Exatas e Tecnológicas

VA Verificação de aprendizagem

WEKA *Waikato Environment for Knowledge Analysis*

1 Introdução

1.1 Justificativa

A evolução na área de administração de empresas demonstra uma mudança no foco das organizações, que se inicia na produção de bens e serviços, em sua estrutura organizacional, nos colaboradores e alcançando, finalmente a sua maior riqueza: o cliente. Os clientes são os ativos mais importantes de uma organização, seja ela privada ou pública. Sem eles, a organização perde o sentido de existir. É por isso que uma organização deve planejar e empregar uma estratégia clara para lidar com os clientes. Uma metodologia que várias empresas utilizam para esta finalidade é o *Customer Relationship Management (CRM)*, expressão em inglês que pode ser entendida como a Gestão de Relacionamento com o Cliente (CHIAVENATO, 2011).

O conceito de CRM é antigo, se entendido como a maneira pela qual os comerciantes de pequenos estabelecimentos costumam lidar com seus clientes típicos, tratando-os pelo nome e conhecendo as suas necessidades, a denominada venda da esquina. A novidade é que através do uso dos recursos tecnológicos, podem-se aplicar as práticas do pequeno comerciante a uma grande quantidade de clientes, deixando transparecer para o cliente a sua importância, por meio de um atendimento personalizado, no qual todos os colaboradores sabem seu nome e suas preferências (GUMMESSON, 2010).

O CRM é uma metodologia que visa entender, antecipar e administrar as necessidades dos clientes potenciais de uma empresa, tratando-se de um conjunto de estratégias, processos, mudanças organizacionais e técnicas pelas quais a empresa

pretende administrar melhor seu próprio empreendimento com relação ao comportamento de seus clientes (MONTEIRO, 2005).

No âmbito de instituições educacionais, o aluno pode ser considerado um cliente, na medida em que é a ele que se presta um serviço. O desempenho do aluno, por outro lado, também é parte desta prestação de serviço, desta vez para a comunidade. Neste sentido, a análise do que leva um aluno a ter um bom desempenho é crucial para uma boa prestação de serviço (FERNANDES, 2007).

Uma instituição educacional possui diversos tipos de relacionamentos com seus alunos (relacionamentos de marketing, acadêmicos, de serviços diversos ou financeiros), e na maioria das instituições esses relacionamentos ocorrem de forma não estruturada e não coordenada. O resultado disso é a redundância de algumas informações sobre os alunos e falta de outras, nos diversos níveis da instituição. O CRM aplicado ao ambiente educacional visa ampliar e fornecer ao gestor educacional conhecimentos sobre os alunos através dos dados e informações gerados durante todo o processo, permitindo com isso um tratamento individualizado, a fim de proporcionar interações diretas, ativas e consistentes com os mesmos (CARLOS, 2009).

Os softwares de CRM possuem três módulos distintos, caracterizados pela tecnologia utilizada: operacional, colaborativo e analítico. O módulo operacional é composto pelos mecanismos de coleta dos dados dos clientes, registros dos contatos entre o cliente e a empresa, gestão de campanhas de marketing. O módulo colaborativo garante que o fluxo de dados coletados nos vários pontos de contatos interaja com o resto da empresa através de um sistema automatizado. O módulo analítico, por sua vez, envolve a utilização de modelos de mineração de dados, a fim de avaliar o valor dos

clientes, entender e prever seu comportamento, analisando os padrões de dados para extrair conhecimento, a fim de aperfeiçoar o relacionamento com o cliente (TSIPTISIS e CHORIANOPOULOS, 2009).

A mineração de dados é uma etapa do processo de descoberta de conhecimento em banco de dados (DCBD), contemplando um conjunto de ferramentas e técnicas que, por meio do uso de algoritmos, são capazes de explorar um conjunto de dados, extraindo ou evidenciando padrões e predizendo conhecimento. Esse conhecimento pode ser apresentado por essas ferramentas de diversas formas: classificação, agrupamentos, hipóteses, regras, árvores de decisão, grafos, ou dendrogramas (FAYYAD *et al.*, 1996).

No processo de DCBD, é imprescindível agregar o conhecimento do negócio ao selecionar os dados para serem submetidos aos métodos e algoritmos de mineração de dados, originando um modelo que englobe e leve em consideração todas as peculiaridades do problema a ser resolvido. Somente dessa forma, pode ser gerada inteligência sobre o negócio, que pode e deve servir de base no processo decisório para gestão de relacionamento com cliente. Quanto mais uma organização conhece sua clientela, mais ela poderá acertar em suas ações e, conseqüentemente, ser mais competitiva (GUMMESSON, 2010).

A mineração de dados vem sendo aplicada cada vez mais em diversas organizações, com a intenção de se obter ganhos e vantagens competitivas nas mais diferentes áreas de atuação como bancos, indústria, telecomunicações, seguros, e em algumas organizações de utilidade pública como saúde, educação e justiça (SCHNEIDER, 2003).

O objetivo de qualquer organização é ser a mais competitiva possível. Neste trabalho, aborda-se uma organização do tipo instituição pública de ensino. O fato de ser pública não significa que não seja necessário preocupar-se com a qualidade, com a concorrência com outras instituições de ensino superior ou em promover ações para evitar evasão de alunos. Além disso, deverá atualizar seus objetivos com as exigências do mercado.

Com o surgimento de outros cursos na área de informática na cidade de Anápolis e na região circunvizinha, houve uma diminuição na concorrência do processo seletivo para ingresso ao Curso de Sistemas de Informação e, conseqüentemente, percebe-se, de forma intuitiva, o quanto o interesse e o rendimento dos alunos vêm diminuindo no decorrer dos últimos anos, conforme dados registrados no novo Projeto Pedagógico do Curso de Sistemas de Informação da Unidade Universitária de Ciências Exatas e Tecnológicas da Universidade Estadual de Goiás (BRITO *et al.*, 2009). É importante, portanto, suprir a instituição de informação inteligente para realizar ações direcionadas, a fim de atrair candidatos com um perfil o mais adequado possível ao curso e, assim, reverter este cenário.

1.2 Objetivos

O objetivo geral desse trabalho é conhecer os processos de descoberta de conhecimento em banco de dados, sobretudo a etapa de mineração de dados, focando nos algoritmos e técnicas de classificação e aplicar este conhecimento em um estudo de caso, para identificar padrões que representam o perfil do aluno do Curso de SI-UnUCET. Pretende-se identificar o perfil do aluno que possui bom rendimento escolar, a partir de dados contidos em seu histórico universitário e em dados do questionário

socioeconômico e cultural obtidos de sua entrada na universidade. Para atingir este objetivo, a tarefa de classificação é a mais indicada.

1.3 Organização do trabalho

O presente trabalho está sistematizado em seis capítulos. O capítulo 2 descreve o embasamento teórico do processo de descoberta do conhecimento em base de dados e o detalhamento da metodologia aplicada no caso de estudo deste trabalho, CRISP-DM. O capítulo 3 apresenta os materiais e métodos utilizados para desenvolver o caso de estudo. No capítulo 4 encontra-se uma breve descrição do Curso de SI-UnUCET-UEG. O capítulo 5 descreve toda aplicação realizada com os dados do curso e os comentários dos resultados encontrados. As conclusões e proposta para trabalhos futuros são discutidos no capítulo 6.

2 Referencial teórico

O mecanismo de descoberta do conhecimento em mineração de dados consiste em uma série de etapas, iniciando com a definição dos objetivos para os quais é aplicado o novo conhecimento, até a exposição do mesmo à alta direção da organização, como apoio à tomada de decisão. A mineração de dados propriamente dita é apenas uma dessas etapas.

2.1 Descoberta de conhecimento em banco de dados - DCBD

O termo *Knowledge Discovery in Databases (KDD)* ou em português Descoberta de Conhecimento em Banco de Dados (DCBD) foi formalizado por Fayyad, em 1989 (FAYYAD *et al.*, 1996), para expressar um processo de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis, a partir de grandes conjuntos de dados. Não trivial significa que normalmente encontra-se complexidade no decorrer da execução do processo. Interativo, significa que o usuário poderá optar pela retomada em qualquer uma das etapas do processo. Iterativo significa que o processo pode repetir diversas vezes para se chegar a um resultado e, a cada vez, gera um resultado parcial, o qual será usado na vez seguinte.

2.1.1 Etapas do processo de DCBD

As etapas definidas por Fayyad em 1996 para o processo de descoberta de conhecimento em banco de dados são apresentadas na figura 1, que apresenta uma breve descrição das seis etapas (FAYYAD *et al.*, 1996).

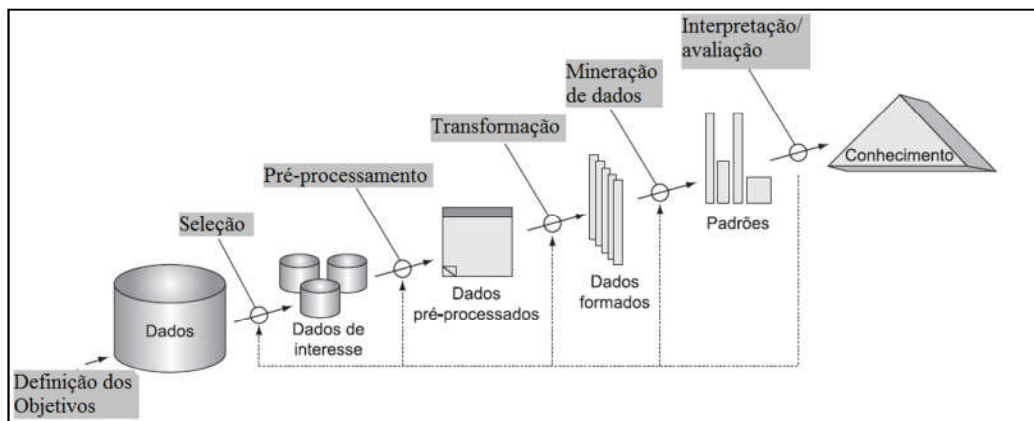


Figura 1 - Etapas do processo DCBD

Fonte: Adaptado de (FAYYAD *et al.*, 1996) p. 41

A primeira etapa é a definição dos objetivos: o usuário define o que ele deseja conhecer e quais são seus objetivos com o conhecimento a ser adquirido. Em seguida, na segunda etapa, é realizada uma seleção dos dados alvos de prospecção, criando um subconjunto dos dados selecionados. A terceira etapa consiste na limpeza dos dados, através de um pré-processamento, visando adequá-los para a etapa da mineração de dados. A quarta etapa visa uma transformação dos dados, a fim de diminuir o número de variáveis envolvidas no processo. Esta transformação trata de valores de atributos e não dos atributos em si. A quinta etapa é a mineração dos dados, a etapa mais importante do processo de descoberta de conhecimento em base de dados. O usuário opta por um método (classificação, regressão, agrupamento e associação) para fazer previsão ou descobrir padrões em sua base de dados. Após a realização da mineração dos dados, as previsões ou os padrões enumerados devem ser interpretados e avaliados, ou seja, deve ser realizado um trabalho de pós-processamento, a fim de descobrir o que consiste a nova descoberta.

Em busca dos melhores resultados, pesquisadores e profissionais na área de DCBD têm proposto diversos processos (*workflow* ou algumas abordagens simples de passo a passo) para maximizar as chances de sucesso na condução de projetos de

mineração de dados. Estes esforços têm levado a vários processos de DCBD, sendo o mais popular o *Cross-Industry Standard Process for Data Mining - CRISP-DM* (FERREIRA, 2005).

2.2 *Cross-industry standard proces for data mining - CRISP-DM*

O processo denominado *Cross-Industry Standard Proces for Data Mining - CRISP-DM* foi proposto em meados da década de 1990 por um consórcio europeu de empresas, para servir como metodologia padrão não proprietária para mineração de dados. Na figura 2, ilustra-se esse processo proposto, que é uma sequência de seis etapas, que inicia com um bom entendimento do negócio e da necessidade do projeto de mineração de dados e finaliza com a implementação da solução que satisfaz a necessidade especificada (CHAPMAN *et al.*, 2000).

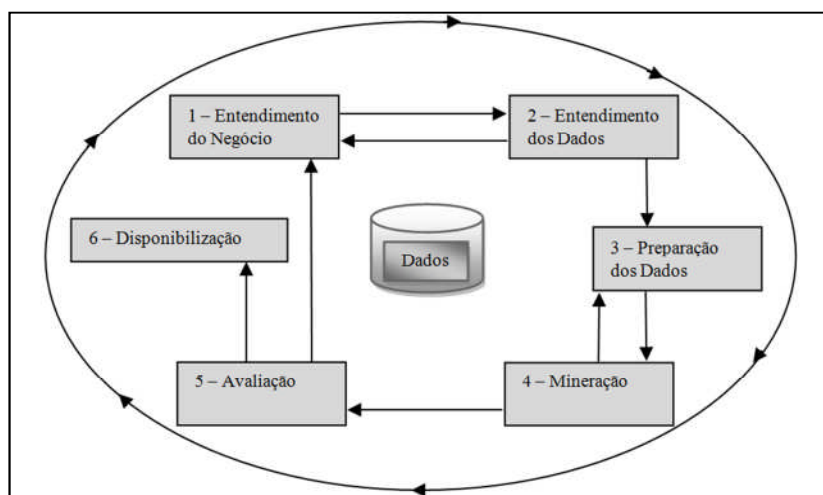


Figura 2 - Processo CRISP-DM
Fonte: Adaptada de (TURBAN *et al.*, 2010) p. 207

Mesmo que essas etapas sejam sequenciais por natureza, geralmente há uma grande realimentação, pelos seguintes motivos: a mineração de dados é impulsionada pela experiência e experimentação, em função da situação do problema; dependendo do

conhecimento e experiência do analista, o processo todo pode ser muito repetitivo e consumir muito tempo. A saída de uma etapa é a entrada da próxima, exigindo uma atenção especial ao realizar qualquer etapa, para não colocar todo o projeto em um caminho errado (CHAPMAN *et al.*, 2000).

2.2.1 Entendimento do negócio

O elemento principal de qualquer projeto de mineração de dados é entender completamente o que se espera do projeto, ou seja, os objetivos do projeto associados às exigências do negócio, com a finalidade de definir o problema de descoberta de conhecimento e traçar um plano preliminar para atingir tais objetivos. Muitas vezes, um projeto possui objetivos conflitantes e restrições que devem ser devidamente equilibradas. O objetivo é descobrir os fatores importantes, no início, que podem influenciar o resultado do projeto, para não encontrar respostas certas para as perguntas erradas, tais como: "quais são as características comuns dos clientes que perdemos para os nossos concorrentes nos últimos tempos?" ou "quais são os perfis típicos dos nossos clientes que agregam mais valor ao nosso negócio?" (CHAPMAN *et al.*, 2000).

2.2.2 Entendimento dos dados

Após a compreensão do negócio, a próxima atividade do processo é identificar os dados relevantes para o projeto, de diversas bases de dados disponíveis, através de técnicas estatísticas e gráficas para poder compreender melhor os dados, tais como simples resumos estatísticos de cada variável (por exemplo, para variáveis numéricas: a média, valor mínimo e máximo, mediana e desvio padrão estão entre as médias calculadas, enquanto que para as variáveis categóricas (qualitativas) o modo e a

tabela de frequência são calculados). Em seguida, faz-se a análise de correlação, diagramas de dispersão, histogramas e boxplots. Uma identificação e seleção das fontes de dados e das variáveis mais relevantes podem tornar mais fácil para os algoritmos de mineração de dados na descoberta dos padrões de conhecimento útil (SCHAEFFER, 2003).

2.2.3 Preparação dos dados

O objetivo desta etapa é preparar os dados identificados na etapa anterior a fim de serem submetidos aos métodos de mineração de dados. Comparando com as outras etapas do CRISP-DM esta fase consome mais tempo e esforço que as outras fases. Cerca de 80 por cento do tempo total gasto em um projeto de mineração de dados é na preparação dos dados. A razão para este enorme esforço despendido nesta etapa é o fato de que os dados do mundo real geralmente são incompletos (falta de valores de atributo, falta certos atributos de interesse, ou contendo apenas dados agregados), são impuros (contendo erros ou discrepância), e inconsistentes (contendo discrepâncias no código ou nomes). A figura 3 ilustra os quatro principais passos necessários para converter os dados do mundo real em conjuntos de dados exploráveis (TURBAN *et al.*, 2010).

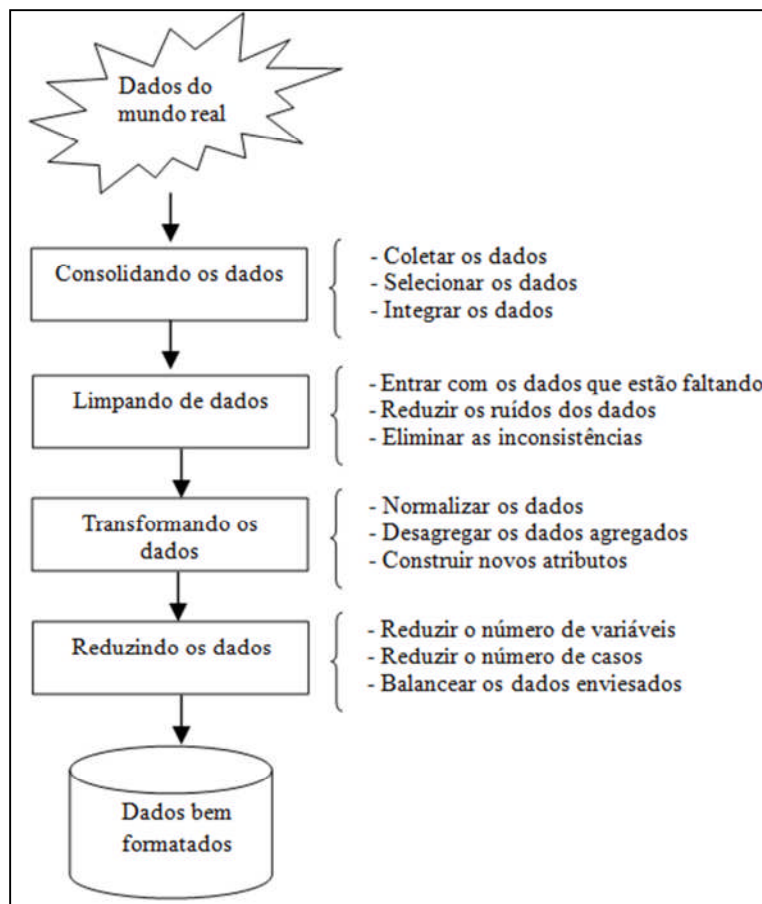


Figura 3 - Etapas de pré-processamento de dados
 Fonte: Adaptado de (TURBAN et al., 2010) p. 209

Na primeira fase de pré-processamento de dados, os dados relevantes são coletados a partir das fontes identificadas (realizado na etapa anterior), os registros necessários e as variáveis são selecionados (com base em um entendimento profundo, os dados desnecessários são retirados) e os registros provenientes de múltiplas fontes de dados são integrados.

Na segunda fase, os valores do conjunto de dados são identificados e tratados. Em alguns casos, encontram-se valores ausentes, resultado de anomalia no conjunto de dados, caso em que eles precisam ser preenchidos com um valor mais provável ou ignorados. Em outros casos, os valores ausentes são parte natural do conjunto de dados (por exemplo, a renda familiar campo é muitas vezes deixada sem resposta por pessoas que estão na camada superior de renda). Nesta etapa também se

retiram os valores discrepantes do conjunto de dados e as inconsistências que são os valores incomuns dentro de uma variável. Além disso, os dados devem ser tratados como um domínio de conhecimento (TAMILSELVI e GIFTA, 2011).

Na terceira fase de pré-processamento de dados, os dados são transformados para melhorar o processamento. Em alguns casos, os dados podem ser normalizados entre um determinado mínimo e máximo, a fim de atenuar a tendência da variável em possuir valores numéricos altos e ficar discrepante em relação a outras variáveis. Outra transformação que ocorre é a desagregação e ou a agregação. Em alguns casos, as variáveis numéricas são convertidas em valores categóricos (por exemplo, baixa, média e alta). Em outros casos, pode-se ter a necessidade de criar uma nova variável no banco de dados existente, a fim de ampliar a informação.

A fase final do pré-processamento de dados é a redução de dados. A redução pode ocorrer na quantidade de variáveis ou na quantidade de registros. Em alguns casos (por exemplo, processamento de imagens, projetos genoma e outros) a quantidade de variáveis pode ser bastante grande, devendo-se reduzir o número de variáveis, originando um subconjunto mais fácil de ser trabalhado e mais relevante. Em outros casos, os conjuntos de dados podem incluir milhões ou bilhões de registros. Apesar do poder de computação aumentar exponencialmente, o processamento de grande número de registros pode não ser prático ou viável. Nesse caso, cria-se uma amostra do subconjunto dos dados para análise. A suposição básica de amostragem é que o subconjunto de dados deve conter todos os padrões relevantes do conjunto de dados completo. Em ambos os casos, o novo subconjunto de dados deve refletir a essência do conjunto de dados original e não ser tendencioso para um subgrupo ou subcategoria.

2.2.4 Construindo um modelo

Nesta etapa, várias técnicas de modelagem são selecionadas e aplicadas a um conjunto de dados já preparados, a fim de atender às necessidades de negócios específicos. A construção do modelo também engloba a avaliação e análise comparativa dos diversos modelos construídos. Como não há um método ideal predefinido para uma tarefa de mineração de dados, deve-se usar uma variedade de tipo de modelo viável, juntamente com uma experiência bem-definida e uma estratégia de avaliação para identificar o "melhor" método para um determinado propósito (CARVALHO, 2009).

Mesmo para um único método ou algoritmo, uma série de parâmetros opcionais precisam ser calibrados para se obter os melhores resultados. Alguns métodos podem ter exigências específicas sobre a forma em que os dados devem ser formatados, muitas vezes fazendo com que o trabalho seja reavaliado com os dados para a etapa de preparação. Assim, dependendo da necessidade do negócio, a tarefa de mineração de dados pode ser preditiva ou descritiva. Cada uma dessas tarefas de mineração de dados pode usar uma variedade de métodos e algoritmos de mineração de dados (PETERMANN, 2006).

2.2.5 Teste e avaliação

Nesta etapa, os modelos desenvolvidos são avaliados quanto à sua precisão e generalidade. Essa etapa avalia e revisa as fases anteriores, para assegurar se o modelo escolhido (ou modelos) atende aos objetivos pré-definidos ou revelar informações adicionais relevantes aos objetivos do negócio. A fim de interpretar corretamente os padrões de conhecimento, é necessário recorrer a uma variedade de tabulações e

técnicas de visualização (por exemplo, tabelas dinâmicas, cruzamento de resultados, gráficos de pizza, histogramas, gráficos de caixa e gráficos de dispersão).

2.2.6 Disponibilização

O desenvolvimento e a avaliação dos modelos não são o fim do projeto de mineração de dados. Mesmo que a finalidade do modelo seja ter uma exploração simples dos dados, o conhecimento obtido a partir de tal exploração deve ser organizado e apresentado de uma forma que o usuário final possa entender e tirar proveito. Dependendo dos requisitos, a fase de divulgação pode ser tão simples como gerar um relatório ou tão complexa como a implementação de um projeto de divulgação dos resultados obtidos para toda empresa (RAJESH, 2011).

A etapa de divulgação também pode incluir atividades de manutenção para os modelos implementados. Como tudo sobre o negócio está em constante mudança, os dados que refletem suas atividades também estão mudando. Ao longo do tempo, os modelos construídos sobre os dados antigos podem se tornar obsoletos, irrelevantes ou enganosos. Assim, monitoramento e manutenção dos modelos são importantes, para que a mineração de dados integre como uma parte importante no dia a dia da organização.

2.3 Mineração de dados

A mineração de dados é um conjunto de ferramentas e técnicas que, por meio do uso de algoritmos, são capazes de explorar um conjunto de dados, extraíndo ou evidenciando padrões e predizendo conhecimento (FAYYAD *et al.*, 1996). O conhecimento extraído pode ser apresentado de diversas formas, tais como: agrupamentos, árvores de decisão, grafos, desde que sejam interpretados por seres

humanos. A importância e os resultados obtidos na aplicabilidade deste conhecimento, principalmente na formulação de estratégias de marketing para aumentar a lucratividade da organização, vêm despertando sua utilização em várias áreas, mesmo que não seja com fins lucrativos (CHIU e TAVELLA, 2008).

A mineração de dados não é uma disciplina nova, mas se torna nova com a interseção das disciplinas de estatística, inteligência artificial, reconhecimento de padrões, aprendizado de máquina, modelagem matemática, banco de dados e ciências da administração. Usando os avanços dessas disciplinas, a mineração de dados esforça-se para progredir na extração de informações e conhecimento úteis, a partir de grandes bases de dados. É um campo emergente que tem atraído a atenção de muitos pesquisadores em um curto espaço de tempo (TURBAN *et al.*, 2010).

Uma empresa que efetivamente utiliza as ferramentas e tecnologias de mineração de dados pode adquirir e manter uma vantagem competitiva estratégica perante ao seu concorrente. A mineração de dados oferece às organizações um ambiente indispensável à tomada de decisão para explorar novas oportunidades de transformação de dados em uma arma estratégica.

2.3.1 Os dados na mineração de dados

Os dados referem-se a uma coleção de fatos obtidos geralmente como resultado de transações, experiências, observações ou experimentos. Os dados podem consistir em letras e sons e, muitas vezes, são vistos como o menor nível de abstração, a partir do qual a informação e o conhecimento são derivados.

Muitas vezes, os dados devem ser preparados para que sejam mais apropriados à análise. Tendo como objetivo melhorar a qualidade dos dados, outros propósitos enfocam a modificação dos dados de modo que se adaptem melhor a uma ferramenta ou técnica específica de mineração de dados. Por exemplo, um dado contínuo, informando o comprimento “15 cm”, pode precisar ser transformado em um com categoria discreta, “curto”, “médio” ou “longo”, para que se possa aplicar uma determinada técnica (TAN et al, 2009).

No mais alto nível de abstração, pode-se classificar os dados como categorizados (qualitativos) e numéricos (quantitativos). Os categorizados podem ser subdivididos em nominal e ordinal, enquanto os numéricos podem ser subdivididos em intervalar e proporcional (TURBAN *et al.*, 2010).

Os dados categorizados ou qualitativos representam o rótulo de múltiplas classes usadas para dividir uma variável em grupos específicos. Exemplos de variáveis categóricas incluem raça, sexo, faixa etária e nível de escolaridade. Embora as duas últimas variáveis também possam ser consideradas de forma numérica, utilizando valores exatos para a idade e para nível educacional, muitas vezes é mais informativo categorizar tais variáveis em um número relativamente pequeno de classes ordenadas. Os dados categorizados também podem ser chamados de dados discretos, que representam um número finito de valores sem continuidade entre eles. Mesmo que os valores utilizados para as variáveis categóricas (ou discretas) sejam variáveis numéricas, estes números não são nada mais do que símbolos e não implicam a possibilidade de calcular valores fracionários (TURBAN *et al.*, 2010).

Os dados nominais contêm medidas de códigos simples atribuídos aos objetos como rótulos, que não são medidas. Por exemplo, a variável estado civil pode ser geralmente classificada como (1) solteiro, (2) casado e (3) divorciado. Dados nominais podem ser representados com valores binomiais de três ou mais valores possíveis (TAN *et al.*, 2009).

Os dados ordinais contêm código referente aos objetos ou eventos como rótulos que também representam a ordem de classificação entre eles. Por exemplo, a variável pontuação de crédito pode ser geralmente classificada como (1) baixo, (2) médio e (3) alta. Relações similares ordenadas podem ser vistas em variáveis tais como: faixa etária (criança, jovem, de meia-idade, idosos) e nível de escolaridade (colégio, faculdade, pós-graduação). Alguns algoritmos de mineração de dados, tais como regressão múltipla logística ordinal, levam em conta essas informações adicionais da ordem de classificação para construir um modelo (TURBAN *et al.*, 2010).

Dados numéricos representam os valores numéricos de uma variável específica. Exemplos de variáveis numéricas são a idade, o número de filhos, o valor da renda familiar, a quilometragem de uma viagem, temperatura, etc. Os valores numéricos podem ser reais ou inteiros. Os dados numéricos também podem ser chamados de dados contínuos e discretos. O número de portas de um carro, por exemplo, é uma variável discreta. Não faz sentido dizer que um carro tem 2,34 portas. Por esta razão, o número de portas possível em um carro não abrange todo o conjunto de números reais. Isso a torna uma variável discreta. Um exemplo de variável contínua seria o volume máximo do porta-malas. Esta variável pode assumir qualquer valor no conjunto dos números reais. Logo, é uma variável contínua. Os dados de medidas também podem ser intervalares, proporcionais ou circulares.

Os dados intervalares são variáveis que podem ser medidas em escalas de intervalo. Um exemplo mais comum são as escalas de temperaturas: a escala em *Celsius* divide-se em 100 partes iguais e a escala *Fahrenheit* divide-se em 180 partes iguais.

Dados proporcionais são variáveis de medida comumente encontradas na física, engenharia, etc. Exemplos de variáveis proporcionais são: a massa de um corpo ou objeto, o comprimento, altura, o ângulo de um plano, os contadores, a corrente elétrica e a quantidade monetária.

Dados do tipo data-hora, textos não estruturados, imagem e áudio precisam ser convertidos em alguma das representações categorizadas ou numéricas, antes que eles possam ser processados por algoritmos de mineração de dados. Os dados também podem ser classificados como estáticos ou dinâmicos (TURBAN *et al.*, 2010).

Alguns métodos de mineração de dados exigem um determinado tipo de dados. Fornecer tipos de dados incompatíveis pode levar a um modelo incorreto ou mais frequentemente, interromper o processo de desenvolvimento do modelo. Para minimizar esta situação, alguns softwares de mineração de dados aceitam uma mistura de variáveis numéricas e nominais e, internamente, fazem as conversões necessárias antes de processar os dados (CHIU e TAVELLA, 2008).

2.3.2 Tarefas de mineração de dados

No decorrer da história, o homem sempre aprendeu observando os padrões que ocorrem em seu meio, formulando hipóteses e testando-as para descobrir regras, e durante séculos estes padrões foram extraídos manualmente. Com a informatização de vários setores, os conjuntos de dados têm crescido em tamanho e complexidade,

tornando-se necessário o uso de ferramentas automatizadas que empregam algoritmos cada vez mais sofisticados, para identificar os padrões existentes nestes grandes conjuntos de dados.

As tarefas de mineração de dados consistem em construir modelos para identificar padrões entre os atributos apresentados no conjunto de dados. Estes modelos são representações matemáticas (uma simples relação linear e ou complexas relações não lineares) que identificam os padrões entre os atributos dos objetos descritos no conjunto de dados (por exemplo, clientes). Alguns destes padrões são descritivos e o objetivo é derivar padrões (correlações, tendências, grupos, trajetórias e anomalias) que resumam os relacionamentos subjacentes nos dados. Outros são preditivos, tendo como objetivo prever o valor de um determinado atributo baseado nos valores de outros atributos. Em geral, a mineração de dados busca identificar os seguintes tipos de padrões:

- **Associação:** é usada para descobrir padrões que descrevam características altamente associadas dentro dos dados, tais como cerveja e fraldas indo juntos em cestas de compras de clientes;
- **Previsão:** provê a natureza das ocorrências futuras de determinados acontecimentos com base no que já aconteceu no passado. Por exemplo, julgar se um paciente possui uma predisposição para uma determinada doença, baseado nos resultados de exames médicos;
- **Agrupamento:** identifica grupos naturais de coisas com base em suas características conhecidas, tal como, por exemplo, descobrir áreas do oceano que possuam um impacto significativo sobre o clima da Terra;

- **Relações sequenciais:** descobrir eventos em tempo ordenado, tal como prever se um cliente bancário com conta corrente abrirá uma conta poupança e uma conta de investimento dentro de um ano (TURBAN *et al.*, 2010).

Em busca destes padrões, a mineração de dados possui três categorias principais de tarefas: previsão, associação e agrupamento. Com base na forma em que os padrões são extraídos dos dados históricos, os algoritmos utilizados para esta finalidade em mineração de dados podem ser classificados como algoritmos de aprendizado supervisionados e algoritmos de aprendizado não supervisionados.

Os algoritmos de aprendizado supervisionado são aqueles que os dados de treinamento incluem tanto os atributos descritivos (ou seja, variáveis independentes ou variáveis de decisão), bem como o atributo de classe (variável de saída, ou seja, variável de resultado).

Em contraste com o aprendizado supervisionado, no não-supervisionado os dados de treinamento incluem somente os atributos descritivos. Na figura 4, mostra-se uma taxonomia simples das tarefas de mineração de dados, juntamente com os métodos de aprendizagem, e os algoritmos mais populares para cada uma das tarefas de mineração de dados.

Mineração de Dados	Método de aprendizagem	Algoritmos mais usados
Previsão	Supervisionado	Classificação e árvore de regressão, ANN, SVM, algoritmos genéticos
Classificação	Supervisionado	Árvore de decisão ANN/MLP, SVM, conjuntos aproximados, algoritmos genéticos
Regressão	Supervisionado	Regressão linear/não linear, árvore de regressão, ANN/MLP, SVM
Associação	Não supervisionado	<u>Apriori</u> , <u>Oner</u> , <u>ZeroR</u> , <u>Eclat</u>
Análise de Relação	Não supervisionado	Expectativa maximizar, algoritmo <u>Apriori</u> , <u>Graficos</u> baseados em correspondência
Análise de seqüência	Não supervisionado	Algoritmo <u>Apriori</u> , técnica FP- <u>Growth</u>
Agrupamento	Não supervisionado	<u>K-means</u> ANN/SOM
Análise de discrepância	Não supervisionado	<u>K-means</u> , Expectativa maximizar (EM)

Figura 4 - Taxonomia das tarefas de mineração de dados
 Fonte: Adaptada de (TURBAN *et al.*, 2010) p. 201

Tarefa de previsão

A tarefa de previsão é referida como o ato de dizer o futuro. Difere da simples adivinhação, tendo em conta as experiências, opiniões, e outras informações relevantes na condução da tarefa de prever. Dependendo da natureza do que está sendo previsto, a previsão pode ser relacionada mais especificamente como classificação, a qual é usada para variáveis discretas (previsão do tempo para amanhã, como “chuva” ou “sol”) e regressão, que é usada para variáveis contínuas (previsão da temperatura para amanhã, é um número real, tal como "35 ° C") (TURBAN *et al.*, 2010).

Tarefa de classificação

A classificação, ou indução supervisionada, talvez seja a tarefa mais comum de todas as tarefas de mineração de dados. O objetivo da classificação é a análise dos dados históricos armazenados em um banco de dados e a geração automática de um

modelo que possa prever o comportamento futuro. Este modelo induzido consiste em generalizações sobre os registros de um conjunto de dados, que ajudam a distinguir classes predefinidas. A esperança é que o modelo possa ser usado para prever as classes de outros registros não classificados e, o mais importante, prever com precisão real os eventos futuros (TURBAN *et al.*, 2010).

As ferramentas comuns de classificação incluem redes neurais ou *artificial neural network* (ANN) e árvores de decisão (aprendizagem de máquina), regressão logística e análise discriminante (a partir da estatística tradicional) e ferramentas emergentes, tais como conjuntos aproximativos, *Support Vector Machines* (SVM), e algoritmos genéticos e técnicas de classificação baseados em regras estatísticas (DUARTE, 2010).

As redes neurais envolvem o desenvolvimento de estruturas matemáticas (que se assemelham a redes neuronais biológicas no cérebro humano) que têm a capacidade de aprender com experiências passadas, apresentadas sob a forma de conjuntos de dados bem estruturados. Eles tendem a ser mais eficazes quando o número de variáveis envolvidas é muito grande e as relações entre elas são complexas e imprecisas. As redes neurais têm desvantagens e vantagens (MARTINS, 1998). Por exemplo, geralmente é muito difícil fazer uma boa lógica para as previsões feitas por uma rede neural. Além disso, as redes neurais tendem à necessidade de um aprendizado considerável. Infelizmente, o tempo necessário para o aprendizado tende a aumentar exponencialmente em relação ao volume dos dados, e, em geral as redes neurais não podem ser treinadas em bases de dados muito grandes. Esses e outros fatores têm limitado a aplicabilidade das redes neurais em domínios ricos de dados.

As árvores de decisão classificam os dados em um número finito de classes com base nos valores das variáveis de entrada. Esta técnica é usada para resolver o problema de classificação, fazendo uma série de questões cuidadosamente organizadas sobre os atributos do registro de teste. Cada vez que fornece uma resposta, um questionamento seguinte é feito, até que se chegue a uma conclusão sobre o rótulo da classe do registro. As questões e suas respostas possíveis podem ser organizadas na forma de uma árvore de decisão, formando uma estrutura hierárquica composta de nodos e arestas direcionadas. A figura 5 ilustra uma árvore de decisão para classificar três variáveis (TAN *et al.*, 2009).

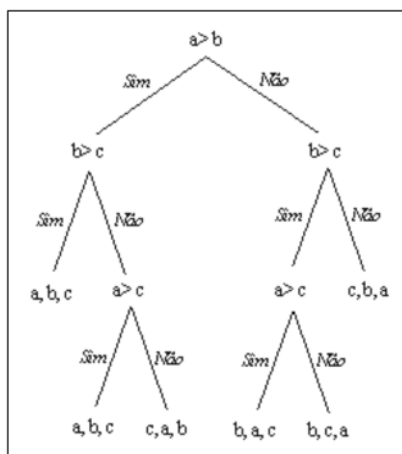


Figura 5 - Árvore de decisão
Fonte: Adapta de (CARVALHO, 2009)

Os nodos de uma árvore de decisão podem ser do tipo raiz, internos e folha ou terminais. Um nodo raiz é aquele em que somente arestas saem dele, enquanto que para os nodos internos entram e saem arestas e, para os nodos folha, somente chegam arestas. Os nodos raiz e internos possuem condições de testes de atributos para separar registros que possuam características diferentes, enquanto os nodos de folha recebem um rótulo de classe, que são usadas para atribuir aos registros que chegam a estes nodos de folha.

A partir de um determinado conjunto de atributos, a construção de árvores de decisão é exponencial, ou seja, se em um conjunto de atributos há dois atributos, poderão ser construídas até quatro árvores diferentes. Porém, algumas árvores são mais precisas que outras e encontrar a árvore ótima é um esforço computacionalmente inviável, por causa do tamanho exponencial do tempo de pesquisa. Entretanto, há alguns algoritmos desenvolvidos para induzir a uma árvore de decisão razoavelmente precisa, embora não perfeita, em uma razoável quantidade de tempo. Os algoritmos de Hunt, ID3, C4.5 e CART são uns desses algoritmos de indução de árvores de decisão (NAYAK e DEVI, 2011).

Tarefa de agrupamento

É o agrupamento de partições de um conjunto de coisas (por exemplo, objetos, eventos etc., presentes dentro da estrutura de um conjunto de dados) em segmentos (ou agrupamentos naturais), cujos membros compartilham características semelhantes. Ao contrário da classificação, no agrupamento os rótulos das classes são desconhecidos. O algoritmo selecionado percorre o conjunto de dados, identificando as semelhanças dos elementos do conjunto com base em suas características e os agrupamentos são estabelecidos. Os agrupamentos são determinados usando um algoritmo do tipo heurístico. Diferentes algoritmos podem ser utilizados para realizar agrupamento no mesmo conjunto de dados. Antes de colocar em uso real a técnica de agrupamento, pode ser necessária a intervenção de um especialista para interpretar, e potencialmente modificar a sugestão de agrupamento proposto pelo algoritmo de agrupamento utilizado. Depois que os agrupamentos são identificados, eles podem ser usados para classificar e interpretar os novos dados (TURBAN *et al.*, 2010). Na figura 6, ilustram-se os grupos encontrados pela tarefa de agrupamento.

A técnica de agrupamento inclui a otimização. O objetivo do agrupamento é a criação de grupos para que os membros dentro de cada grupo tenham similaridade máxima e os membros em grupos diferentes de similaridade mínima. As técnicas mais utilizadas incluem agrupamento k-médias (a partir de estatísticas) e mapas auto-organizáveis ou *self-organizing maps* (SOM), que são uma arquitetura de rede neural originalmente desenvolvida por Kohonen em 1982 (RAJESH, 2011).

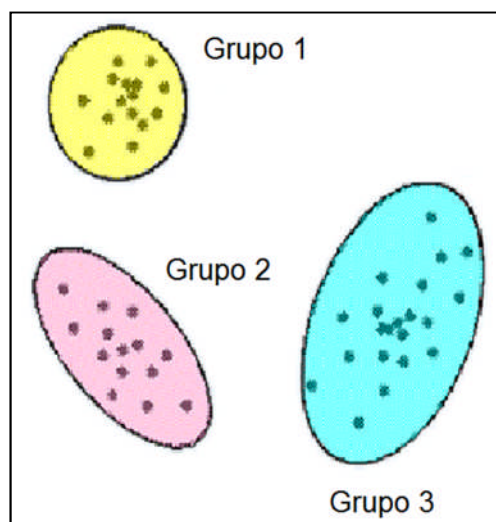


Figura 6 - Grupos encontrados pela tarefa de agrupamento
Fonte: Adaptada de (SCHNEIDER, 2003)

As empresas costumam utilizar eficazmente os seus sistemas de mineração de dados, para realizar a segmentação do mercado com a análise de agrupamento. A análise de agrupamento é um meio de identificação de classe do item e para que os itens do mesmo grupo que tenham mais em comum uns com os outros do que com os itens de outros grupos. A sua aplicabilidade pode ser usada na segmentação de clientes, direcionando a comercialização de produtos adequados para os segmentos na hora, no formato e preço certo. A análise de agrupamento também é usada para identificar grupo natural de eventos ou objetos, de modo que um conjunto de características comuns desses grupos possa ser identificado para descrevê-los (TURBAN *et al.*, 2010).

Tarefa de associação

A associação é uma abordagem baseada na descoberta de relacionamentos importantes escondidos em conjuntos de dados. Os relacionamentos descobertos podem ser representados na forma de regras. Por exemplo, após a análise tem-se a seguinte regra:

- “Dos Clientes que compram o produto A, que ocorre em 47% das vezes, estes também, comprarão o produto B. Isto ocorre em 25% das vezes.”

Uma regra de associação possui os seguintes componentes: condição, resultado, fator de suporte, fator de confiança e *lift* (SCHNEIDER, 2003). Os componentes identificados no exemplo dado são:

- Condição: “clientes que compram produto A”;
- Resultado: “comprarão o produto B”;
- Fator de suporte: “ocorre em 47% das vezes”, é o número de vezes que determinado item ou conjunto de itens aparecem em transações distintas, relacionadas com o número total de transações operadas;
- Fator de confiança: é o número de transações que suportam o resultado em relação ao número de transações que suportam a condição: “Isto ocorre em 25% das vezes”;
- *Lift*: é uma estatística adicional para verificação de dependência entre as variáveis envolvidas em uma determinada regra de associação. Quanto maior for o *lift*, maior será a dependência entre os atributos e, provavelmente, maior será o valor da regra.

O objetivo desta técnica é contabilizar as ocorrências associadas às combinações possíveis entre itens. Primeiramente, identificam-se as transações, procurando quantificar as ocorrências de um determinado item, devido à presença de um ou de outros na transação. Pode-se considerar que ela trabalha em um primeiro momento, buscando encontrar o fator de suporte para, posteriormente, testar todas as regras, verificando o fator de confiança.

A complexidade da utilização dessa regra depende do número de variáveis envolvidas. Se a procura for um item no resultado e outro na condição, então a performance é melhor do que se a procura for por mais de um item, tanto no resultado quanto na condição. Na figura 7, ilustram-se os dados ao centro, associados às regras das laterais.

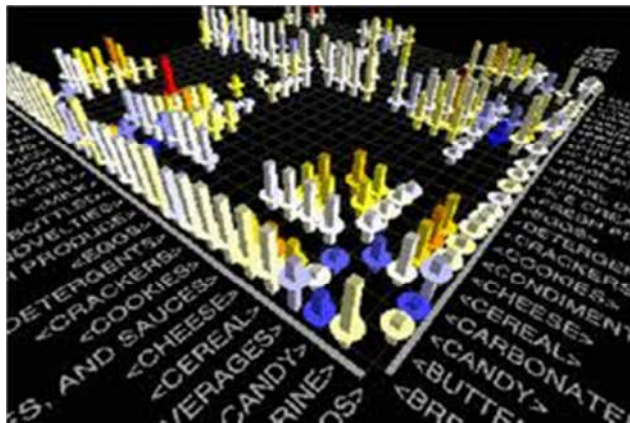


Figura 7 - Dados associados a uma regra
Fonte: Adaptada de (TAMILSELVI e GIFTA, 2011)

Graças às tecnologias automatizadas de coleta de dados, tais como *scanners* de código de barras, o uso de regras de associação para descobrir regularidades entre os produtos nas operações de grande escala registradas nos supermercados tornou-se uma tarefa de descoberta de conhecimento comum no setor de varejo. No contexto da

indústria de varejo, a mineração de regras de associação é frequentemente chamada de análise de cesta básica (TURBAN *et al.*, 2010).

Os algoritmos Apriori, FP-Growth, OneR, ZeroR e Eclat entre outros foram desenvolvidos com o objetivo de gerar regras de associação entre itens de dados. O mais utilizado é o Apriori (SCHNEIDER, 2003).

2.4 Outros processos de DCBD

Para ser aplicado com êxito, um projeto de DCBD deve ser encarado como um processo que segue uma metodologia padronizada e não como um conjunto de ferramentas e técnicas automatizadas. Além de CRISP-DM, há outra metodologia conhecida desenvolvida pelo Instituto SAS, chamado SEMMA (NISBET, 2009). A sigla SEMMA significa Selecionar (*Sample*), Explorar (*Explore*), Modificar (*Modify*), Modelar (*Model*) e Avaliar (*Assess*) – SEMMA.

A principal diferença entre o CRISP-DM e SEMMA é que o CRISP-DM tem uma abordagem mais abrangente, incluindo a compreensão do negócio e dos dados relevantes para projetos de mineração de dados, enquanto o SEMMA, implicitamente, assume que os objetivos do projeto de mineração de dados foram identificados e compreendidos previamente.

Analisando a enquete sobre qual processo os visitantes do site <www.kdnuggets.com> usam para realizar a mineração de dados, que foi aplicada em 2007, 2004 e 2002, a metodologia mais usada foi a CRISP-DM, seguida por SEMMA. Outra observação é que o uso da metodologia SEMMA apresenta um crescimento no período analisado (KDNUGGETS, 2007).

2.5 Aplicação de mineração de dados

A mineração de dados se tornou uma ferramenta popular na resolução de muitos problemas complexos nas organizações. A sua utilização tem sido muito bem sucedida e útil em muitas áreas, algumas das quais são apresentadas nesta seção. O objetivo de uma empresa ao utilizar mineração de dados é a resolução de um problema urgente ou explorar uma oportunidade de negócio emergente, a fim de criar uma vantagem competitiva sustentável (TURBAN *et al.*, 2010).

A mineração de dados ajuda o setor bancário nas seguintes possibilidades:

- Detectar o processo de pedido de empréstimo prevendo com precisão os prováveis inadimplentes;
- Detectar fraude de cartão de crédito e transações bancárias on-line;
- Identificar formas de maximizar as vendas aos clientes de produtos e serviços que eles são mais propensos a comprar;
- Otimizar o retorno de caixa através da previsão exata do fluxo de caixa em entidades bancárias (por exemplo, caixas eletrônicos, agências bancárias).

No setor de varejo, mineração de dados é usada para :

- Prever volumes precisos de vendas no varejo a fim de determinar os níveis de estoque correto;
- Identificar as relações entre as vendas de diferentes produtos (com análise de cesta básica) para melhorar o *layout* da loja na exposição dos produtos relacionados a fim de otimizar as vendas;

- Prever o consumo dos diferentes tipos produtos (com base nas condições sazonais e ambientais) para otimizar a logística e, conseqüentemente, maximizar as vendas;
- Descobrir padrões interessantes na circulação dos produtos (especialmente para os produtos que têm uma vida útil limitada, pois eles são propensos à expiração, perecibilidade e contaminação) em uma cadeia de suprimentos, através da análise dos dados coletados pelos sensores de RFID (*Radio-Frequency IDentification*).

As indústrias usam a mineração de dados para:

- Prever falhas de máquinas antes que elas ocorram através da utilização de dados sensoriais;
- Identificar as anomalias e semelhanças nos sistemas de produção para otimizar a capacidade de produção;
- Descobrir novos padrões para identificar e melhorar a qualidade do produto.

Os corretores e os comerciantes do setor imobiliário usam a mineração de dados para:

- Prever quando e quanto os preços dos imóveis vão mudar;
- Prever o intervalo e a direção da flutuação de valores;
- Avaliar o efeito de questões específicas e eventos sobre os movimentos do mercado global;

- Identificar e prevenir atividades fraudulentas no comércio imobiliário.

A indústria de seguros (KAHANE *et al.*, 2007) utiliza técnicas de mineração de dados para:

- Prever a quantidade de uso da cobertura médica para um melhor planejamento de negócios;
- Determinar a taxa de opcional com base na análise dos pedidos e dados de clientes;
- Prever quais consumidores estão mais propensos a atividades fraudulentas e à inadimplência financeira.

A mineração de dados é usada na informática para:

- Prever falhas de disco rígido bem antes que elas realmente ocorrem;
- Identificar e filtrar conteúdos indesejados da Web e mensagens de e-mail;
- Detectar e prevenir segurança nas redes de computadores;
- Identificar os softwares potencialmente inseguros.

A mineração de dados também tem inúmeras aplicações militares, como por exemplo:

- Prever o custo do movimento de militares e equipamentos em uma manobra;

- Prever ações do adversário e, conseqüentemente, desenvolver estratégias mais bem sucedidas para os enfrentamentos militares;
- Prever o consumo de recursos para um melhor planejamento.

A mineração de dados tem uma variedade de usos na indústria de viagens. É usada com sucesso para:

- Prever as vendas dos diferentes serviços (tipos de assento nos aviões, os tipos de quartos em hotel/resort, os tipos de carros em locadoras de automóveis), a fim de prestar um serviço de forma otimizada para maximizar as receitas, em função do tempo de diferentes operações (vulgarmente conhecido como “*yield management*” gestão de receitas);
- Prever a demanda em diferentes localidades para melhor alocar os recursos organizacionais;
- Identificar os clientes mais rentáveis e prestar-lhes serviços personalizados para sua fidelização.

A mineração de dados tem um número de aplicações na área de planos de saúde. Ela é usada para:

- Identificar as pessoas sem seguro de saúde e os fatores subjacentes a este fenômeno indesejável;
- Identificar as relações custo-benefício-contrato entre os diferentes tratamentos para desenvolver estratégias mais eficazes;

- Prever o tempo de procura em locais diferentes para otimizar o serviço alocar recursos organizacionais.

O uso de mineração de dados em medicina é encarado como um complemento para a pesquisa médica tradicional, determinando o que é essencialmente clínico e biológico na natureza (PATTANAIK e GHOSH, 2010). A análise de mineração de dados pode:

- Identificar novos padrões para melhorar a capacidade de sobrevivência dos pacientes com câncer;
- Prever as taxas de sucesso de pacientes submetidos a transplante de órgãos para desenvolver melhores políticas de doações de órgãos;
- Identificar as funções dos diferentes genes no cromossomo humano (conhecida como genômica);
- Descobrir as relações entre os sintomas e doenças (como doenças e tratamentos bem sucedidos) para ajudar os médicos com informações para tomar decisões corretas em tempo hábil.

A mineração de dados é utilizada com sucesso pela indústria de entretenimento para:

- Analisar os dados do telespectador para decidir quais programas que serão mostrados em horário nobre e como maximizar os lucros, por saber onde inserir publicidade;

- Prever o sucesso financeiro de filmes antes que eles sejam produzidos, para tomar decisões de investimento e otimizar os retornos;
- Prever a demanda em diferentes locais e tempos, para programar eventos de entretenimento, otimizando a alocação de recursos.

A mineração de dados tem inúmeras aplicações para segurança pública, sendo frequentemente usada para:

- Identificar padrões de comportamentos terroristas ou criminosos;
- Descobrir padrões de criminalidade (por exemplo, locais, horários, comportamentos criminosos e outros atributos relacionados) para ajudar a resolver casos criminais em tempo hábil;
- Prever e eliminar possíveis ataques químicos e biológicos na infraestrutura crítica da nação, por meio da análise de dados coletados através de sensoriais especiais para esses fins;
- Identificar e bloquear ataques maliciosos de informação na infraestrutura crítica da nação (muitas vezes chamado à guerra de informação).

Um exemplo de uso da mineração de dados em esporte é o projeto desenvolvido pela Associação de Basquete Nacional dos Estados Unidos (NBA). Consiste em uma aplicação que coleta as observações registradas pela equipe técnica do time com seus respectivos instantes, em um computador, durante as partidas. Após essas coletas, uma aplicação de mineração de dados é utilizada para descobrir padrões interessantes nos dados dos jogos de basquete. As interpretações de padrões são

facilitadas pelo recurso que o usuário tem em relacionar os vídeos aos padrões, dando subsídios para a comissão técnica escolher um time mais adequado.

A gestão de relacionamento com o cliente (*customer relationship management* – CRM) é uma maneira diferenciada de realizar o marketing tradicional. O objetivo de CRM é criar relacionamento de um-para-um (*one-on-one*) entre uma empresa e seus clientes, através do desenvolvimento de uma compreensão profunda de suas necessidades e desejos. Este relacionamento é construído ao longo do tempo, por meio de uma variedade de operações (por exemplo, informações de produtos, vendas, pedidos de serviços, chamadas de garantia, etc.). Elas acumulam enormes quantidades de dados. Quando combinado com os atributos demográficos e sócioeconômicos, tornam os dados ricos em informações. Utilizando a mineração de dados a empresa pode:

- Identificar os compradores com potencial para comprar novos produtos /serviços (isto é, o perfil do cliente);
- Compreender as causas raízes das insatisfações do cliente, a fim de melhorar a retenção de clientes (ou seja, promover a fidelização do cliente);
- Identificar a frequência de compra a fim de torná-la a menor possível;
- Identificar os clientes mais rentáveis e as suas necessidades preferenciais para fortalecer os relacionamentos e para maximizar as vendas.

Finalmente, na educação, há pesquisas em mineração de dados que se preocupam com o desenvolvimento de métodos para explorar os tipos de dados exclusivos que vêm de ambiente educacional; também se utilizam esses métodos para entender melhor os alunos e os contextos que circundam o aprendizado (MINING, 2011). Algumas aplicabilidades de EMD (*Educational Data Mining*) são:

- Identificar o perfil do aluno;
- Melhorar os modelos de conhecimento do estudante durante seu aprendizado nas diferentes áreas, tais como: língua estrangeira, geometria, química, física e outras, resultando na redução considerável do tempo gasto pelos alunos para desenvolver suas habilidades acadêmicas;
- Viabilizar a expansão do conhecimento científico relacionado aos estados emocionais do aluno (e.g. motivado, frustrado, confuso, etc);
- Utilizar softwares educacionais capazes de apontar os fatores que influenciam o comportamento do aluno e identificar aspectos sutis, muitas vezes imperceptíveis, dando subsídios para reformular os softwares de ensino a distância.

2.6 Outros trabalhos de mineração de dados na educação

Dois trabalhos tratam da mesma temática dentro da educação: a evasão de alunos do ensino superior. O título do primeiro é “Utilização da extração de conhecimento de bases de dados para identificar padrões de evasão de alunos de graduação da UNICAMP” (MARTINS, 1998). E o segundo é “Evasão no ensino superior: um estudo utilizando a mineração de dados como ferramenta de gestão do

conhecimento em um banco de dados referente à graduação de engenharia.”. A proposta do presente trabalho também está ligada diretamente com o tema dos dois trabalhos. Ao encontrar o perfil do aluno aplicado, os gestores do curso poderão desenvolver ações de marketing para atrair candidatos com o perfil encontrado, evitando-se o baixo rendimento e possível evasão do curso.

Nenhum deles utiliza a metodologia CRISP-DM como processo de DCBD. A ferramenta de mineração de dados utilizada pelo primeiro trabalho é a ferramenta C4.5 e o outro utiliza a ferramenta da Oracle ORACLE Data Miner. A ferramenta empregada neste trabalho é o software WEKA. A tarefa predominante é a classificação entre os trabalhos utilizando árvore de decisão.

3 Material e métodos

Neste capítulo, apresentam-se o material utilizado e os procedimentos realizados no desenvolvimento desta pesquisa. A metodologia empregada para a descoberta de conhecimento em banco de dados é a CRISP-DM, constituída por seis etapas. Algumas ferramentas computacionais são utilizadas como auxílio na aplicação dos métodos. O material empregado são os dados dos alunos do Curso de Sistemas de Informação da Unidade Universitária Ciências Exatas e Tecnológicas (UnUCET) da Universidade Estadual de Goiás.

3.1 Problema da pesquisa

O problema a ser explorado refere-se à identificação de padrões, que representam o perfil do aluno que não foi reprovado em nenhuma disciplina durante sua graduação no Curso de Sistemas de Informação da UnUCET, identificado e denominado aluno aplicado. A identificação do perfil do aluno é encontrada através da aplicabilidade do conhecimento apropriado do processo de descoberta de conhecimento em bases de dados.

Dois fatos influenciaram a escolha do material a ser estudado. O primeiro é que, a partir de 2008, durante as reuniões do colegiado do curso, o principal assunto em pauta tem sido sobre o aumento na quantidade de alunos dependentes ou desinteressados pelas disciplinas, gerando baixo rendimento acadêmico, levando a acreditar que alguns alunos não se adequam ao perfil do curso. Outro fato é que o número de candidatos optantes pelo curso de SI-UnUCET tem diminuído a cada ano. No projeto pedagógico de 2009 (BRITO *et al.*, 2009), justifica-se este fato pelo

surgimento de novos cursos na área de informática em Anápolis e Goiânia. A proposta deste trabalho é identificar o perfil do aluno aplicado, empregando os dados dos ingressantes de 2002 a 2007. O perfil encontrado é fundamental para propor mudanças no processo seletivo do curso e promover ações de marketing para atrair candidatos com as características encontradas.

3.2 Hipótese

A aplicação do processo de descoberta do conhecimento em banco de dados utilizando a tarefa classificação poderá identificar o perfil do ingressante do Curso de SI da UnUCET, por meio dos dados acadêmicos e do questionário socioeconômico e cultural.

3.3 Metodologias aplicadas

O estudo está limitado na utilização da metodologia CRISP-DM para descobrir conhecimento em banco de dados. Na etapa de mineração de dados é empregada a tarefa de classificação usando algoritmos de árvores de decisão, através do software WEKA (*Waikato Environment for Knowledge Analysis*). A identificação de perfil é própria de classificação. Os dados acadêmicos e do vestibular dos alunos foram disponibilizados em formato de planilha eletrônica, importados para um sistema de gerenciamento de banco de dados (SGBD), onde foram feitas todas as limpezas e adaptações dos dados que são utilizados pelo WEKA.

3.4 Materiais usados

Para realizar o estudo de caso, os dados foram disponibilizados pela Pró-Reitoria de Graduação da UEG e Núcleo de Seleção da UEG, atendendo à solicitação feita pelo autor do trabalho. As ferramentas computacionais foram: SGBD Microsoft Office Access 2007 para limpeza e preparação dos dados e o WEKA para realizar a etapa de mineração de dados.

3.4.1 Ferramenta WEKA

WEKA (*Waikato Environment for Analysis Knowledge*) é uma ferramenta desenvolvida por pesquisadores da Universidade de *Waikato* na Nova Zelândia, sendo uma das mais populares entre a comunidade de mineração de dados. Ela possui código aberto, interface amigável, agrega um conjunto de algoritmos de classificação, regras de associação, regressão, pré-processamento e agrupamento, todos implementados em JAVA.

Todas as técnicas do Weka são baseadas na suposição de que os dados estão disponíveis como um arquivo simples ou relação, onde cada registro é descrito por um número fixo de atributos numéricos ou nominais. A operacionalidade pode se dar através de linhas de comando ou usando interface gráfica, que possui quatro botões: *Explorer*, *Experimenter*, *KnowledgeFlow*, *SimpleCLI*, descritos abaixo.

- *SimpleCLI*: Fornece uma interface de linha de comando simples que permite a execução de comandos WEKA. Geralmente é utilizado em sistemas operacionais que não possuem uma interface própria;

- *KnowledgeFlow*: Este ambiente contém as mesmas funções do Explorer, porém com uma interface *drag-and-drop* (arrastar e largar), promovendo aprendizagem incremental;
- *Experimenter*: Um ambiente que serve para realização de experiências e testes estatísticos entre os modelos encontrados;
- *Explorer*: Lugar composto por seis guias para explorar os dados. *Preprocess* é a primeira a ser habilitada para o usuário escolher e modificar os dados que serão submetidos aos algoritmos de mineração. Somente após inserir os dados é que as guias *Classify*, *Cluster*, *Associate*, *Select attributes* e *Visualize* serão ativadas e estarão prontas para serem usadas.

Para inserir os dados no WEKA, é necessário colocá-los em um formato que seja entendido por ele. O seu método preferido para carregar os dados é o formato de arquivo de atributo-relação (ARFF), onde é possível definir o tipo de dados, intervalo, nome da coluna e o que cada uma possui. Na guia *Preprocess*, é possível realizar filtros e visualizar informações estatísticas dos dados selecionados. Na figura 8, ilustra-se a leitura de arquivo arff.

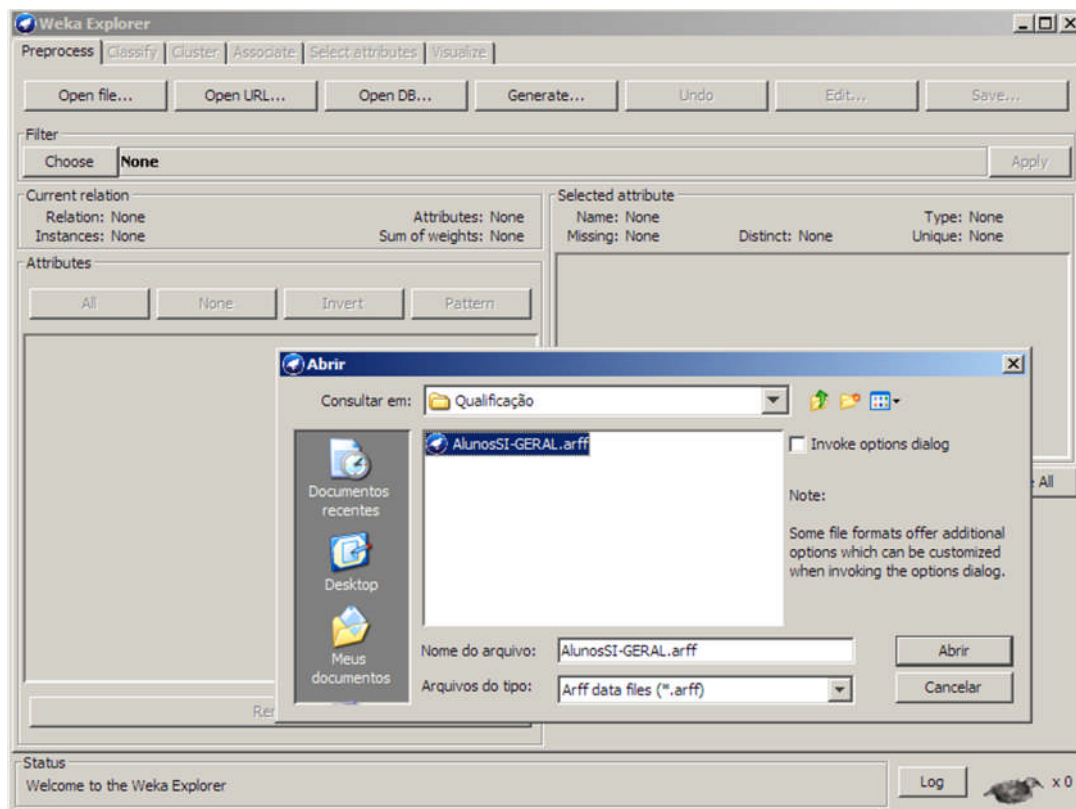


Figura 8 - Tela do WEKA lendo um arquivo arff

As tarefas de classificação e regressão são providas pela aba *Classify*, onde é possível escolher o algoritmo J48, que é uma versão implementada no WEKA do algoritmo C4.5 desenvolvido por Ross Quinlan (WITTEN *et al.*, 2011). Este algoritmo constrói uma árvore de decisão, usando a abordagem *top-down*, em que o atributo mais significativo, ou seja, o mais generalizado, quando comparado a outros atributos do conjunto, é considerado raiz da árvore. Na sequência da construção, o próximo nó da árvore será o segundo atributo mais significativo, e assim sucessivamente, até gerar o nó folha, que representa o atributo alvo da instância. Após a construção, o algoritmo inicia um processo de poda, a fim de reduzir o excesso de ajustes (*overfitting*) aos dados de treinamento (BASGALUPP, 2010). Ao clicar no botão *Choose* mostrado na figura 9, sobre o nome J48 aparece um formulário para alterar os parâmetros padrões deste algoritmo. Na figura 10, mostra-se o formulário.

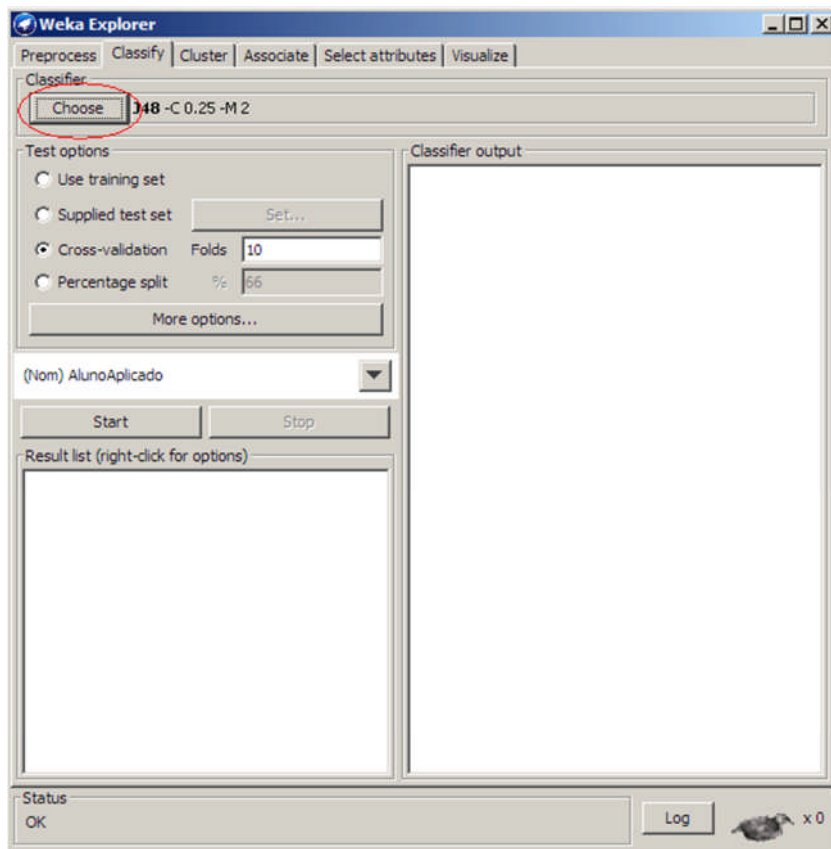


Figura 9 - Botão que lista os algoritmos de classificadores disponíveis.

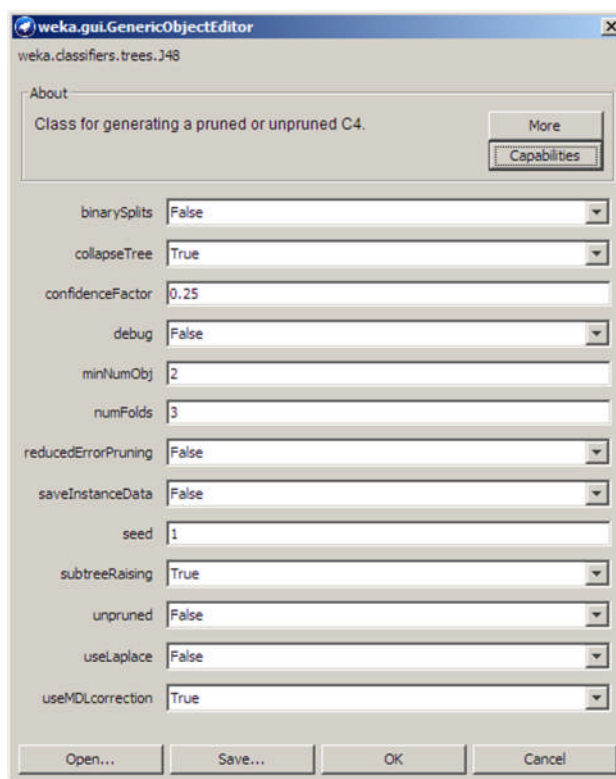


Figura 10 - Parâmetros do classificador J48

As opções são:

- *BinarySplits* – Usar divisor binário em atributos nominal ao construir árvores;
- *CollapseTree* – Remover atributos que não reduzem erro de treinamento;
- *ConfidenceFactor* – Fator de confiança utilizado na poda (valores menores implicam em mais poda);
- *Debug* – Se verdadeiro, o classificador emite informações adicionais na console;
- *MinNumObj* – Número mínimo de instâncias por folha;
- *NumFolds* – Determina a quantidade de dados utilizados para reduzir erro na poda. Uma dobra é utilizada na poda, o resto para o crescimento da árvore;
- *ReducedErrorPruning* – Reduzir poda, se houver erro;
- *SaveInstanceData* – Salvar os dados de treinamento para visualização;
- *Seed* – Quantidade de descendência de dados para randomizar a redução de erros na poda;
- *SubtreeRaising* – Considerar a operação elevação de subárvore na poda;
- *Unpruned* – Se a poda é realizada;
- *UseLaplace* – Se a contagem de folhas for suavizada com base em um lugar;

- *UseMDLcorrection* – Usa correção MDL quando encontrar divisão de atributos numéricos.

Após alterar os parâmetros desejados, escolhe-se o atributo que determine a classe, caso seja classificação e executa-se o algoritmo escolhido. Posteriormente, é disponibilizado um relatório com os resultados da classificação, divididos nas seguintes seções:

- *Run information*: Mostra as informações utilizadas pelo modelo de aprendizagem tais como: nome da relação, instâncias, atributos e tipo de teste que foram utilizados no processo;
- *Classifier model (full training set)*: Mostra uma representação textual do modelo de classificação que foi produzido;
- *Summary*: Mostra um resumo das estatísticas: a precisão do modelo quanto a sua capacidade de prever a classe predeterminada;
- *Detailed Accuracy By Class*: Apresenta a precisão de cada classe do classificador de previsão;
- *Confusion Matrix*: Exibe a matriz confusão: o número de classificações corretas em oposição às classificações perdidas para cada classe.

4 Descrição do estudo de caso

4.1 O Curso de Sistemas de Informação da UnUCET-UEG

O curso de Bacharelado em Sistemas de Informação tem por objetivo a formação de recursos humanos para serem empreendedores ativos e éticos, com os fins devidos para desenvolver, implantar e gerenciar sistemas de informação, de modo a atender às demandas das organizações e da sociedade, propiciando aos seus alunos formação em Ciências da Computação, em Administração de Empresas e Sistemas de Informação, enfatizando aspectos teóricos e práticos (SBC, 2003).

Em 2002, iniciou-se o Curso de Sistemas de Informação da UnUCET, substituindo o Curso Superior de Tecnologia em Processamento de Dados, que teve início em 1986. A proposta inicial do Curso de Sistemas de Informação era atender às necessidades sociais e de mercado da época, já que o Curso de Tecnologia em Processamento de Dados necessitava de algumas atualizações e adequações. Além disso, havia as orientações do Ministério da Educação para que fossem substituídos os cursos de natureza tecnólogo para outros da área de bacharelado em computação e informática e também recomendações da Sociedade Brasileira da Computação (SBC) (SBC, 2003).

O Conselho Universitário da UEG, através da Resolução CsU 34/2001, aprovou o funcionamento do Curso de Sistemas de Informação na Unidade Universitária de Ciências Exatas e Tecnológicas, iniciando o processo de extinção do Curso de Tecnologia em Processamento de Dados. E a Resolução CsU 029/2005 aprovou e convalidou o Projeto Pedagógico Unificado do Curso de Sistemas de

Informação, com 3.280 horas/aulas, em algumas unidades da UEG, inclusive na UnUCET. Na tabela 1, ilustra-se a matriz curricular de 2002. O curso foi reconhecido pelo Conselho Estadual de Educação através da portaria 1.436 de 10 de novembro de 2006, D.O.E. em 13/11/2006, até a data 31/12/2009.

Tabela 1 – Matriz Curricular 2002

Série	Disciplina	Carga Horária Semanal
1ª	Álgebra Linear	2
	Cálculo Diferencial e Integral	4
	Eletricidade e Lógica Digital	2
	Inglês Instrumental	2
	Introdução a Computação	2
	Linguagem e Técnicas em Programação	4
	Organização, Sistemas e Métodos	4
	Princípios de Sistemas de Informação e Teoria da Computação	4
2ª	Análise de Sistemas	4
	Arquitetura de Computadores	2
	Bancos de Dados	4
	Estruturas de Dados	4
	Metodologia da Pesquisa Científica	2
	O Profissional de Informática e a Sociedade	2
	Programação I	4
	Sistemas Operacionais	2
3ª	Contabilidade e Econômica para computação	4
	Engenharia de Software	4
	Novas Aplicações em Sistemas de Informações	4
	Programação II	4
	Projeto de Graduação I	4
	Redes de Computadores	4
	Conceitos de Inteligência Artificial	2

4ª	Direito e Ética em Informática	4
	Empreendedorismo em Informática	2
	Planejamento e Gestão de Sistemas de Informação	4
	Probabilidade e Estatística	4

Fonte Adaptada de (BRITO *et al.*, 2009) p. 63

A estrutura administrativa acadêmica é composta por três coordenações principais, que são a Coordenação de Curso, a Coordenação Adjunta de Estágio e a Coordenação Adjunta de Trabalho de Conclusão de Curso (TCC). Além das três coordenações, o curso conta com dois secretários: um secretário acadêmico, responsável pelo registro e controle acadêmico dos conteúdos, frequências e aproveitamentos dos discentes e um secretário de curso, responsável por auxiliar as coordenações e os professores do curso, gerenciando a comunicação oficial das três coordenações, distribuição e controle das mensagens eletrônicas oficiais e auxílio nas atividades administrativas correlatas ao curso.

O colegiado do Curso de Sistemas de Informação é composto pelos professores que ministram disciplinas regulares no curso e alunos indicados pelo centro acadêmico de Sistemas de Informação para o mandato de um ano letivo. O colegiado se reúne para deliberar assuntos de interesse do curso. É a instância superior do curso nos termos previstos no regimento geral da UEG.

O registro das frequências e notas dos discentes é realizado pelo professor em um sistema *online* “Veritas”. Além disso, cada aluno possui um dossiê (uma pasta) onde são arquivados todos os documentos relativos à sua vida acadêmica, tais como guias de matrícula, solicitações diversas, documentação pessoal e outros, mantidos pelo secretário acadêmico.

Em 2009, o curso teve sua grade curricular reformulada, atendendo às novas exigências do mercado e algumas recomendações do Conselho Estadual de Educação. A modalidade é bacharelado com regime de funcionamento seriado anual, sendo que na matriz de 2002 o tempo mínimo de integralização é de 4 anos e máximo de 7 anos; e na matriz de 2009, o tempo mínimo é de 4 anos e máximo de 6 anos. Na matriz de 2002, a carga horária total é 3280 horas/aulas e na matriz de 2009 a carga horária total é 3030 horas. De segunda a sexta, as aulas são ministradas no período noturno e, aos sábados, no período matutino. A forma de ingresso é anual e são oferecidas 40 vagas no processo seletivo (BRITO *et al.*, 2009).

A proposta do novo projeto pedagógico é ter um egresso com postura ativa, capaz de propor e implementar o alinhamento entre as tecnologias de informação e demandas das organizações e da sociedade, dando ênfase no processo e nas boas práticas de desenvolvimento de software e gestão de sistemas de informação, buscando agregar valores em processos decisórios, em análise de cenários e em implementação de novas estratégias organizacionais.

A matriz curricular ilustrada na tabela 2 é resultado do trabalho para atender às recomendações do Conselho Estadual de Educação e de pesquisas realizadas em matrizes de outras instituições de ensino superior, discussões com os egressos atuantes no mercado de trabalho e deliberada pelo colegiado do curso.

Tabela 2 - Matriz Curricular 2009

Série	Disciplina	Carga Horária Semanal
	Álgebra Linear	2
	Cálculo Diferencial e Integral	4

1 ^a	Eletricidade e Lógica Digital	2
	Fundamentos de Matemática para a Computação	2
	Inglês Instrumental	2
	Linguagem e Técnicas em Programação	4
	Princípios da Administração	4
	Sistemas de Informação, a Profissão e a Sociedade	4
2 ^a	Análise de Sistemas	4
	Arquitetura de Computadores	2
	Bancos de Dados	4
	Estruturas de Dados	4
	Metodologia da Pesquisa Científica	2
	Programação I	4
	Sistemas Operacionais	2
	Teoria da Computação	2
3 ^a	Engenharia de Software	4
	Novas Aplicações em Sistemas de Informações	4
	Prática Interdisciplinar em Desenvolvimento de Software	4
	Probabilidade e Estatística	4
	Programação II	4
	Redes de Computadores	4
4 ^a	Gestão Organizacional e Financeira	4
	Direito e Ética em Informática	4
	Auditoria e Segurança em Sistemas de Informação	2
	Empreendedorismo em Informática	2
	Interação homem-máquina	2
	Planejamento e Gestão de Sistemas de Informação	4

Fonte Adaptada de (BRITO *et al.*, 2009) p. 65

Como descrito no projeto do curso, a integralização é prevista em 4 anos, desde que o aluno seja aprovado em todas as disciplinas. O critério de aprovação é o mesmo de todos os cursos da UEG. O aluno é aprovado na disciplina com frequência igual ou superior a 75% das aulas efetivamente ministradas e que obtiver média parcial

(MP) igual ou superior a 7,0 ou média final (MF) igual ou superior a 5,0. A MP é a média aritmética das Verificações de Aprendizagem (VA) aplicadas no final de cada bimestre. A MF é a média aritmética entre a MP e a nota da prova final (PF). A PF é uma prova aplicada ao aluno que obtiver MP inferior a 7,0 e igual ou superior a 3,0. Caso a MP seja menor do que 3,0, o aluno estará reprovado na disciplina.

Na tabela 3, apresenta-se a quantidade de alunos matriculados entre os anos de 2005 a 2009 em cada série. Em termos gerais, observa-se que há uma pequena variação no número de alunos matriculados.

Tabela 3 - Alunos matriculados de 2005 à 2009

Ano	Matriculados por Série				Total
	1ª Série	2ª Série	3ª Série	4ª Série	
2005	39	47	35	26	147
2006	43	57	36	29	165
2007	54	40	41	31	166
2008	52	32	42	38	164
2009	37	41	31	47	156

Fonte Adaptada de (BRITO *et al.*, 2009) p.103

Na tabela 4, apresentam-se indicativos do curso no decorrer de alguns alunos para demonstrar a origem dos ingressantes, trancamentos, taxa de desistentes, abandonos, graduados e transferência de outras Unidade da UEG ou IES.

Tabela 4 - Indicativos do Curso de SI UnUCET-UEG

	2005	2006	2007	2008
Ingressantes por Processo Seletivo	38	33	40	39
Matriculados	165	166	164	156
Trancamentos de matrícula	0	3	0	3
Desistentes/Abandonos	13	7	6	12
Graduados	16	24	27	27
Ingressantes por outras vias	0	7	2	0
Transferência para outra UnU ou IES	0	0	0	0

Fonte Adaptada de (BRITO *et al.* 2009) p.103

Quanto ao perfil do ingressante, não há uma definição. O candidato deve ter concluído o ensino médio ou equivalente ou ter a possibilidade de conclusão até o ato da matrícula e ser aprovado no Processo Seletivo. Caso as vagas não sejam preenchidas, poderá ocorrer ingresso ao curso por transferência de outra unidade universitária da UEG ou de outra instituição de ensino superior ou por portadores de diploma. Os casos de *ex officio* independem de vagas.

4.2 Processo seletivo

O processo seletivo do curso é realizado pelo Núcleo de Seleção da UEG, juntamente com outros cursos anualmente. Publica-se um edital contendo as instruções para o preenchimento da ficha de inscrição, valor da taxa de inscrição, relação de documentos exigidos, os cursos com suas respectivas vagas, o conteúdo programático e tipo das provas, as obras literárias exigidas, calendário e locais das provas, os critérios de avaliação e classificação, procedimentos para matrículas, entre outras informações relevantes aos candidatos.

A prova de Conhecimentos Gerais é comum a todos os cursos, composta de 80 questões objetivas, com proposições de múltipla escolha e abrangendo conteúdos das seguintes disciplinas: Língua Portuguesa e Literatura Brasileira, Língua Estrangeira Moderna (Inglês ou Espanhol), Geografia, História, Matemática, Física, Química e Biologia. Além da prova de Conhecimentos Gerais, os candidatos para o curso de Sistemas de Informação têm Redação e Matemática como prova Discursiva Específica.

No ato de inscrição ao processo seletivo, o candidato é convidado e motivado a responder um Questionário Socioeconômico e Cultural, para traçar o perfil

da clientela que pretende ingressar na instituição. É esclarecido ao candidato que não há respostas certas ou erradas, mas que é preciso respondê-las com sinceridade. As informações fornecidas não influenciam na avaliação do candidato a uma vaga na UEG. O questionário socioeconômico e cultural é formulado com o objetivo de conhecer o perfil dos candidatos. É notório que na sociedade há mudanças nos comportamentos das pessoas a cada ano, motivadas pelas evoluções tecnológicas ou situações financeiras e o questionário também muda com a tentativa de melhor identificar o perfil dos ingressantes. No anexo A, apresenta-se a evolução dos questionários utilizados nos processos seletivos de 2002 a 2007, anos utilizados como material de pesquisa deste trabalho.

A correção da prova objetiva de Conhecimentos Gerais, que tem caráter eliminatório e classificatório, é feita por meio eletrônico, enquanto que a correção das provas de Redação e Discursiva Específica é feita por bancas formadas por professores das respectivas áreas. Somente são corrigidas as provas dos candidatos classificados na prova de Conhecimentos Gerais, em ordem decrescente dos pontos obtidos em uma proporção de três candidatos por vagas. No caso de haver mais de um candidato no último lugar com o mesmo somatório de pontos, todos eles serão classificados para correção das provas de redação e discursiva específica, mesmo que seja ultrapassado o número de três candidatos por vaga.

A classificação final dos candidatos é feita pela soma total dos pontos obtidos por cada candidato nas provas de Conhecimentos Gerais, Redação e Discursiva Específica, em ordem decrescente dos resultados, até o limite de vagas oferecidas para cada curso para o qual o candidato é considerado Classificado. Após o limite de vagas, o candidato é considerado “Excedente”. Em caso de empate prevalecem os seguintes

critérios: maior número de pontos na prova Discursiva Específica, maior número de pontos na prova de Redação e idade.

Na tabela 5, apresenta-se a concorrência por vagas do curso de SI-UnUCET-UEG referentes aos anos de ingressos de 2002 à 2009. Observa-se que o número de candidatos tem diminuído a cada ano.

Tabela 5 - Concorrência do processo seletivo do período de 2001 à 2008

2001	2002	2003	2004	2005	2006	2007	2008
22,73	18,1	16,5	15	11,33	11,83	10,66	9,56

Fonte Adaptada de (BRITO *et al.*, 2009) p.102

5 Aplicação de DCBD para identificar o perfil do aluno

Apresenta-se neste capítulo, o estudo de caso realizado com os dados dos alunos do Curso de Sistemas de Informação da UnUCET. A organização das seções segue as fases definidas da metodologia de descoberta de conhecimento em base de dados CRISP-DM.

Na metodologia CRISP-DM, a primeira fase de um processo de descoberta de conhecimento em base de dados (DCBD) é o entendimento do negócio e possui quatro tarefas principais, que são: a determinação dos objetivos do negócio, a avaliação da situação, a determinação dos objetivos da mineração dos dados e a produção do plano do projeto.

5.1 Determinação dos objetivos do negócio

Um dos objetivos prático do curso de Sistemas de Informação da UnUCET-UEG é formar discentes, em 4 anos, aptos a desenvolver e gerenciar sistemas de informação no mercado de trabalho.

Pela carência de informações estatísticas em relação aos alunos, as informações disponíveis para a realização da DCBD são as respostas do questionário socioeconômico e cultural preenchidos no ato da inscrição do vestibular e as notas obtidas no vestibular das provas aplicadas (objetiva, específica e redação), além da média final de cada disciplina cursada durante o curso. Todos estes dados são utilizados

para criar uma categorização de alunos, definido como “aluno aplicado” que é aquele que não foi reprovado em nenhuma disciplina, e “aluno não aplicado” definido como aquele que obteve pelo menos uma reprovação. Para ser considerado reprovado na disciplina sua média final é inferior a 5 e aprovado superior ou igual a 5 e ter freqüentado 75% das aulas.

Pretende-se aplicar o DCBD para identificar padrões de comportamento dos egressos, utilizando os dados acadêmicos e do processo seletivo. Através do resultado obtido tem-se o perfil dos alunos aplicados, que é disponibilizado para a gestão do curso que, o após análise das informações, desenvolve ações para direcionar a sua estratégia de marketing nos canais mais adequados, com a finalidade de atrair candidatos com o perfil fornecido, para os próximos processos seletivos. Devido a essas ações, o processo seletivo pode ser mais direcionado e, conseqüentemente, o curso poderá ter alunos com potencial para integralizar a matriz curricular em 4 anos.

5.2 Avaliação da situação

A pessoa indicada para ser o analista de dados para interpretar e aprovar os resultados obtidos é o coordenador do curso (WINCK, 2007). Os dados acadêmicos são disponibilizados pela PrG-UEG, no formato de planilha eletrônica e os dados do vestibular também são fornecidos no mesmo formato pelo Núcleo de Seleção da UEG.

Após uma limpeza dos dados fornecidos, o quantitativo útil de registros é pequeno, porém significativo, dando condições para aplicar a mineração e encontrar o perfil do aluno aplicado.

O principal risco do projeto é não encontrar padrão nenhum que identifique o perfil do “aluno aplicado” ou que o encontre com baixa probabilidade. Porém, existem várias perguntas no questionário que podem ser combinadas, de tal forma que a etapa de mineração de dados possa ser repetida várias vezes, até que seja encontrado um modelo satisfatório. Em último caso, os motivos da experiência mal sucedida devem ser relatados e explicados.

As principais terminologias são: aluno aplicado (I) aquele que não foi reprovado em nenhuma disciplina durante sua graduação; aluno não aplicado (N) aquele que foi reprovado pelo menos em uma disciplina. “Objetiva” é nota da prova de conhecimentos gerais; “específica” é a nota das provas específicas de Língua Portuguesa, Matemática e Física; “redação” é nota da prova de redação; Q1 a Q32 correspondem aos nomes das perguntas do questionário socioeconômico e cultural; Q1-A, Q1-B e Q1-C corresponde às alternativas de cada pergunta.

5.3 Determinação dos objetivos da mineração dos dados

A categorização dos alunos (aplicado e não aplicado) é conhecida pelo fato de ter integralizado ou não o curso em 4 anos. Pretende-se obter um modelo de mineração de dados que mapeie os atributos do Questionário Socioeconômico e Cultural e das notas do vestibular para um dos rótulos de classes “aluno aplicado” e “aluno não aplicado”. A tarefa da mineração dos dados mais apropriada é a classificação, uma vez que já conhece a variável alvo e o que se pretende identificar são as variáveis explicativas que induzem o conjunto de dados à classe. A técnica de classificação escolhida é a árvore de decisão, por possuir uma boa precisão de classificação, pela rapidez na criação de modelo de aprendizagem e pela facilidade de

assimilação de sua estrutura com a representação do modelo do mundo real, onde os rótulos de classes são os nodos folha e os outros atributos são os nodos internos ou o nodo raiz que indicam as regras de decisões para chegar ao nodo folha (TAN *et al.*, 2009).

5.4 Preparação dos dados

Os dados acadêmicos são fornecidos pela Pró-Reitoria de Graduação da UEG no formato de planilha eletrônica sem normalização, contendo 6081 linhas, sendo que cada uma delas fornece informações da disciplina que o aluno cursou. Na tabela 6, é ilustrada a relação dos atributos dessa planilha.

Tabela 6 - Atributos dos dados acadêmicos

Nome da Coluna – Atributos	Descrição
Matriz	Identificação da matriz curricular que está matriculado
Matrícula	Numero da matrícula do aluno
Ano/Semestre Inicial	Ano que o aluno iniciou no curso
Nome	Nome do aluno
Data Nascimento	Data de Nascimento do aluno
Rua	Rua onde reside o aluno
Complemento	Complemento do endereço do aluno
Bairro	Bairro ou setor onde o aluno reside
CEP	CEP do endereço do aluno
Disciplina	Nome da disciplina
Ano Matrícula	Ano da matrícula na disciplina
Matriz da Disciplina	Matriz da disciplina
Situação	Situação do aluno na disciplina aprovado ou reprovado
Total de Faltas	Total de Faltas do aluno na disciplina
Media Final	Media final da disciplina
Frequência	Percentual da frequência do aluno na disciplina

1ªVA	Nota da 1ª verificação de aprendizagem
2ªVA	Nota da 2ª verificação de aprendizagem
3ªVA	Nota da 3ª verificação de aprendizagem
4ª VA	Nota da 4ª verificação de aprendizagem
Exame Final	Nota da prova final

As informações dos processos seletivos fornecidas pelo Núcleo de Seleção da UEG também são fornecidas no formato de planilha eletrônica, contendo informações dos processos seletivos de 2002/1 até 2011/1 de toda a UEG. Cada vestibular possui uma guia diferente, totalizando 19 guias, cada uma contendo aproximadamente 409.123 linhas, e cada linha armazena a resposta de uma pergunta do Questionário Socioeconômico e Cultural dos candidatos classificados e excedentes. A tabela 7 ilustra as colunas desta planilha.

Tabela 7 - Colunas da planilha dados dos processos seletivos

Nome da Coluna	Descrição
Situação	Situação do candidato no processo seletivo
Turma	A turma que o candidato concorreu, é composta pelo nome do curso a unidade universitária e turno do curso
CodInscricao	Código da inscrição ao vestibular
Nome	Nome do candidato
DataNascimento	Data de Nascimento do candidato
Endereço	Endereço do candidato
Bairro	Bairro do candidato
CidadeRes	Cidade da residência do candidato
UF	Unidade federativa
Objetiva	Nota que o candidato obteve na prova Conhecimento Gerais
Discursiva	Nota que o candidato obteve na prova Discursiva
Redação	Nota que o candidato obteve na Redação
CodQuestão	Código da questão do Questionário Socioeconômico e Cultural
Questão	Descrição da questão do Questionário Socioeconômico e Cultural

Alternativa	Alternativa que o candidato respondeu
-------------	---------------------------------------

Estas duas planilhas contêm os dados de entrada a serem submetidas à mineração. Antes de explorar os dados, essas planilhas são importadas para o Microsoft Access, ambiente onde se realizam algumas normalizações, preparos e limpeza desses dados.

Os dados acadêmicos são importados para a tabela “AlunosDadosAcademicos”. O primeiro passo é identificar os alunos reprovados em mais de uma disciplina e os que não são reprovados em nenhuma. Com estes dois grupos de alunos é possível criar um novo atributo na tabela que receberá a letra “S” para os alunos que integralizaram o curso em 4 anos e a letra “N” para os que levaram mais tempo para concluir o curso.

O segundo passo é separar os dados do curso de SI da UnUCET-UEG em uma nova tabela com o mesmo *layout*, conforme apresentado na tabela 7, denominada “vestibulandos”. O terceiro é encontrar a correspondência dos registros dos alunos matriculados com os da tabela “vestibulandos”, colocando a matrícula em nova coluna com o mesmo nome. O quarto passo é criar uma nova coluna na tabela “vestibulando” com o nome “alunoaplicado”, que recebe os dados da tabela “AlunosDadosAcademicos”.

O quinto passo é a análise detalhada dos questionários socioeconômicos e culturais aplicados nos processos seletivos, apresentados no anexo A. Constata-se que, no decorrer dos anos, há adequações das questões e das alternativas, criando-se uma dificuldade para selecionar as questões, porque não são as mesmas alternativas e nem é o mesmo o número da questão (posicionamento no questionário) em todos os anos,

variando de um ano para outro. Por exemplo, o posicionamento da questão “O que influenciou na sua opção de curso?” aparece como a vigésima quinta pergunta em um ano e como a vigésima terceira em outro, e a mesma situação ocorre com outras perguntas.

O sexto passo consiste em limpar e adequar o questionário socioeconômico, criando uma tabela “QuestoesVestibular” no banco de dados do Microsoft Access com o *layout* apresentado na tabela 8.

Tabela 8 - Layout da tabela “QuestoesVestibular”

Coluna	Descrição
Identificação	Número seqüência que identifica o registro como único
Questão	Nome da coluna na tabela “vestibulandos” que armazena a resposta da questão
QuestaoOriginal	Pergunta
AlternativaOriginal	Alternativa
AlternativaRepresentada	Codificação para representa os agrupamentos das alternativas

No anexo B, apresentam-se os registros da tabela “QuestoesVestibular”, após serem submetidos ao processo de limpeza e pré-processamento.

Para agrupar as alternativas apresentadas no anexo B, são considerados os conceitos de taxa de ganho (TAN *et al.*, 2009). Quanto maior for o número de alternativas, menor é a quantidade de registros associados e, conseqüentemente aumentam as informações divididas que, por sua vez, reduzem a taxa de ganho. A taxa de ganho é comprovada nos diferentes ciclos de mineração realizados neste estudo de uso.

5.5 Mineração de dados

De acordo com a metodologia CRISP-DM, o processo de DCDB pode ser compreendido em fases, que podem ser refeitas até que se alcance um resultado satisfatório. A seguir são apresentados os atributos escolhidos, os ciclos realizados com diferentes entradas, a quantidade de registros que sempre é a mesma, mudando os atributos, os agrupamentos de alternativas e mudança nos parâmetros do algoritmo J48.

5.5.1 Atributos para mineração de dados

Após o processo de limpeza dos dados, são escolhidos os atributos: nota da prova objetiva, nota da prova discursiva, nota da redação, o rótulo da classe aluno aplicado e 10 questões do questionário socioeconômico cultural. As questões foram escolhidas considerando o conhecimento e experiência que o proponente possui do curso de SI-UnUCET, pois o processo de descoberta de conhecimento através da mineração de dados não é um processo de relação de causa e efeito, ou seja, não tem como comprovar cientificamente quais são os efeitos de uma causa. É um processo que dá subsídio para um profissional conhecedor do negócio identificar os efeitos de uma provável causa.

No anexo B são listadas todas as perguntas com suas respectivas alternativas do questionário socioeconômico cultural, e entre elas são escolhidas as seguintes:

1. Na tabela 9 são mostradas as alternativas agrupadas da questão Q6: “Como você se preparou para este processo seletivo?”.

Tabela 9 – Alternativas agrupadas da questão Q6

Alternativa original	Alternativa representada
Estudei apenas durante a 3ª série regular	Q6-A
Estudei sozinho	Q6-A
Fiz cursinho por menos de um ano.	Q6-B
Fiz cursinho por um ano ou mais.	Q6-B
Fiz cursinho por um ano	Q6-B
Em branco	Q6-N

2. Na tabela 10, são mostradas as alternativas agrupadas da questão Q13:

“Em que tipo de estabelecimento escolar fez seus estudos?”.

Tabela 10 - Alternativas agrupadas da questão Q13

Alternativa original	Alternativa representada
Somente particular	Q13-A
Maior parte particular.	Q13-A
Somente público.	Q13-B
Maior parte público.	Q13-B
Resposta em branco	Q13-N

3. Na tabela 11, são mostradas as alternativas agrupadas da questão Q18:

“Qual a sua idade completa?”.

Tabela 11 - Alternativas agrupadas da questão Q18

Alternativa original	Alternativa representada
Até 17 anos.	Q18-A
18 anos	Q18-A
19 anos	Q18-B
20 a 22 anos	Q18-B
23 a 26 anos.	Q18-C
27 a 30 anos.	Q18-C
Acima de 30 anos	Q18-C
Em branco	Q18-N

4. Na tabela 12, são mostradas as alternativas agrupadas da questão Q20:

“O que influenciou na sua opção de curso?”.

Tabela 12 - Alternativas agrupadas da questão Q20

Alternativa original	Alternativa representada
Já trabalho na área.	Q20-A
Prestígio social.	Q20-B
Maior retorno financeiro	Q20-B

Maiores oportunidades de trabalho.	Q20-B
Relação de candidatos-vaga	Q20-C
Por outros motivos	Q20-C
Contribuição para a sociedade.	Q20-C
Influência de família e/ou terceiros	Q20-C
Relação candidatos-vaga.	Q20-C
Em branco	Q20-N

5. Na tabela 13, são mostradas as alternativas agrupadas da questão Q22

“Onde reside atualmente?”.

Tabela 13 - Alternativas agrupadas da questão Q22

Alternativa original	Alternativa representada
Anápolis	Q22-A
Região Metropolitana de Goiânia	Q22-B
Goiânia	Q22-B
Sudoeste goiano.	Q22-C
Sul goiano	Q22-C
Outros estados.	Q22-C
Centro goiano - Eixo BR-153.	Q22-C
Cidades de outros estados	Q22-C
Entorno do Distrito Federal	Q22-C
Nordeste goiano.	Q22-C
Norte goiano	Q22-C
Oeste goiano - Eixo GO-060.	Q22-C
Interior de outros estados	Q22-C
Sudeste goiano - Estrada de Ferro.	Q22-C
Outra cidade do interior do Estado de Goiás	Q22-C
Capital do outros estados	Q22-C
Em branco	Q22-N

6. Na tabela 14, são mostradas as alternativas agrupadas da questão Q26:

“Qual é a renda mensal de seu grupo familiar?”.

Tabela 14 - Alternativas agrupadas da questão Q26

Alternativa original	Alternativa representada
Um salário mínimo	Q26-A
Quatro a cinco salários mínimos	Q26-B
Dois a três salários mínimos.	Q26-B
Seis a sete salários mínimos	Q26-C
Oito a dez salários mínimos.	Q26-C
Acima de dez salários mínimos.	Q26-C
Em branco	Q26-N

7. Na tabela 15, são mostradas as alternativas agrupadas da questão Q27:

“Qual é a sua participação na vida econômica do seu grupo familiar?”.

Tabela 15 - Alternativas agrupadas da questão Q27

Alternativa original	Alternativa representada
Não trabalho.	Q27-A
Trabalho, mas recebo ajuda financeira da família ou de outras pessoas.	Q27-A
Trabalho, mas recebo ajuda financeira	Q27-A
Trabalho, e contribuo para sustento da família	Q27-A
Não trabalho	Q27-A
Trabalho e sou o principal responsável pelo sustento da família	Q27-B
Trabalho e sou responsável pelo sustento da família	Q27-B
Trabalho para o meu próprio sustento.	Q27-B
Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família	Q27-B
Resposta em branco	Q27-N

8. Na tabela 16, são mostradas as alternativas agrupadas da questão Q30:

“Qual o nível de escolaridade de seu pai ou responsável?”.

Tabela 16 - Alternativas agrupadas da questão Q30

Alternativa original	Alternativa representada
Ensino Superior completo.	Q30-A
Pós-graduação.	Q30-A
Ensino Superior incompleto.	Q30-A
Nenhuma instrução.	Q30-B
Ensino Fundamental completo.	Q30-B
Ensino Fundamental incompleto	Q30-B
Ensino Médio completo.	Q30-B
Ensino Médio incompleto.	Q30-B
Nenhuma instrução	Q30-B
Resposta em branco	Q30-N

9. Na tabela 17, são mostradas as alternativas agrupadas da questão Q31:

“Qual o nível de escolaridade de sua mãe ou responsável?”.

Tabela 17 - Alternativas agrupadas da questão Q31

Alternativa original	Alternativa representada
Ensino Superior completo.	Q31-A
Pós-graduação.	Q31-A
Ensino Superior incompleto.	Q31-A
Nenhuma instrução.	Q31-B

Ensino Fundamental completo.	Q31-B
Ensino Fundamental incompleto	Q31-B
Ensino Médio completo.	Q31-B
Ensino Médio incompleto.	Q31-B
Nenhuma instrução	Q31-B
Resposta em branco	Q31-N

10. Na tabela 18, são mostradas as alternativas agrupadas da questão Q33:

“Qual o seu sexo?”.

Tabela 18 - Alternativas agrupadas da questão Q33

Alternativa original	Alternativa representada
Feminino	Q33-A
Masculino	Q33-B

Após a criação do arquivo ARFF com estes atributos, são inseridos no WEKA para mineração, na aba “*preprocess*” na opção “*open file*” o arquivo “Alunos-SI-01.ARFF” é lido para dentro do ambiente do WEKA. Na figura 11, apresentam-se os dados, a estatística dos atributos e seus valores com suas respectivas quantidades.

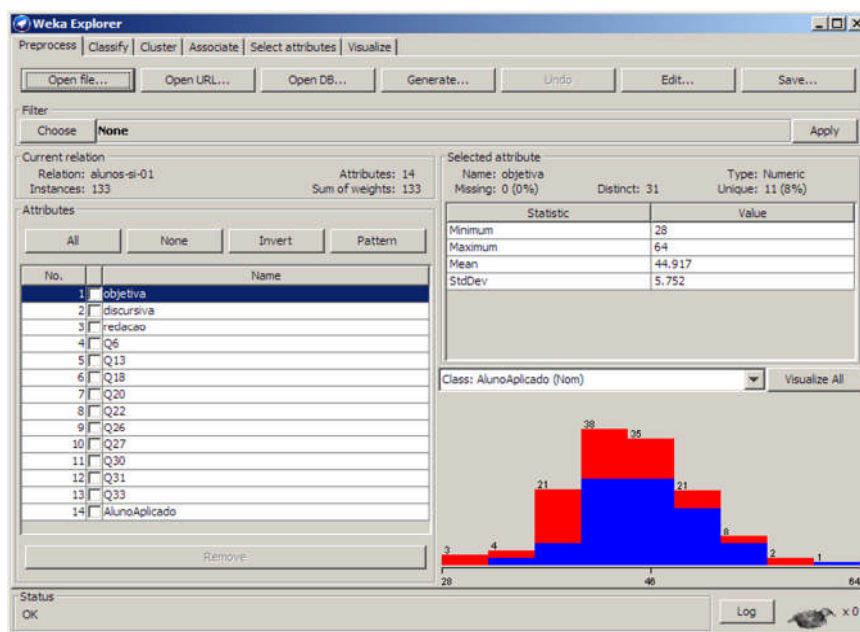


Figura 11 - Dados estatísticos dos atributos do arquivo alunos-si-01

Um dos mecanismos de poda utilizados pelo algoritmo J48 é baseado na comparação das taxas de estimativas de erro de cada subárvore e do nó folha. São processados sucessivos testes a partir do nó raiz da árvore, e se a estimativa de erro indicar que a árvore será mais precisa se os nós descendentes (filhos) de um determinado nó forem eliminados, então estes nós descendentes serão eliminados e o nó indicado passará a ser o novo nó folha. O fator de confiança é a taxa de estimativa de erro, quanto menor for, maior será a poda.

Matriz de confusão é uma matriz quadrada que contém informações para o entendimento do resultado do algoritmo de classificação, que são:

- A quantidade de instâncias classificadas corretamente;
- A quantidade de instâncias classificadas erroneamente;
- A quantidade de instâncias que o algoritmo acreditava ser de um tipo (eleito) e na verdade foram classificadas como de outro tipo (não eleito)

5.5.2 Primeiro ciclo de mineração de dados

Neste ciclo são usados todos os atributos descritos na seção anterior sem nenhuma alteração. Através do algoritmo J48 com fator de confiança padrão de 25% e podendo realizar poda, encontrou-se o resultado mostrado na figura 12. Na matriz confusão, percebe-se que selecionando 67 da classe dos alunos aplicados, há uma possibilidade de 12 erros. E ao selecionar 45 da classe dos alunos não aplicados há uma possibilidade de 9 erros.

Na figura 13, mostra-se a árvore de decisão resultante da mineração deste primeiro ciclo, que classificou 112 instâncias corretamente e que representa 84.21% dos 133 registros treinados.

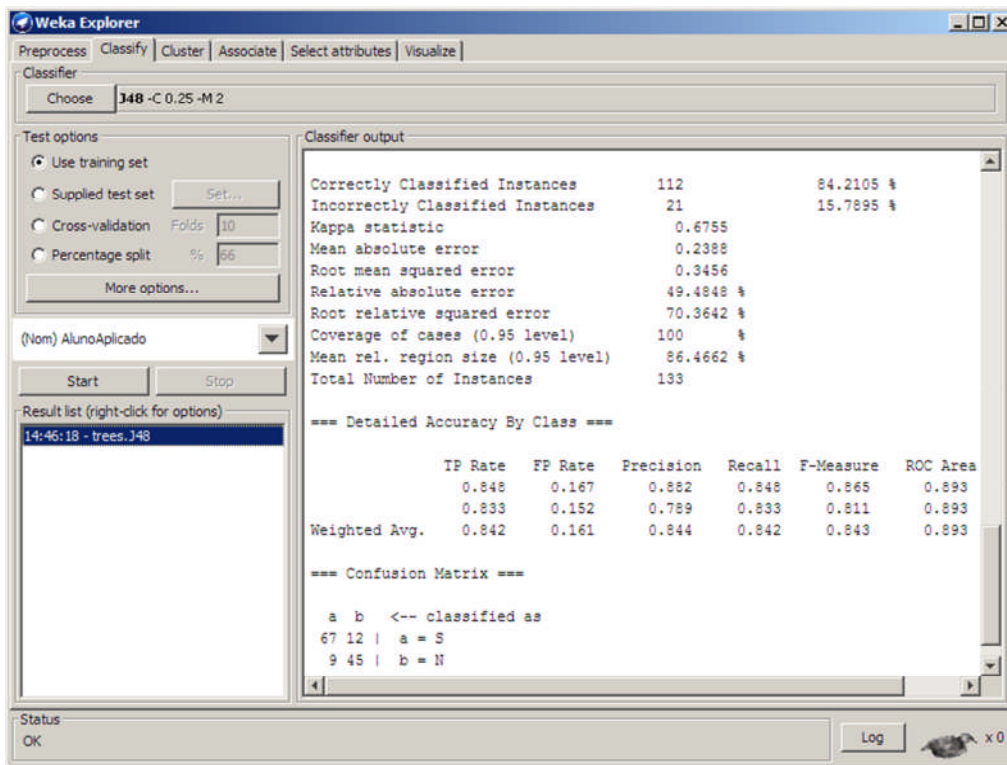


Figura 13 - Resultado da Classificação do arquivo alunos-si-01

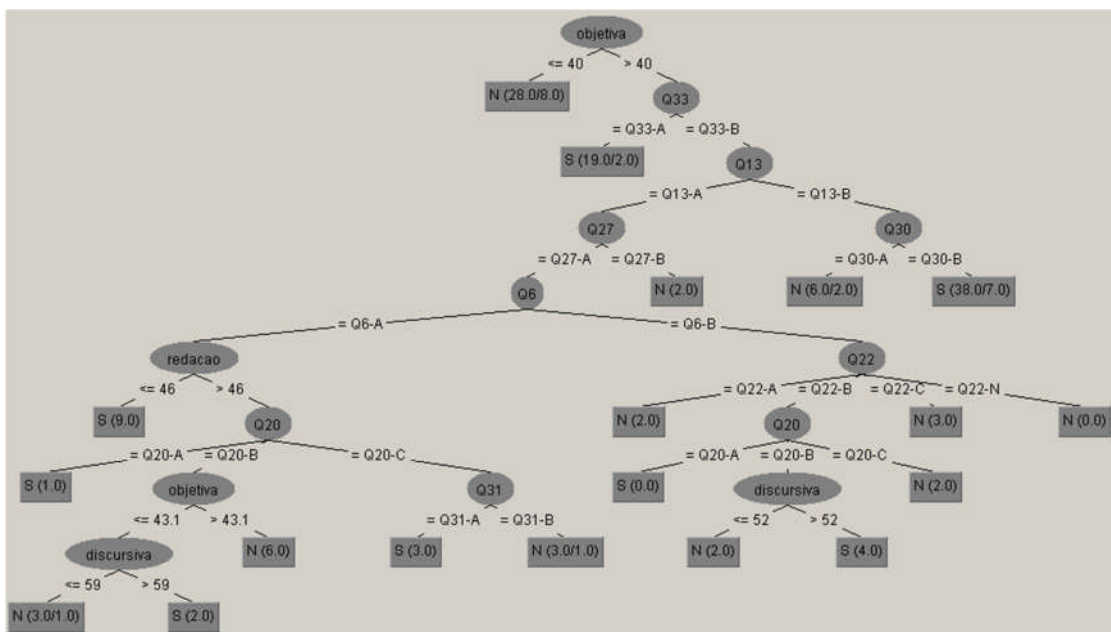


Figura 12 - Árvore de decisão do arquivo alunos-si-01

5.5.3 Segundo ciclo de mineração de dados

Neste ciclo, a única diferença reside no agrupamento das questões Q30 e Q31, utilizadas para comprovar os conceitos de taxa de ganho. Essas duas questões foram concatenadas em uma, tendo uma similaridade entre elas, considerando que se a mãe ou o pai tiverem nível de ensino superior, isso influencia o candidato, motivo este que justifica a junção das duas questões. A rotina que gera o arquivo ARFF considera a resposta Q30-A tanto para a questão Q30, quanto para a questão Q31, caso a resposta seja afirmativa em qualquer uma das duas questões.

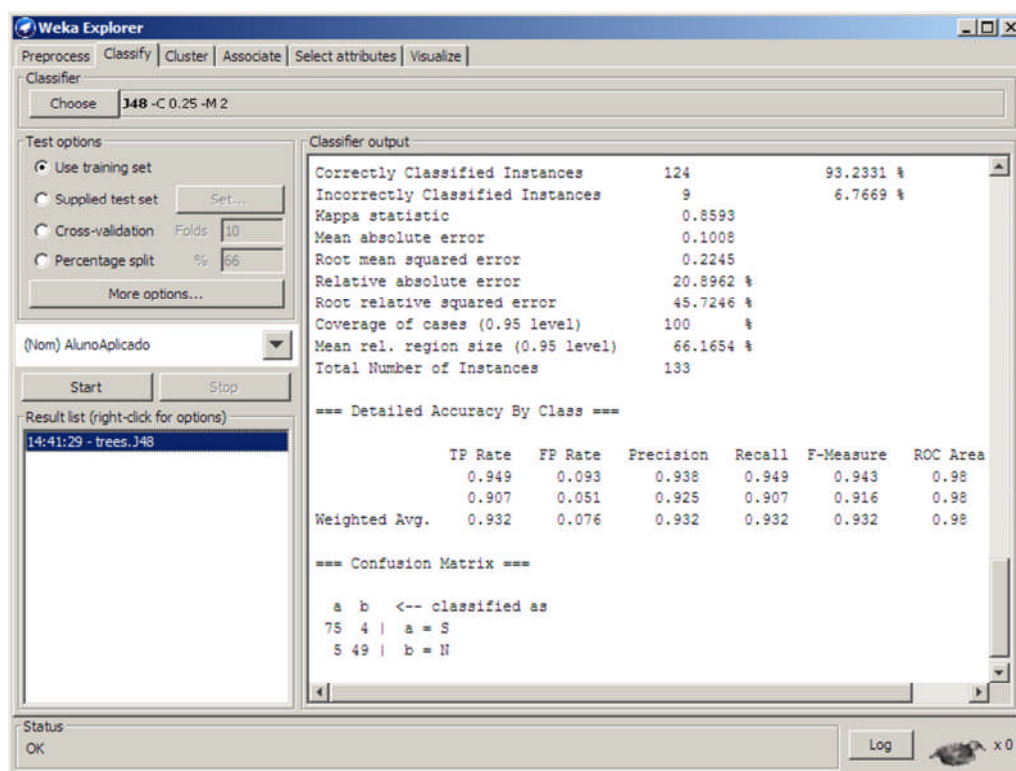


Figura 14 - Resultado da classificação do arquivo alunos-si-02

Na figura 14, mostra-se o resultado da classificação, sendo que 124 instâncias foram classificadas corretamente, totalizando 93.23% dos 133 registros. Na matriz confusão, são 75 da classe dos alunos aplicados com uma possibilidade de 4 erros e 49 da classe dos alunos não aplicados, com uma possibilidade de 5 erros.

Comparando os resultados do primeiro ciclo com o segundo, o ciclo que utilizou os conceitos de taxa de ganho para dividir os atributos obteve o melhor modelo. Na tabela 19, ilustram-se os indicadores, mostrando que no segundo ciclo ocorreram 12 instâncias classificadas corretas a mais que no primeiro ciclo.

Tabela 19 – Comparativo entre o 1º e o 2º ciclo

	1º ciclo	2º ciclo
Instâncias classificadas corretas	112	124
Instâncias classificadas incorretas	21	9

5.5.4 Terceiro ciclo de mineração de dados

Após exaustivos experimentos em desconsiderar atributos e alternar agrupamento de alternativas, tais como por exemplo, na questão Q18, foi realizado um novo agrupamento das faixa etárias e em outro experimento não foi usado o atributo Q8, e assim sucessivamente, com vários outros experimentos, alternando atributos e combinando os mesmos, sempre na busca do melhor resultado através da taxa de ganho. Ou seja, na escolha de atributos e na alternância de agrupamento das alternativas, somente foram encontrados resultados inferiores ao do 2º ciclo. Em função disso, foi utilizada outra estratégia em busca do melhor modelo, através da mudança de parâmetros do classificador J48.

5.5.5 Quarto ciclo de mineração de dados

O primeiro passo foi alterar o valor do fator de confiança de 25% para 15%, conseqüentemente, dando a ordem ao algoritmo para aumentar a poda. Entretanto, o

resultado não foi satisfatório em termos de classificação das instâncias, conforme é mostrado na figura 15; porém, construiu-se uma árvore menor, ou seja, com menor número de nós intermediários, como se mostra na figura 16. Dentre os 14 atributos iniciais, com este parâmetro, o algoritmo elaborou um modelo com apenas 5 atributos, resultando na queda das instâncias classificadas corretamente.

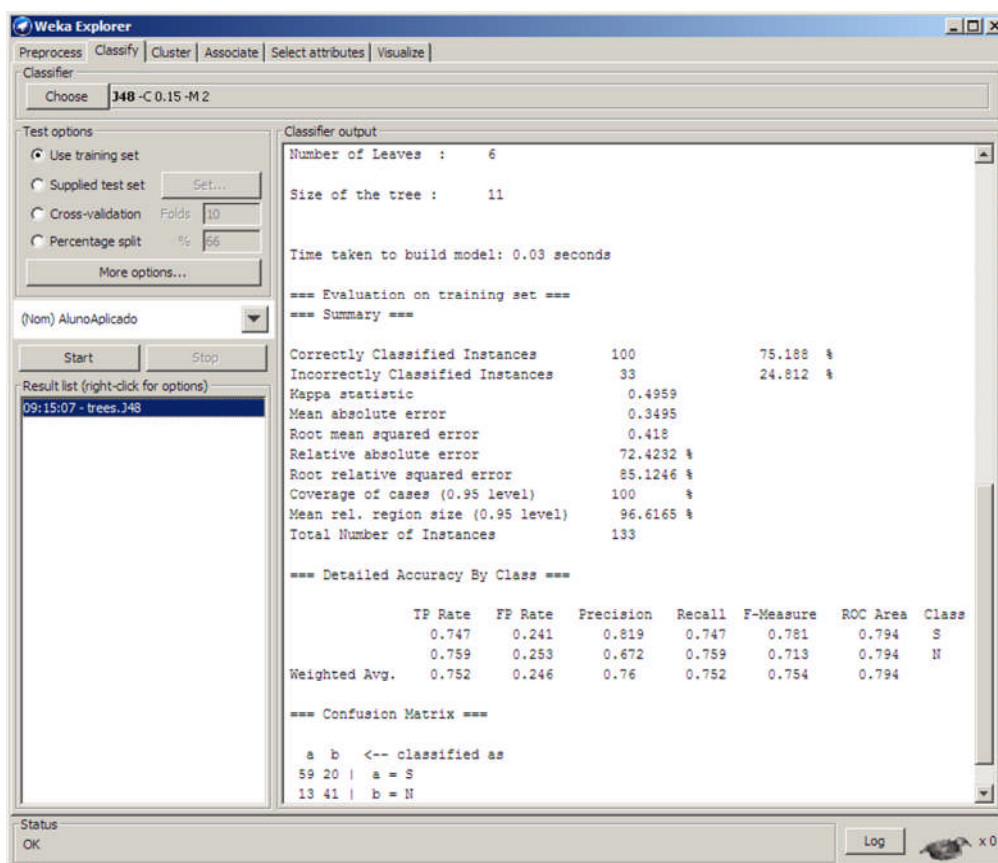


Figura 15 - Estatística do resultado da classificação usando fator de confiança 15%.

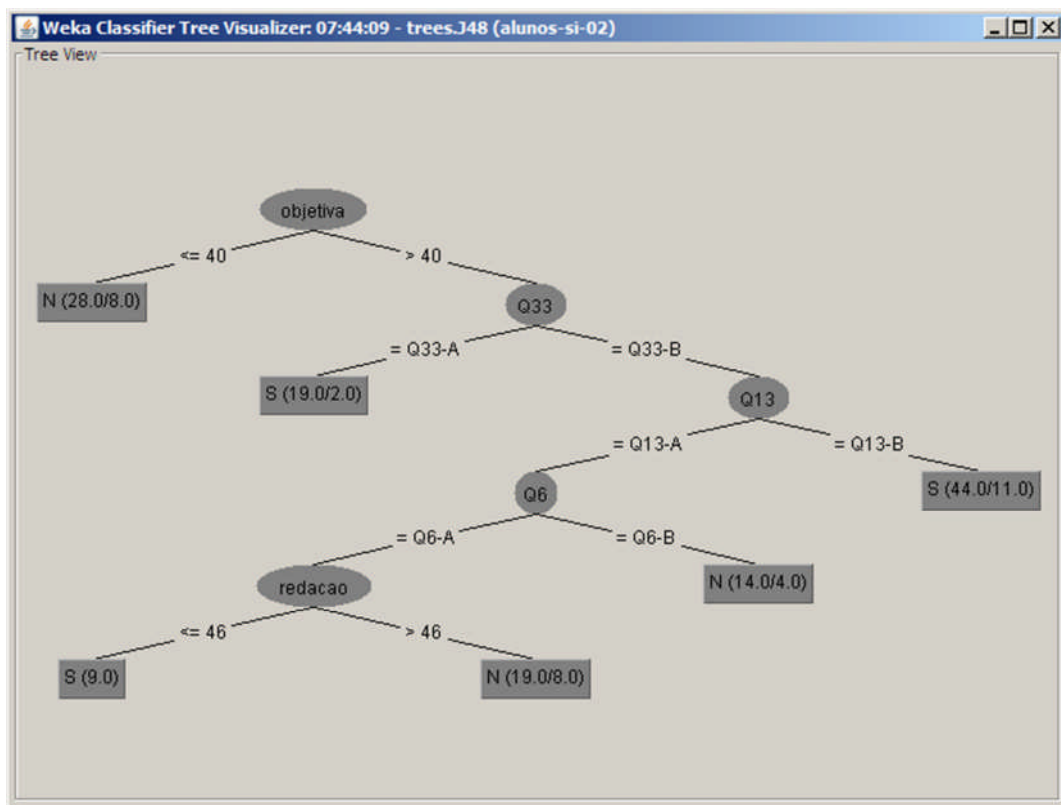


Figura 16 - Árvore de decisão com fator de confiança de 15% .

5.5.6 Quinto ciclo de mineração de dados

Neste ciclo, o parâmetro *unpruned* é configurado com *true*, dando a ordem para o algoritmo não realizar poda. Conforme se mostra na figura 17, são 128 instâncias classificadas corretamente, prevalecendo 96,24% dos registros treinados. Porém, a árvore resultante é maior que a dos outros resultados, conforme a figura 18.

5.5.7 Resultado da mineração de dados

Analisando os resultados, percebe-se que o resultado estatístico do 5º ciclo é o melhor entre os ciclos, conforme ilustra o gráfico da figura 19.

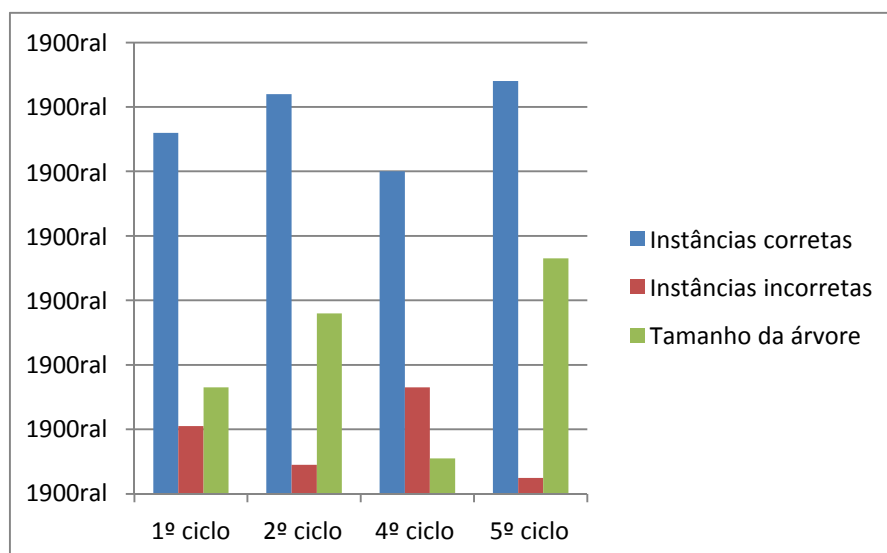


Figura 19 - Gráfico comparativo dos ciclos de mineração de dados

Observando-se a árvore resultante do 4º ciclo, percebe-se que sua interpretabilidade é muito mais fácil que a da árvore do 5º ciclo, e isso ocorre devido ao índice alto de poda realizado no 4º ciclo. Ao mesmo tempo, seu índice de instâncias corretas é melhor. Mesmo não sendo uma solução ótima, o 5º ciclo apresenta resultados satisfatórios em tempo factível.

Uma das últimas tarefas da metodologia CRISP-DM é apresentar o resultado do conhecimento encontrado e, dessa forma, a partir da análise das regras de classificação encontradas, é possível verificar padrões referentes ao perfil do ingressante ao Curso de SI-UnUCET-UEG.

Analisando a árvore do 5º ciclo, verifica-se a existência de três folhas com resultados expressivos entre os demais. Os percursos do nó raiz até as folhas com os três melhores resultados da classe “alunos aplicados” são apresentados na tabela 20.

Tabela 20 - Percurso das melhores folhas

Folha	Percurso	Registros classificados
1	objetiva > 40, Q33-B, Q13-B, Q26-B, Q20-B e Q30-B	15
2	objetiva > 40, Q33-B, Q13-A, Q27-A, Q6-A redação <= 46	9
3	objetiva > 40, Q33-A e Q20-B	9

Nos três percursos, a nota da prova objetiva é maior que 40, e a representação dos atributos do questionário socioeconômico cultural que influenciaram nos resultados são: Q6-A não fez cursinho, Q13-B fez 2º grau em escola pública, Q13-A fez 2º grau em escola particular, Q20-B maiores oportunidade de trabalho foi o que influenciou sua escolha pelo curso, Q26-B renda mensal familiar de dois a cinco salários mínimos, Q27-A não trabalha, Q30-B o pai ou mãe sem ensino superior, Q33-B masculino, Q33-A feminino.

O perfil do aluno da folha 1 é do sexo masculino, fez seu 2º grau em escola pública, ingressa sem fazer cursinho, o que influencia sua escolha pelo curso são as oportunidades de trabalho e seus pais não possuem ensino superior.

Essas são as informações fornecidas aos gestores para que possam usá-las na tomada de decisão, ao promoverem campanhas de marketing para prospectar novos alunos para o curso.

6 Conclusão

O CRISP-DM é um processo fácil de ser utilizado na prática, de forma eficaz e eficiente, em relação à realização de descoberta de conhecimento em banco de dados, começando com a reflexão do objetivo em descobrir o conhecimento e terminando na divulgação do conhecimento encontrado.

No entendimento do negócio, que é a primeira etapa, percebe-se a necessidade de conhecer bem o objetivo que se pretende alcançar, porque é esta fase que direciona o projeto e define qual tarefa de mineração de dados é utilizada. No caso do estudo deste trabalho, utilizou-se a classificação através de algoritmo de árvore de decisão. Na segunda fase, no entendimento dos dados, verifica-se a falta que faz a padronização e centralização dos dados em uma organização.

A preparação dos dados é a fase mais trabalhosa de todo o processo de descoberta de conhecimento em base de dados. É necessário o domínio de ferramentas computacionais que auxiliam na limpeza dos dados e o conhecimento em análise de dados facilita a realização dessa fase.

Na construção do modelo, percebe-se a necessidade do conhecimento que compõe o referencial teórico deste trabalho, mesmo usando uma ferramenta pronta como o WEKA. No caso deste trabalho, é necessário saber escolher os atributos que forneçam a melhor taxa de ganho, e quais são as consequências em usar ou não a poda. São estes entre outros parâmetros disponibilizados pelo WEKA que devem ser calibrados de acordo com os resultados. No estudo de caso, foi encontrado um modelo

com percentual de 96.24% classificados corretamente, sendo um resultado bastante aceitável.

Conclui-se que o aluno aplicado é aquele que, de forma geral, obteve nota maior que 40 na prova objetiva, do sexo masculino, fez o 2º grau na rede pública de ensino, não fez cursinho, seus pais não possuem ensino superior e o que influencia sua escolha pelo curso são as oportunidades de trabalho.

O modelo encontrado pode ser aprimorado através do acesso ao banco de dados do sistema acadêmico e do processo seletivo, para que possa mudar o critério da formação da classe de alunos aplicados. Incluir outros cursos similares para verificar padrões gerais, usando outra tarefa de mineração de dados para complementar o conhecimento obtido, como por exemplo, a associação. Predizer a classe dos alunos novos, para que o coordenador do curso adote estratégias na tentativa de mudar o cenário previsto, caso seja negativo.

Para que o modelo encontrado não se torne obsoleto, será necessária sua manutenção, através da utilização dos novos dados gerados no decorrer dos anos.

A aplicabilidade de mineração de dados na área de educação é uma área emergente, que pode trazer respostas para melhorar a qualidade do ensino, e é conhecida também como *Educational Data Mining (EDM)*. Como o objetivo do estudo de caso é identificar o perfil do egresso do Curso de Sistemas de Informação da UnUCET-UEG que concluiu o curso em 4 anos, torna-se uma contribuição para a melhoria no ensino.

Para trabalhos futuros, sugere-se o desenvolvimento de um sistema que possa integrar um sistema acadêmico com o WEKA ou outra ferramenta de mineração de dados e automatizar a divulgação do resultado, possibilitando a democratização do conhecimento encontrado no ambiente educacional.

Referências

- BASGALUPP, M. P., 2010, **LEGAL-Tree: Um algoritmo genético multi-objetivo lexicográfico para indução de árvores de decisão**. Tese de D.Sc., ICMC/USP São Carlos, São Carlos, SP, Brasil.
- BRITO, J. D. R., LIMA, J. O., SANTANA, M. R. C., et al. **Projeto pedagógico do curso de sistemas de informação**. UEG. Anápolis - GO. 2009.
- CARLOS, J. **Gestão escolar e novas tecnologias**. <http://www.icapti.com.br>, 2009. Disponível em: <<http://www.icapti.com.br/Artigos/detalhes.asp?PublicacaoID=139>>. Acesso em: 01 nov. 2011.
- CARVALHO, L. D. S., 2009, **Mineração de Dados Aplicada ao Entendimento do Comportamento do Consumidor para dar Suporte ao Processo de Tomada de Decisões**. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- CHAPMAN, P., CLINTON, J., KERBER, R., et al. CRISP-DM 1.0, Step-by-step data mining guide. **CRoss Industry Standard Process for Data Mining**, 2000. Disponível em: <<http://www.crisp-dm.org/CRISPWP-0800.pdf>>. Acesso em: 18 abr. 2011.
- CHIAVENATO, I. **Introdução à Teoria Geral da Administração**. 8^a Edição. ed. São Paulo: Campus, 2011.
- CHIU, S.; TAVELLA, D. **Data Mining and Market Intelligence for Optimal Marketing Returns**. Burlington: Elsevier Inc, 2008.
- DUARTE, K. B., 2010, **Programação genética aplicada no processo de descoberta de conhecimento em base dados de rede de pesquisa**. Dissertação de M.Sc., EEEC/UFG, Goiânia, GO, Brasil.
- FAYYAD, U. M., PIATETSKY-SHAPIRO, G., SMUTH, P., et al. **Advances in Knowledge Discovery and Data Mining**. Menlo Park: AAAI Press, 1996.

FERNANDES, M. P., 2007, **Descoberta de Conhecimento em base de dados e estratégias de relacionamento com cliente: Um estudo no setor de serviço.**

Dissertação de M.Sc., PPGA/MACKENZIE, São Paulo, SP, Brasil.

FERREIRA, J. B., 2005, **Mineração de dados na retenção de clientes em telefonia celular.** Dissertação de M.Sc., Departamento de Engenharia Elétrica/PUC-RJ, Rio de Janeiro, RJ, Brasil.

GUMMESSON, E. **Marketing de Relacionamento Total.** Porto Alegre: Bookman, 2010.

KAHANE, Y. LEVIN, N., MEIRI, R., et al. Applying Data Mining Technology for Insurance Rate Making: An Example of Automobile Insurance. **Asia-Pacific Journal of Risk and**, v. 2, 2007. ISSN Iss. 1, Article 3.

KDNUGGETS. KDnuggets : Polls : Data Mining Methodology (Aug 2007), 01 ago. 2007. Disponível

em:<http://www.kdnuggets.com/polls/2007/data_mining_methodology.htm>. Acesso em: 20 maio 2010, 20:30:10.

MARTINS, C. D. S., 1998, **Utilização da extração de conhecimento de bases de dados para identificar padrões de evasão de alunos de graduação da UNICAMP.** Dissertação de M.Sc., Instituto de Informática/PUC-Campinas, Campinas, SP, Brasil.

MINING, I. W. G. O. E. D. International Educational Data Mining Society, 15 July 2011. Disponível em: <<http://www.educationaldatamining.org/>>. Acesso em: 14 dezembro 2011, 18:30:30.

MONTEIRO, A. M. M. **A Era do Relacionamento: call center, contact center, telemarketing.** São Paulo: Fênix Editora, 2005.

NAYAK, G.; DEVI, S. "A SURVEY ON PRIVACY PRESERVING DATA MINING: APPROACHES AND TECHNIQUES". **International Journal of Engineering Science and Technology (IJEST)**, p. 2127-2133, 2011.

NISBET, R.; ELDER, J.; MINER, G. **Handbook of Statistical Analysis and Data Mining Applications**. Burlington: Elsevier Inc, 2009.

PATTANAIAK, S.; GHOSH, P. P. “Role of Data Mining in E-Payment systems”, **International Journal of Computer Science and Information Security**, Balasore, v. 7, Fevereiro 2010.

PETERMANN, R. J., 2006, **Modelo de Mineração de Dados para Classificação de Clientes em Telecomunicações**. Dissertação de M.Sc., Faculdade de Engenharia Elétrica/PUC-RS, Porto Alegre, RS, Brasil.

RAJESH. “Application of Spatial Data Mining for Agriculture”. **International Journal of Computer Applications**, Fevereiro 2011.

SBC, D. D. E. Currículo de Referência da SBC para Cursos de Graduação em Computação e Informática. **SBC-Sociedade Brasileira de Computação**, 2003.

Disponível em:

<http://www.sbc.org.br/index.php?option=com_jdownloads&Itemid=195&task=finish&cid=185&catid=36>. Acesso em: 14 out. 2011, 20:45:15.

SCHAEFFER, A. A., 2003, **Data Mining no Varejo: estudo de caso para loja de materiais de construção**. Dissertação de M.Sc., Instituto de Informática/UFRGS, Porto Alegre, RS, Brasil.

SCHNEIDER, L. F., 2003, **Aplicação do Processo de Descoberta de Conhecimento em Dados do Poder Judiciário do Estado do Rio Grande do Sul**. Dissertação de M.Sc., Instituto de Informática/UFRGS, Porto Alegre, RS, Brasil.

TAMILSELVI, J., GIFTA, B. “Handling Duplicate Data in Data Warehouse for Data Mining”. **International Journal of Computer Applications**, Fevereiro 2011.

TAN, P.-N., STEINBACH, M.; KUMAR, V. **Introdução ao Data Mining**. Rio de Janeiro: Editora Ciência Moderna Ltda, 2009.

TSIPTSIS, K.; CHORIANOPOULOS, A. **Data Mining Techniques in CRM: Inside Customer Segmentation**. Chichester: John Wiley & Sons Ltd, 2009.

TURBAN, E., SHARDA, R., DELEN, D., et al. **Decision Support and Business Intelligence Systems**. 9ªEdição. ed. New Jersey: Pearson , 2010.

WINCK, A. T., 2007, **Processo de KDD para auxílio à reconfiguração de ambientes virtualizados**. Dissertação de M.Sc., Faculdade de Informática/PUC-RS, Porto Alegre, RS, Brasil.

WITTEN, I. H.; EIBE, F.; HALL, M. A. **Data mining: practical machine learning tools and techniques**. 3ªEdição. ed. Burlington: Elsevier Inc, 2011.

Anexos

Anexo A – Questionários Socioeconômico e Cultural

Questionário Socioeconômico e Cultural utilizado nos vestibulares dos anos 2002 e 2003.

- 1 - Qual o seu sexo?
 - 1 - Masculino
 - 2 - Feminino
- 2 - Qual a sua idade completa?
 - 1 - Até 17 anos
 - 2 - 18 anos
 - 3 - 19 anos
 - 4 - 20 a 22 anos
 - 5 - 23 a 26 anos
 - 6 - 27 a 30 anos
 - 7 - Acima de 30 anos
- 3 - Qual seu estado civil?
 - 1 - Solteiro (a)
 - 2 - Casado (a)
 - 3 - Viúvo (a)
 - 4 - Outros
- 4 - Onde reside atualmente?
 - 1 - Anápolis
 - 2 - Goiânia
 - 3 - Outra cidade do interior do Estado de Goiás
 - 4 - Capital de outros estados
 - 5 - Interior de outros estados
- 5 - Qual o tipo de residência de sua família?
 - 1 - Própria
 - 2 - Alugada
 - 3 - Cedida
 - 4 - Outros
- 6 - Como você mora atualmente?
 - 1 - Sozinho, de aluguel
 - 2 - Sozinho, em residência própria
 - 3 - Em república
 - 4 - Em pensionato ou pensão
 - 5 - Em entidade assistencial
 - 6 - Com a própria família
 - 7 - Outros
- 7 - Que meio de transporte você utiliza?
 - 1 - Carro próprio
 - 2 - Carro da família
 - 3 - Coletivo
 - 4 - Moto
 - 5 - Outros meios
- 8 - Quantas pessoas contribuem para a obtenção da renda de sua família?
 - 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Mais de seis
- 9 - Quantas pessoas vivem da renda mensal de seu grupo familiar?
 - 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Acima de sete
- 10 - Qual seu salário?
 - 1 - Nenhum
 - 2 - Menos de um salário mínimo
 - 3 - Um salário mínimo
 - 4 - Dois a três salários mínimos
 - 5 - Quatro a seis salários mínimos
 - 6 - Sete a nove salários mínimos
 - 7 - Dez ou mais salários mínimos
- 11 - Qual é a renda mensal de seu grupo familiar?
 - 1 - Um salário mínimo
 - 2 - Dois a três salários mínimos
 - 3 - Quatro a cinco salários mínimos
 - 4 - Seis a sete salários mínimos
 - 5 - Oito a dez salários mínimos
 - 6 - Acima de dez salários mínimos
- 12 - Qual é a sua participação na vida econômica do seu grupo familiar?
 - 1 - Não trabalho
 - 2 - Trabalho, mas recebo ajuda financeira da família ou de outras pessoas
 - 3 - Trabalho para o meu próprio sustento
 - 4 - Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família
 - 5 - Trabalho e sou o principal responsável pelo sustento da família
- 13 - Como pretende se manter durante o curso universitário?
 - 1 - Trabalhando
 - 2 - Com recursos da família
 - 3 - Com bolsa de estudos
 - 4 - Com ajuda dos outros
- 14 - Qual o nível de escolaridade de seu pai ou responsável?
 - 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação

- 15 - Qual o nível de escolaridade de sua mãe ou responsável?
- 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 16 - Em que tipo de estabelecimento escolar fez seus estudos?
- 1 - Somente público
 - 2 - Somente particular
 - 3 - Maior parte público
 - 4 - Maior parte particular
- 17 - Em que turno você fez integralmente ou em sua maior parte o Ensino Médio?
- 1 - Diurno
 - 2 - Noturno
- 18 - Qual modalidade de Ensino Médio você concluiu ou concluirá?
- 1 - Atual Ensino Médio
 - 2 - 2º grau compacto
 - 3 - Magistério
 - 4 - Supletivo
 - 5 - Técnico-profissionalizante
- 19 - Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?
- 1 - Nenhuma
 - 2 - Uma
 - 3 - Duas
 - 4 - Três
 - 5 - Mais de três
- 20 - Como você se preparou para o vestibular?
- 1 - Fiz cursinho por um ano
 - 2 - Fiz cursinho por menos de um ano
 - 3 - Estudei apenas durante a 3ª série regular
 - 4 - Estudei sozinho
- 21 - Além de livros didáticos, quantos livros você lê por mês?
- 1 - Nenhum
 - 2 - Um a dois
 - 3 - Três a cinco
 - 4 - Acima de cinco
- 22 - Qual o tempo que você reserva diariamente para seus estudos?
- 1 - Nenhuma hora
 - 2 - Uma hora
 - 3 - Duas horas
 - 4 - Três horas
 - 5 - Quatro horas
 - 6 - Tempo integral
- 23 - Há quantos anos está tentando ingressar no Ensino Superior?
- 1 - Este é o primeiro ano
 - 2 - Um ano
 - 3 - Dois anos
 - 4 - Três anos
 - 5 - Mais de três anos
- 24 - Você já iniciou algum curso superior?
- 1 - Não
 - 2 - Sim, mas abandonei
 - 3 - Sim, e já o concluí
 - 4 - Sim, já concluí um e estou cursando outro
 - 5 - Sim, já concluí um e abandonei o outro
- 25 - O que influenciou na sua opção de curso?
- 1 - Maiores oportunidades de trabalho
 - 2 - Relação candidatos-vagas
 - 3 - Já trabalho na área
 - 4 - A família e/ou terceiros
 - 5 - Maior retorno financeiro
 - 6 - Prestígio social
 - 7 - Contribuição para a sociedade
 - 8 - Outros motivos
- 26 - Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?
- 1 - Não domino nenhuma língua estrangeira
 - 2 - Domino o inglês
 - 3 - Domino o espanhol
 - 4 - Domino o francês
 - 5 - Domino o italiano
 - 6 - Domino mais de uma língua estrangeira
- 27 - De quais atividades abaixo você mais participa?
- 1 - Artísticas e culturais
 - 2 - Religiosas
 - 3 - Político-partidárias
 - 4 - Esportivas
 - 5 - Outras
- 28 - Qual o meio de comunicação que você mais utiliza para se manter informado?
- 1 - Jornal escrito
 - 2 - Jornal falado (TV)
 - 3 - Jornal falado (rádio)
 - 4 - Internet
 - 5 - Revistas
 - 6 - Outras fontes
- 29 - Você usa microcomputador?
- 1 - Não
 - 2 - Sim, de casa
 - 3 - Sim, do trabalho
 - 4 - Sim, de amigos
 - 5 - Sim, de outros locais
- 30 - Você ficou sabendo do Processo Seletivo da UEG por intermédio de?
- 1 - Jornais
 - 2 - TV
 - 3 - Escolas de Ensino Médio
 - 4 - Cursinho
 - 5 - Outros
- 31 - Qual a razão principal pela qual você escolheu a UEG?
- 1 - Pela qualidade dos cursos oferecidos
 - 2 - Pelo número de candidatos/vaga
 - 3 - Por ser gratuita
 - 4 - Pela proximidade de minha residência
- 32 - Se você tem sugestões para o aperfeiçoamento deste, processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:
- 1 - Da inscrição
 - 2 - Das provas
 - 3 - Dos cursos e vagas
 - 4 - Da matrícula
 - 5 - Das áreas de ensino
 - 6 - Dos conteúdos programáticos
 - 7 - Sugiro alteração em vários itens
 - 8 - Não sugiro alteração alguma

Questionário socioeconômico Cultural utilizado no vestibular do ano 2004.

- 1 - Qual o seu sexo?
 - 1 - Masculino
 - 2 - Feminino
- 2 - Qual a sua idade completa?
 - 1 - Até 17 anos
 - 2 - 18 anos
 - 3 - 19 anos
 - 4 - 20 a 22 anos
 - 5 - 23 a 26 anos
 - 6 - 27 a 30 anos
 - 7 - Acima de 30 anos
- 3 - Qual seu estado civil?
 - 1 - Solteiro (a)
 - 2 - Casado (a)
 - 3 - Viúvo (a)
 - 4 - Outros
- 4 - Onde reside atualmente?
 - 1 - Anápolis
 - 2 - Goiânia
 - 3 - Outra cidade do interior do Estado de Goiás
 - 4 - Capital de outros estados
 - 5 - Interior de outros estados
- 5 - Qual o tipo de residência de sua família?
 - 1 - Própria
 - 2 - Alugada
 - 3 - Cedida
 - 4 - Outros
- 6 - Como você mora atualmente?
 - 1 - Sozinho, de aluguel
 - 2 - Sozinho, em residência própria
 - 3 - Em república
 - 4 - Em pensionato ou pensão
 - 5 - Em entidade assistencial
 - 6 - Com a própria família
 - 7 - Outros
- 7 - Que meio de transporte você utiliza?
 - 1 - Carro próprio
 - 2 - Carro da família
 - 3 - Coletivo
 - 4 - Moto
 - 5 - Outros meios
- 8 - Quantas pessoas contribuem para a obtenção da renda de sua família?
 - 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Mais de seis
- 9 - Quantas pessoas vivem da renda mensal de seu grupo familiar?
 - 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Acima de sete
- 10 - Qual seu salário?
 - 1 - Nenhum
 - 2 - Menos de um salário mínimo
 - 3 - Um salário mínimo
 - 4 - Dois a três salários mínimos
 - 5 - Quatro a seis salários mínimos
 - 6 - Sete a nove salários mínimos
 - 7 - Dez ou mais salários mínimos
- 11 - Qual é a renda mensal de seu grupo familiar?
 - 1 - Um salário mínimo
 - 2 - Dois a três salários mínimos
 - 3 - Quatro a cinco salários mínimos
 - 4 - Seis a sete salários mínimos
 - 5 - Oito a dez salários mínimos
 - 6 - Acima de dez salários mínimos
- 12 - Qual é a sua participação na vida econômica do seu grupo familiar?
 - 1 - Não trabalho
 - 2 - Trabalho, mas recebo ajuda financeira da família ou de outras pessoas
 - 3 - Trabalho para o meu próprio sustento
 - 4 - Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família
 - 5 - Trabalho e sou o principal responsável pelo sustento da família
- 13 - Como pretende se manter durante o curso universitário?
 - 1 - Trabalhando
 - 2 - Com recursos da família
 - 3 - Com bolsa de estudos
 - 4 - Com ajuda dos outros
- 14 - Qual o nível de escolaridade de seu pai ou responsável?
 - 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 15 - Qual o nível de escolaridade de sua mãe ou responsável?
 - 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 16 - Em que tipo de estabelecimento escolar fez seus estudos?
 - 1 - Somente público
 - 2 - Somente particular
 - 3 - Maior parte público
 - 4 - Maior parte particular
- 17 - Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?
 - 1 - Nenhuma
 - 2 - Uma
 - 3 - Duas

- 4 - Três Processo
5 - Mais de três
- 18 - Como você se preparou para o vestibular?
1 - Fiz cursinho por um ano
2 - Fiz cursinho por menos de um ano
3 - Estudei apenas durante a 3ª série regular
4 - Estudei sozinho
- 19 - Além de livros didáticos, quantos livros você lê por mês?
1 - Nenhum
2 - Um a dois
3 - Três a cinco
4 - Acima de cinco
- 20 - Qual o tempo que você reserva diariamente para seus estudos?
1 - Nenhuma hora
2 - Uma hora
3 - Duas horas
4 - Três horas
5 - Quatro horas
6 - Tempo integral
- 21 - Há quantos anos está tentando ingressar no Ensino Superior?
1 - Este é o primeiro ano
2 - Um ano
3 - Dois anos
4 - Três anos
5 - Mais de três anos
- 22 - Você já iniciou algum curso superior?
1 - Não
2 - Sim, mas abandonei
3 - Sim, e já o conclui
4 - Sim, já conclui um e estou cursando outro
5 - Sim, já conclui um e abandonei o outro
- 23 - O que influenciou na sua opção de curso?
1 - Maiores oportunidades de trabalho
2 - Relação candidatos-vagas
3 - Já trabalho na área
4 - A família e/ou terceiros
5 - Maior retorno financeiro
6 - Prestígio social
7 - Contribuição para a sociedade
- 24 - Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?
1 - Não domino nenhuma língua estrangeira
2 - Domino o inglês
3 - Domino o espanhol
4 - Domino o francês
5 - Domino o italiano
6 - Domino mais de uma língua estrangeira
- 25 - De quais atividades abaixo você mais participa?
1 - Artísticas e culturais
2 - Religiosas
3 - Político-partidárias
4 - Esportivas
- 26 - Qual o meio de comunicação que você mais utiliza para se manter informado?
1 - Jornal escrito
2 - Jornal falado (TV)
3 - Jornal falado (rádio)
4 - Internet
5 - Revistas
- 27 - Você usa microcomputador?
1 - Não
2 - Sim, de casa
3 - Sim, do trabalho
4 - Sim, de amigos
5 - Sim, de outros locais
- 28 - Você ficou sabendo do Processo Seletivo da UEG por intermédio de?
1 - Jornais
2 - TV
3 - Escolas de Ensino Médio
4 - Cursinho
5 - Internet
6 - Amigos, vizinhos ou parentes
- 29 - Qual a razão principal pela qual você escolheu a UEG?
1 - Pela qualidade dos cursos oferecidos
2 - Pelo número de candidatos/vaga
3 - Por ser gratuita
4 - Pela proximidade de minha residência
- 30 - Se você tem sugestões para o aperfeiçoamento deste, processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:
1 - Da inscrição
2 - Das provas
3 - Dos cursos e vagas
4 - Da matrícula
5 - Das áreas de ensino
6 - Dos conteúdos programáticos
7 - Sugiro alteração em vários itens
8 - Não sugiro alteração alguma
- 31 - Você é a favor da reserva de cotas para o ingresso nas Universidades Brasileiras?
1 - Sim
2 - Não
- 32 - Você aprova as mudanças no Provão do MEC?
1 - Sim
2 - Não

Questionário socioeconômico Cultural utilizado no vestibular do ano 2005.

- 1 - Qual o seu sexo?
1 - Masculino
2 - Feminino
- 2 - Qual a sua idade completa?
1 - Até 17 anos
2 - 18 anos
3 - 19 anos
4 - 20 a 22 anos

- 5 - 23 a 26 anos
6 - 27 a 30 anos
7 - Acima de 30 anos
- 3 - Qual seu estado civil?
1 - Solteiro (a)
2 - Casado (a)
3 - Viúvo (a)
4 - Outros
- 4 - Onde reside atualmente?
1 - Anápolis
2 - Goiânia
3 - Outra cidade do interior do Estado de Goiás
4 - Capital de outros estados
5 - Interior de outros estados
- 5 - Qual o tipo de residência de sua família?
1 - Própria
2 - Alugada
3 - Cedida
4 - Outros
- 6 - Como você mora atualmente?
1 - Sozinho, de aluguel
2 - Sozinho, em residência própria
3 - Em república
4 - Em pensionato ou pensão
5 - Em entidade assistencial
6 - Com a própria família
7 - Outros
- 7 - Que meio de transporte você utiliza?
1 - Carro próprio
2 - Carro da família
3 - Coletivo
4 - Moto
5 - Outros meios
- 8 - Quantas pessoas contribuem para a obtenção da renda de sua família?
1 - Uma
2 - Duas
3 - Três a quatro
4 - Cinco a seis
5 - Mais de seis
- 9 - Quantas pessoas vivem da renda mensal de seu grupo familiar?
1 - Uma
2 - Duas
3 - Três a quatro
4 - Cinco a seis
5 - Acima de sete
- 10 - Qual seu salário?
1 - Nenhum
2 - Menos de um salário mínimo
3 - Um salário mínimo
4 - Dois a três salários mínimos
5 - Quatro a seis salários mínimos
6 - Sete a nove salários mínimos
7 - Dez ou mais salários mínimos
- 11 - Qual é a renda mensal de seu grupo familiar?
1 - Um salário mínimo
2 - Dois a três salários mínimos
3 - Quatro a cinco salários mínimos
4 - Seis a sete salários mínimos
5 - Oito a dez salários mínimos
6 - Acima de dez salários mínimos
- 12 - Qual é a sua participação na vida econômica do seu grupo familiar?
1 - Não trabalho
2 - Trabalho, mas recebo ajuda financeira da família ou de outras pessoas
3 - Trabalho para o meu próprio sustento
4 - Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família
5 - Trabalho e sou o principal responsável pelo sustento da família
- 13 - Como pretende se manter durante o curso universitário?
1 - Trabalhando
2 - Com recursos da família
3 - Com bolsa de estudos
4 - Com ajuda dos outros
- 14 - Qual o nível de escolaridade de seu pai ou responsável?
1 - Nenhuma instrução
2 - Ensino Fundamental incompleto
3 - Ensino Fundamental completo
4 - Ensino Médio incompleto
5 - Ensino Médio completo
6 - Ensino Superior incompleto
7 - Ensino Superior completo
8 - Pós-graduação
- 15 - Qual o nível de escolaridade de sua mãe ou responsável?
1 - Nenhuma instrução
2 - Ensino Fundamental incompleto
3 - Ensino Fundamental completo
4 - Ensino Médio incompleto
5 - Ensino Médio completo
6 - Ensino Superior incompleto
7 - Ensino Superior completo
8 - Pós-graduação
- 16 - Em que tipo de estabelecimento escolar fez seus estudos?
1 - Somente público
2 - Somente particular
3 - Maior parte público
4 - Maior parte particular
- 17 - Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?
1 - Nenhuma
2 - Uma
3 - Duas
4 - Três
5 - Mais de três
- 18 - Como você se preparou para o vestibular?
1 - Fiz cursinho por um ano
2 - Fiz cursinho por menos de um ano
3 - Estudei apenas durante a 3ª série regular
4 - Estudei sozinho
- 19 - Além de livros didáticos, quantos livros você lê por mês?
1 - Nenhum
2 - Um a dois
3 - Três a cinco
4 - Acima de cinco
- 20 - Qual o tempo que você reserva diariamente para seus estudos?
1 - Nenhuma hora
2 - Uma hora
3 - Duas horas
4 - Três horas
5 - Quatro horas
6 - Tempo integral
- 21 - Há quantos anos está tentando ingressar no Ensino Superior?
1 - Este é o primeiro ano
2 - Um ano
3 - Dois anos
4 - Três anos
5 - Mais de três anos

- 22 - Você já iniciou algum curso superior?
- 1 - Não
 - 2 - Sim, mas abandonei
 - 3 - Sim, e já o conclui
 - 4 - Sim, já conclui um e estou cursando outro
 - 5 - Sim, já conclui um e abandonei o outro
- 23 - O que influenciou na sua opção de curso?
- 1 - Maiores oportunidades de trabalho
 - 2 - Relação candidatos-vagas
 - 3 - Já trabalho na área
 - 4 - A família e/ou terceiros
 - 5 - Maior retorno financeiro
 - 6 - Prestígio social
 - 7 - Contribuição para a sociedade
- 24 - Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?
- 1 - Não domino nenhuma língua estrangeira
 - 2 - Domino o inglês
 - 3 - Domino o espanhol
 - 4 - Domino o francês
 - 5 - Domino o italiano
 - 6 - Domino mais de uma língua estrangeira
- 25 - De quais atividades abaixo você mais participa?
- 1 - Artísticas e culturais
 - 2 - Religiosas
 - 3 - Político-partidárias
 - 4 - Esportivas
- 26 - Qual o meio de comunicação que você mais utiliza para se manter informado?
- 1 - Jornal escrito
 - 2 - Jornal falado (TV)
 - 3 - Jornal falado (rádio)
 - 4 - Internet
 - 5 - Revistas
- 27 - Você usa microcomputador?
- 1 - Não
 - 2 - Sim, de casa
 - 3 - Sim, do trabalho
 - 4 - Sim, de amigos
 - 5 - Sim, de outros locais
- 28 - Você ficou sabendo do Processo Seletivo da UEG por intermédio de?
- 1 - Jornais
 - 2 - TV
 - 3 - Escolas de Ensino Médio
 - 4 - Cursinho
 - 5 - Internet
 - 6 - Amigos, vizinhos ou parentes
- 29 - Qual a razão principal pela qual você escolheu a UEG?
- 1 - Pela qualidade dos cursos oferecidos
 - 2 - Pelo número de candidatos/vaga
 - 3 - Por ser gratuita
 - 4 - Pela proximidade de minha residência
- 30 - Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:
- 1 - Da inscrição
 - 3 - Dos cursos e vagas
 - 4 - Da matrícula
 - 5 - Das áreas de ensino
 - 6 - Dos conteúdos programáticos
 - 7 - Sugiro alteração em vários itens
 - 8 - Não sugiro alteração alguma
- 31 - Você é a favor da reserva de cotas para o ingresso na Universidade Estadual de Goiás – UEG?
- 1 - Sim
 - 2 - Não
- 32 - Como você se considera:
- 1 - Branco (a)
 - 2 - Negro (a)
 - 3 - Pardo (a) mulato (a)
 - 4 - Amarelo (a) (de origem oriental)
 - 5 - Indígena

Questionário socioeconômico Cultural utilizado no vestibular do ano 2006.

- 1 - Qual o seu sexo?
- 1 - Masculino
 - 2 - Feminino
- 2 - Qual a sua idade completa?
- 1 - Até 17 anos
 - 2 - 18 anos
 - 3 - 19 anos
 - 4 - 20 a 22 anos
 - 5 - 23 a 26 anos
 - 6 - 27 a 30 anos
 - 7 - Acima de 30 anos
- 3 - Qual seu estado civil?
- 1 - Solteiro (a)
 - 2 - Casado (a)
 - 3 - Viúvo (a)
 - 4 - Outros
- 4 - Onde reside atualmente?
- 1 - Anápolis
 - 2 - Goiânia
 - 3 - Outra cidade do interior do Estado de Goiás
 - 4 - Capital de outros estados
 - 5 - Interior de outros estados
- 5 - Qual o tipo de residência de sua família?
- 1 - Própria
 - 2 - Alugada
 - 3 - Cedida
 - 4 - Outros
- 6 - Como você mora atualmente?
- 1 - Sozinho, de aluguel
 - 2 - Sozinho, em residência própria
 - 3 - Em república
 - 4 - Em pensionato ou pensão
 - 5 - Em entidade assistencial
 - 6 - Com a própria família
 - 7 - Outros
- 7 - Que meio de transporte você utiliza?
- 1 - Carro próprio
 - 2 - Carro da família
 - 3 - Coletivo
 - 4 - Moto
 - 5 - Outros meios
- 8 - Quantas pessoas contribuem para a obtenção da renda

- de sua família?
- 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Mais de seis
- 9 - Quantas pessoas vivem da renda mensal de seu grupo familiar?
- 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Acima de seis
- 10 - Qual seu salário?
- 1 - Nenhum
 - 2 - Menos de um salário mínimo
 - 3 - Um salário mínimo
 - 4 - Dois a três salários mínimos
 - 5 - Quatro a seis salários mínimos
 - 6 - Sete a nove salários mínimos
 - 7 - Dez ou mais salários mínimos
- 11 - Qual é a renda mensal de seu grupo familiar?
- 1 - Um salário mínimo
 - 2 - Dois a três salários mínimos
 - 3 - Quatro a cinco salários mínimos
 - 4 - Seis a sete salários mínimos
 - 5 - Oito a dez salários mínimos
 - 6 - Acima de dez salários mínimos
- 12 - Qual é a sua participação na vida econômica do seu grupo familiar?
- 1 - Não trabalho
 - 2 - Trabalho, mas recebo ajuda financeira da família ou de outras pessoas
 - 3 - Trabalho para o meu próprio sustento
 - 4 - Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família
 - 5 - Trabalho e sou o principal responsável pelo sustento da família
- 13 - Como pretende se manter durante o curso universitário?
- 1 - Trabalhando
 - 2 - Com recursos da família
 - 3 - Com bolsa de estudos
 - 4 - Com ajuda dos outros
- 14 - Qual o nível de escolaridade de seu pai ou responsável?
- 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 15 - Qual o nível de escolaridade de sua mãe ou responsável?
- 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 16 - Em que tipo de estabelecimento escolar fez seus estudos?
- 1 - Somente público
 - 2 - Somente particular
 - 3 - Maior parte público
 - 4 - Maior parte particular
- 17 - Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?
- 1 - Nenhuma
 - 2 - Uma
 - 3 - Duas
 - 4 - Três
 - 5 - Mais de três
- 18 - Como você se preparou para o vestibular?
- 1 - Fiz cursinho por um ano
 - 2 - Fiz cursinho por menos de um ano
 - 3 - Estudei apenas durante a 3ª série regular
 - 4 - Estudei sozinho
- 19 - Além de livros didáticos, quantos livros você lê por mês?
- 1 - Nenhum
 - 2 - Um a dois
 - 3 - Três a cinco
 - 4 - Acima de cinco
- 20 - Qual o tempo que você reserva diariamente para seus estudos?
- 1 - Nenhuma hora
 - 2 - Uma hora
 - 3 - Duas horas
 - 4 - Três horas
 - 5 - Quatro horas
 - 6 - Tempo integral
- 21 - Há quantos anos está tentando ingressar no Ensino Superior?
- 1 - Este é o primeiro ano
 - 2 - Um ano
 - 3 - Dois anos
 - 4 - Três anos
 - 5 - Mais de três anos
- 22 - Você já iniciou algum curso superior?
- 1 - Não
 - 2 - Sim, mas abandonei
 - 3 - Sim, e já o conclui
 - 4 - Sim, já conclui um e estou cursando outro
 - 5 - Sim, já conclui um e abandonei o outro
- 23 - O que influenciou na sua opção de curso?
- 1 - Maiores oportunidades de trabalho
 - 2 - Relação candidatos-vagas
 - 3 - Já trabalho na área
 - 4 - A família e/ou terceiros
 - 5 - Maior retorno financeiro
 - 6 - Prestígio social
 - 7 - Contribuição para a sociedade
- 24 - Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?
- 1 - Não domino nenhuma língua estrangeira
 - 2 - Domino o inglês
 - 3 - Domino o espanhol
 - 4 - Domino o francês
 - 5 - Domino o italiano
 - 6 - Domino mais de uma língua estrangeira
- 25 - De quais atividades abaixo você mais participa?
- 1 - Artísticas e culturais
 - 2 - Religiosas
 - 3 - Político-partidárias
 - 4 - Esportivas
- 26 - Qual o meio de comunicação que você mais utiliza para se manter informado?
- 1 - Jornal escrito
 - 2 - Jornal falado (TV)
 - 3 - Jornal falado (rádio)
 - 4 - Internet
 - 5 - Revistas

- 27 - Você usa microcomputador?
- 1 - Não
 - 2 - Sim, de casa
 - 3 - Sim, do trabalho
 - 4 - Sim, de amigos
 - 5 - Sim, de outros locais
- 28 - Você ficou sabendo do Processo Seletivo da UEG por intermédio de?
- 1 - Jornais
 - 2 - TV
 - 3 - Escolas de Ensino Médio
 - 4 - Cursinho
 - 5 - Internet
 - 6 - Amigos, vizinhos ou parentes
- 29 - Qual a razão principal pela qual você escolheu a UEG?
- 1 - Pela qualidade dos cursos oferecidos
 - 2 - Pelo número de candidatos/vaga
 - 3 - Por ser gratuita
 - 4 - Pela proximidade de minha residência
- 30 - Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:
- 1 - Da inscrição
 - 2 - Das provas
 - 3 - Dos cursos e vagas
 - 4 - Da matrícula
 - 5 - Das áreas de ensino
 - 6 - Do sistema de cotas
 - 7 - Dos conteúdos programáticos
 - 8 - Sugiro alteração em vários itens
 - 9 - Não sugiro alteração alguma
- 31 - Que condição você aponta como obstáculo de acesso ao Ensino Superior?
- 1 - Econômica
 - 2 - Racial
 - 3 - Social
 - 4 - Geográfica
- 32 - Como você se considera:
- 1 - Branco (a)
 - 2 - Negro (a)
 - 3 - Pardo (a) mulato (a)
 - 4 - Amarelo (a) (de origem oriental)
 - 5 - Indígena
- 33 - Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?
- 1 - Não.
 - 2 - Apenas de um Processo Seletivo.
 - 3 - Mais de um Processo Seletivo.

Questionário socioeconômico Cultural utilizado no vestibular do ano 2007.

- 1- Qual o seu sexo?
- 1 - Masculino
 - 2 - Feminino
- 2 - Qual a sua idade completa?
- 1 - Até 17 anos
 - 2 - 18 anos
 - 3 - 19 anos
 - 4 - 20 a 22 anos
 - 5 - 23 a 26 anos
 - 6 - 27 a 30 anos
 - 7 - Acima de 30 anos
- 3 - Qual seu estado civil?
- 1 - Solteiro (a)
 - 2 - Casado (a)
 - 3 - Viúvo (a)
 - 4 - Outros
- 4 - Onde reside atualmente?
- 1 - Anápolis
 - 2 - Goiânia
 - 3 - Outra cidade do interior do Estado de Goiás
 - 4 - Capital de outros estados
 - 5 - Interior de outros estados
- 5 - Qual o tipo de residência de sua família?
- 1 - Própria
 - 2 - Alugada
 - 3 - Cedida
 - 4 - Outros
- 6 - Como você mora atualmente?
- 1 - Sozinho, de aluguel
 - 2 - Sozinho, em residência própria
 - 3 - Em república
 - 4 - Em pensionato ou pensão
 - 5 - Em entidade assistencial
 - 6 - Com a própria família
 - 7 - Outros
- 7 - Que meio de transporte você utiliza?
- 1 - Carro próprio
 - 2 - Carro da família
 - 3 - Coletivo
 - 4 - Moto
 - 5 - Outros meios
- 8 - Quantas pessoas contribuem para a obtenção da renda de sua família?
- 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Mais de seis
- 9 - Quantas pessoas vivem da renda mensal de seu grupo familiar?
- 1 - Uma
 - 2 - Duas
 - 3 - Três a quatro
 - 4 - Cinco a seis
 - 5 - Acima de seis
- 10 - Qual seu salário?
- 1 - Nenhum
 - 2 - Menos de um salário mínimo
 - 3 - Um salário mínimo
 - 4 - Dois a três salários mínimos
 - 5 - Quatro a seis salários mínimos
 - 6 - Sete a nove salários mínimos
 - 7 - Dez ou mais salários mínimos
- 11 - Qual é a renda mensal de seu grupo familiar?
- 1 - Um salário mínimo
 - 2 - Dois a três salários mínimos
 - 3 - Quatro a cinco salários mínimos
 - 4 - Seis a sete salários mínimos
 - 5 - Oito a dez salários mínimos
 - 6 - Acima de dez salários mínimos

- 12 - Qual é a sua participação na vida econômica do seu grupo familiar?
- 1 - Não trabalho
 - 2 - Trabalho, mas recebo ajuda financeira da família ou de outras pessoas
 - 3 - Trabalho para o meu próprio sustento
 - 4 - Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família
 - 5 - Trabalho e sou o principal responsável pelo sustento da família
- 13 - Como pretende se manter durante o curso universitário?
- 1 - Trabalhando
 - 2 - Com recursos da família
 - 3 - Com bolsa de estudos
 - 4 - Com ajuda dos outros
- 14 - Qual o nível de escolaridade de seu pai ou responsável?
- 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 15 - Qual o nível de escolaridade de sua mãe ou responsável?
- 1 - Nenhuma instrução
 - 2 - Ensino Fundamental incompleto
 - 3 - Ensino Fundamental completo
 - 4 - Ensino Médio incompleto
 - 5 - Ensino Médio completo
 - 6 - Ensino Superior incompleto
 - 7 - Ensino Superior completo
 - 8 - Pós-graduação
- 16 - Em que tipo de estabelecimento escolar fez seus estudos?
- 1 - Somente público
 - 2 - Somente particular
 - 3 - Maior parte público
 - 4 - Maior parte particular
- 17 - Qual o curso de Ensino Médio que você concluiu ou está concluindo?
- 1 - Técnico (Eletrônica, Química, Contabilidade Agrícola, etc)
 - 2 - Magistério
 - 3 - Colegial, antigo científico ou clássico
 - 4 - Supletivo
- 18 - Em que não que você concluiu ou concluirá o Ensino Médio?
- 1 - 2006
 - 2 - 2005
 - 3 - 2004
 - 4 - 2003
 - 5 - 1991-2002
 - 6 - Antes de 1991
- 19 - Em que período você cursou o Ensino Médio?
- 1 - Todo diurno
 - 2 - Maior parte diurno
 - 3 - Todo noturno
 - 4 - Maior parte noturno
- 20 - Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?
- 1 - Nenhuma
 - 2 - Uma
 - 3 - Duas
 - 4 - Três
 - 5 - Mais de três
- 21 - Como você se preparou para o vestibular?
- 1 - Fiz cursinho por um ano
 - 2 - Fiz cursinho por menos de um ano
 - 3 - Estudei apenas durante a 3ª série regular
 - 4 - Estudei sozinho
- 22 - Além de livros didáticos, quantos livros você lê por mês?
- 1 - Nenhum
 - 2 - Um a dois
 - 3 - Três a cinco
 - 4 - Acima de cinco
- 23 - Da relação abaixo, qual a disciplina com que você tem maior afinidade?
- 1 - Língua Portuguesa
 - 2 - Matemática
 - 3 - Geografia
 - 4 - História
 - 5 - Química
 - 6 - Biologia
 - 7 - Física
 - 8 - Língua Estrangeira
- 24 - Qual o tempo que você reserva diariamente para seus estudos?
- 1 - Nenhuma hora
 - 2 - Uma hora
 - 3 - Duas horas
 - 4 - Três horas
 - 5 - Quatro horas
 - 6 - Tempo integral
- 25 - Há quantos anos está tentando ingressar no Ensino Superior?
- 1 - Este é o primeiro ano
 - 2 - Um ano
 - 3 - Dois anos
 - 4 - Três anos
 - 5 - Mais de três anos
- 26 - Você já iniciou algum curso superior?
- 1 - Não
 - 2 - Sim, mas abandonei
 - 3 - Sim, e já o conclui
 - 4 - Sim, já conclui um e estou cursando outro
 - 5 - Sim, já conclui um e abandonei o outro
- 27 - Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?
- 1 - Não
 - 2 - Sim, em uma
 - 3 - Sim, em duas
 - 4 - Sim, em três
 - 5 - Sim, em mais de três
- 28 - O que influenciou na sua opção de curso?
- 1 - Maiores oportunidades de trabalho
 - 2 - Relação candidatos-vagas
 - 3 - Já trabalho na área
 - 4 - A família e/ou terceiros
 - 5 - Maior retorno financeiro
 - 6 - Prestígio social
 - 7 - Contribuição para a sociedade
- 29 - O que você espera obter num curso superior?
- 1 - Nível Superior para melhorar a atividade que já desenvolve
 - 2 - Formação profissional teórica voltada para futuro emprego
 - 3 - Formação teórica voltada para o ensino e a pesquisa
 - 4 - Aumento de conhecimento e cultura geral
 - 5 - Consciência crítica que possibilite a interação na sociedade
 - 6 - Diploma de nível superior

- 7 - Outro motivo
- 30 - Se aprovado(a), qual será a distância entre a sua provável moradia e a UEG?
- 1 - Até 20 Km
 - 2 - De 20 a 50 Km
 - 3 - De 50 a 100 Km
 - 4 - Mais de 100 Km
- 31 - Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?
- 1 - Com a sua própria família
 - 2 - Com os parentes
 - 3 - Sozinho(a)
 - 4 - Pensão, pensionato ou república
 - 5 - Outra
- 32 - Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?
- 1 - Não domino nenhuma língua estrangeira
 - 2 - Domino o inglês
 - 3 - Domino o espanhol
 - 4 - Domino o francês
 - 5 - Domino o italiano
 - 6 - Domino mais de uma língua estrangeira
- 33 - De quais atividades abaixo você mais participa?
- 1 - Artísticas e culturais
 - 2 - Religiosas
 - 3 - Político-partidárias
 - 4 - Esportivas
- 34 - Qual o meio de comunicação que você mais utiliza para se manter informado?
- 1 - Jornal escrito
 - 2 - Jornal falado (TV)
 - 3 - Jornal falado (rádio)
 - 4 - Internet
 - 5 - Revistas
- 35 - Você usa microcomputador?
- 1 - Não
 - 2 - Sim, de casa
 - 3 - Sim, do trabalho
 - 4 - Sim, de amigos
 - 5 - Sim, de outros locais
- 36 - Você ficou sabendo do Processo Seletivo da UEG por intermédio de?
- 1 - Jornais
 - 2 - TV
 - 3 - Escolas de Ensino Médio
 - 4 - Cursinho
 - 5 - Internet
 - 6 - Amigos, vizinhos ou parentes
- 37 - Qual a razão principal pela qual você escolheu a UEG?
- 1 - Pela qualidade dos cursos oferecidos
 - 2 - Pelo número de candidatos/vaga
 - 3 - Por ser gratuita
 - 4 - Pela proximidade de minha residência
- 38 - Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:
- 1 - Da inscrição
 - 2 - Das provas
 - 3 - Dos cursos e vagas
 - 4 - Da matrícula
 - 5 - Das áreas de ensino
 - 6 - Do sistema de cotas
 - 7 - Dos conteúdos programáticos
 - 8 - Sugiro alteração em vários itens
 - 9 - Não sugiro alteração alguma
- 39 - Que condição você aponta como obstáculo de acesso ao Ensino Superior?
- 1 - Econômica
 - 2 - Racial
 - 3 - Social
 - 4 - Geográfica
- 40 - Como você se considera:
- 1 - Branco (a)
 - 2 - Negro (a)
 - 3 - Pardo (a) mulato (a)
 - 4 - Amarelo (a) (de origem oriental)
 - 5 - Indígena
- 41 - Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?
- 1 - Não.
 - 2 - Apenas de um Processo Seletivo.
 - 3 - Mais de um Processo Seletivo.

Anexo B – Alternativas categorizadas do questionário socioeconômico e cultural

Conteúdo da tabela “Questões Vestibular” que representa as perguntas do questionário socioeconômico e cultura que tiveram suas alternativas categorizadas para serem utilizadas no terceiro ciclo da mineração de dados.

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q1	Além de livros didáticos, quantos livros você lê por ano?	Um a dois	Q1-A
Q1	Além de livros didáticos, quantos livros você lê por ano?	Um a dois.	Q1-A
Q1	Além de livros didáticos, quantos livros você lê por ano?	Três a cinco	Q1-B
Q1	Além de livros didáticos, quantos livros você lê por ano?	Três a cinco.	Q1-B
Q1	Além de livros didáticos, quantos livros você lê por ano?	Acima de cinco.	Q1-B
Q1	Além de livros didáticos, quantos livros você lê por ano?	Acima de cinco	Q1-B
Q1	Além de livros didáticos, quantos livros você lê por ano?	Nenhum	Q1-C
Q1	Além de livros didáticos, quantos livros você lê por ano?	Nenhum.	Q1-C
Q1	Além de livros didáticos, quantos livros você lê por ano?	Em branco	Q1-N
Q10	De quais atividades você mais participa?	Artísticas e culturais.	Q10-A
Q10	De quais atividades abaixo você mais participa?	Artísticas e culturais.	Q10-A
Q10	De quais atividades você mais participa?	Artísticas e culturais	Q10-A
Q10	De quais atividades abaixo você mais participa?	Artísticas e culturais	Q10-A
Q10	De quais atividades abaixo você mais participa?	Esportivas	Q10-B
Q10	De quais atividades você mais participa?	Esportivas	Q10-B
Q10	De quais atividades abaixo você mais participa?	Esportivas.	Q10-B
Q10	De quais atividades você mais participa?	Esportivas.	Q10-B
Q10	De quais atividades abaixo você mais participa?	Outras	Q10-C
Q10	De quais atividades você mais participa?	Político-partidárias	Q10-D
Q10	De quais atividades abaixo você mais participa?	Político-partidárias.	Q10-D
Q10	De quais atividades abaixo você mais participa?	Político partidárias	Q10-D
Q10	De quais atividades abaixo você mais participa?	Religiosas	Q10-E
Q10	De quais atividades você mais participa?	Religiosas	Q10-E
Q10	De quais atividades abaixo você mais participa?	Religiosas.	Q10-E
Q10	De quais atividades você mais participa?	Religiosas.	Q10-E
Q10	De quais atividades você mais participa?	Em branco	Q10-N
Q10	De quais atividades abaixo você mais participa?	Resposta em branco	Q10-N
Q10	De quais atividades abaixo você mais participa?	Em branco	Q10-N
Q10	De quais atividades abaixo você mais participa?	Marcação Multipla	Q10-N
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	até 2003.	Q11-A
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	Antes de 2002.	Q11-A
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	1991-2002	Q11-A
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	Antes de 2002	Q11-A
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	até 2003	Q11-A
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2003.	Q11-B
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2003	Q11-B
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2004.	Q11-C
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2004	Q11-C
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2005	Q11-D

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2005.	Q11-D
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2006	Q11-E
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2006.	Q11-E
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2007	Q11-F
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2007.	Q11-F
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2008	Q11-G
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2008.	Q11-G
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2009	Q11-H
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2009.	Q11-H
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2010.	Q11-I
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	2010	Q11-I
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	Antes de 1991	Q11-J
Q11	Em que ano você concluiu ou concluirá o Ensino Médio?	Em branco	Q11-N
Q12	Em que período você cursou o Ensino Médio?	Todo diurno.	Q12-A
Q12	Em que período você cursou o Ensino Médio?	Todo diurno	Q12-A
Q12	Em que turno você fez integralmente ou em sua maior parte o Ensino Médio?	Diurno	Q12-A
Q12	Em que período você cursou o Ensino Médio?	Maior parte diurno	Q12-A
Q12	Em que período você cursou o Ensino Médio?	Maior parte diurno.	Q12-A
Q12	Em que período você cursou o Ensino Médio?	Todo noturno	Q12-B
Q12	Em que período você cursou o Ensino Médio?	Todo noturno.	Q12-B
Q12	Em que turno você fez integralmente ou em sua maior parte o Ensino Médio?	Noturno	Q12-B
Q12	Em que período você cursou o Ensino Médio?	Maior parte noturno	Q12-B
Q12	Em que período você cursou o Ensino Médio?	Maior parte noturno.	Q12-B
Q12	Em que turno você fez integralmente ou em sua maior parte o Ensino Médio?	Resposta em branco	Q12-N
Q12	Em que período você cursou o Ensino Médio?	Em branco	Q12-N
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Somente particular	Q13-A
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Maior parte particular.	Q13-A
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Somente particular.	Q13-A
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Maior parte particular	Q13-A
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Somente público.	Q13-B
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Maior parte público.	Q13-B
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Maior parte público	Q13-B
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Somente público	Q13-B
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Resposta em branco	Q13-N
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Em branco	Q13-N
Q13	Em que tipo de estabelecimento escolar fez seus estudos?	Outros	Q13-N
Q14	Qual seu salário?	Nenhum	Q14-A
Q14	Qual seu salário?	Nenhum.	Q14-A
Q14	Qual seu salário?	Um salário mínimo.	Q14-B
Q14	Qual seu salário?	Menos de um salário mínimo	Q14-B
Q14	Qual seu salário?	Menos de um salário mínimo.	Q14-B
Q14	Qual seu salário?	Um salário mínimo	Q14-B
Q14	Qual seu salário?	Quatro a seis salários mínimos.	Q14-C
Q14	Qual seu salário?	Dois a três salários mínimos	Q14-C
Q14	Qual seu salário?	Sete a nove salários mínimos	Q14-C
Q14	Qual seu salário?	Quatro a seis salários mínimos	Q14-C
Q14	Qual seu salário?	Dois a três salários mínimos.	Q14-C
Q14	Qual seu salário?	Dez ou mais salários mínimos	Q14-C
Q14	Qual seu salário?	Resposta em branco	Q14-N
Q14	Qual seu salário?	Em branco	Q14-N
Q15	Em sua vida escolar, até o Ensino Médio, você teve	Uma	Q15-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
	alguma reprovação?		
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Duas	Q15-A
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Duas.	Q15-A
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Mais de três	Q15-A
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Três	Q15-A
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Uma.	Q15-A
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Nenhuma.	Q15-B
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Nenhuma	Q15-B
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Em branco	Q15-N
Q15	Em sua vida escolar, até o Ensino Médio, você teve alguma reprovação?	Resposta em branco	Q15-N
Q16	Qual seu estado civil?	Casado (a)	Q16-A
Q16	Estado civil:	Casado	Q16-A
Q16	Estado civil?	Casado (a).	Q16-A
Q16	Estado civil:	Solteiro	Q16-B
Q16	Estado civil?	Solteiro (a).	Q16-B
Q16	Qual seu estado civil?	Solteiro (a)	Q16-B
Q16	Qual seu estado civil?	Viúvo (a)	Q16-B
Q16	Estado civil:	Em branco	Q16-N
Q16	Qual seu estado civil?	Resposta em branco	Q16-N
Q16	Qual seu estado civil?	Outros	Q16-N
Q16	Estado civil?	Outros.	Q16-N
Q16	Estado civil?	Em branco	Q16-N
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Um ano	Q17-A
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Este é o primeiro ano	Q17-A
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Este é o primeiro ano.	Q17-A
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Um ano.	Q17-A
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Três anos.	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Três anos	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Mais de três anos.	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Mais de três anos	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Dois anos	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Dois anos.	Q17-B
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Em branco	Q17-N
Q17	Há quantos anos está tentando ingressar no Ensino Superior?	Resposta em branco	Q17-N
Q18	Qual a sua idade completa?	18 anos	Q18-A
Q18	Idade completa?	18 anos.	Q18-A
Q18	Idade completa?	Até 17 anos.	Q18-A
Q18	Idade completa:	18 anos	Q18-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q18	Qual a sua idade completa?	Até 17 anos	Q18-A
Q18	Idade completa:	Até 17 anos	Q18-A
Q18	Qual a sua idade completa?	20 a 22 anos	Q18-B
Q18	Idade completa?	20 a 22 anos.	Q18-B
Q18	Qual a sua idade completa?	19 anos	Q18-B
Q18	Idade completa:	19 anos	Q18-B
Q18	Idade completa:	20 a 22 anos	Q18-B
Q18	Idade completa?	19 anos.	Q18-B
Q18	Idade completa?	23 a 26 anos.	Q18-C
Q18	Idade completa:	27 a 30 anos	Q18-C
Q18	Qual a sua idade completa?	27 a 30 anos	Q18-C
Q18	Idade completa?	27 a 30 anos.	Q18-C
Q18	Idade completa:	Acima de 30 anos	Q18-C
Q18	Qual a sua idade completa?	Acima de 30 anos	Q18-C
Q18	Idade completa:	23 a 26 anos	Q18-C
Q18	Qual a sua idade completa?	23 a 26 anos	Q18-C
Q18	Idade completa?	Acima de 30 anos.	Q18-C
Q18	Idade completa:	Em branco	Q18-N
Q18	Idade completa?	Em branco	Q18-N
Q18	Qual a sua idade completa?	Resposta em branco	Q18-N
Q19	Meio de transporte que você utiliza?	Coletivo.	Q18-B
Q19	Meio de transporte que você utiliza:	Outros meios	Q18-B
Q19	Meio de transporte que você utiliza?	Outros meios.	Q18-B
Q19	Meio de transporte que você utiliza:	Coletivo	Q18-B
Q19	Meio de transporte que você utiliza:	Em branco	Q18-N
Q19	Meio de transporte que você utiliza?	Em branco	Q18-N
Q19	Meio de transporte que você utiliza?	Carro próprio.	Q19-A
Q19	Meio de transporte que você utiliza:	Carro da família	Q19-A
Q19	Meio de transporte que você utiliza:	Carro próprio	Q19-A
Q19	Meio de transporte que você utiliza:	Moto	Q19-A
Q19	Meio de transporte que você utiliza?	Carro da família.	Q19-A
Q19	Meio de transporte que você utiliza?	Moto.	Q19-A
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o inglês	Q2-A
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o inglês.	Q2-A
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o inglês	Q2-A
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino mais de uma língua estrangeira	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o Italiano	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino mais de uma língua estrangeira	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o francês.	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o espanhol	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino mais de uma Língua Estrangeira	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o espanhol.	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o espanhol	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino mais de uma língua estrangeira.	Q2-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Domino o Italiano	Q2-B
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Não domino nenhuma Língua Estrangeira	Q2-C
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Não domino nenhuma língua estrangeira	Q2-C
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Não domino nenhuma língua estrangeira	Q2-C
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Não domino nenhuma língua estrangeira.	Q2-C
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Resposta em branco	Q2-N
Q2	Com relação ao domínio da língua estrangeira moderna, em que situação você se enquadra?	Em branco	Q2-N
Q20	O que influenciou na sua opção de curso?	Já trabalho na área.	Q20-A
Q20	O que influenciou na sua opção de curso?	Já trabalho na área	Q20-A
Q20	O que influenciou na sua opção de curso?	Prestígio social.	Q20-B
Q20	O que influenciou na sua opção de curso?	Maior retorno financeiro	Q20-B
Q20	O que influenciou na sua opção de curso?	Prestígio social	Q20-B
Q20	O que influenciou na sua opção de curso?	Maiores oportunidades de trabalho.	Q20-B
Q20	O que influenciou na sua opção de curso?	Maiores oportunidades de trabalho	Q20-B
Q20	O que influenciou na sua opção de curso?	Maior retorno financeiro.	Q20-B
Q20	O que influenciou na sua opção de curso?	A família e/ou terceiros	Q20-C
Q20	O que influenciou na sua opção de curso?	Relação de candidatos-vaga	Q20-C
Q20	O que influenciou na sua opção de curso?	Relação candidatos/vaga	Q20-C
Q20	O que influenciou na sua opção de curso?	Por outros motivos	Q20-C
Q20	O que influenciou na sua opção de curso?	Contribuição para a sociedade.	Q20-C
Q20	O que influenciou na sua opção de curso?	A família e/ou terceiros.	Q20-C
Q20	O que influenciou na sua opção de curso?	Relação candidatos/vaga.	Q20-C
Q20	O que influenciou na sua opção de curso?	Contribuição para a sociedade	Q20-C
Q20	O que influenciou na sua opção de curso?	Relação de candidatos/vaga	Q20-C
Q20	O que influenciou na sua opção de curso?	Influência de família e/ou terceiros	Q20-C
Q20	O que influenciou na sua opção de curso?	Relação candidatos-vaga.	Q20-C
Q20	O que influenciou na sua opção de curso?	A de família e/ou terceiros	Q20-C
Q20	O que influenciou na sua opção de curso?	Em branco	Q20-N
Q20	O que influenciou na sua opção de curso?	Resposta em branco	Q20-N
Q21	O que você espera obter num curso superior?	Consciência crítica que possibilite a interação na sociedade.	Q21-A
Q21	O que você espera obter num curso superior?	Outro motivo	Q21-A
Q21	O que você espera obter num curso superior?	Outro motivo.	Q21-A
Q21	O que você espera obter num curso superior?	Formação teórica voltada para o ensino e a pesquisa.	Q21-A
Q21	O que você espera obter num curso superior?	Formação teórica voltada para o ensino e a pesquisa	Q21-A
Q21	O que você espera obter num curso superior?	Aumento de conhecimento e cultura geral	Q21-A
Q21	O que você espera obter num curso superior?	Aumento de conhecimento e cultura geral.	Q21-A
Q21	O que você espera obter num curso superior?	Consciência crítica que possibilite a interação na sociedade	Q21-A
Q21	O que você espera obter num curso superior?	Diploma de nível superior.	Q21-A
Q21	O que você espera obter num curso superior?	Diploma de nível superior	Q21-A
Q21	O que você espera obter num curso superior?	Formação profissional teórica voltada para futuro emprego	Q21-B
Q21	O que você espera obter num curso superior?	Formação profissional teórica voltada para futuro emprego.	Q21-B
Q21	O que você espera obter num curso superior?	Nível Superior para melhorar a atividade que já desenvolve	Q21-B
Q21	O que você espera obter num curso superior?	Nível superior para melhorar a atividade	Q21-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
		que já desenvolve.	
Q21	O que você espera obter num curso superior?	Em branco	Q21-N
Q22	Onde reside atualmente?	Anápolis	Q22-A
Q22	Onde reside atualmente?	Região Metropolitana de Goiânia	Q22-B
Q22	Onde reside atualmente?	Região Metropolitana de Goiânia.	Q22-B
Q22	Onde reside atualmente?	Goiânia	Q22-B
Q22	Onde reside atualmente?	Sudoeste goiano.	Q22-C
Q22	Onde reside atualmente?	Sul goiano.	Q22-C
Q22	Onde reside atualmente?	Sul goiano	Q22-C
Q22	Onde reside atualmente?	Outros estados.	Q22-C
Q22	Onde reside atualmente?	Centro goiano - Eixo BR-153.	Q22-C
Q22	Onde reside atualmente?	Cidades de outros estados	Q22-C
Q22	Onde reside atualmente?	Entorno do Distrito Federal.	Q22-C
Q22	Onde reside atualmente?	Entorno do Distrito Federal	Q22-C
Q22	Onde reside atualmente?	Nordeste goiano.	Q22-C
Q22	Onde reside atualmente?	Norte goiano	Q22-C
Q22	Onde reside atualmente?	Norte goiano.	Q22-C
Q22	Onde reside atualmente?	Oeste goiano - Eixo GO-060	Q22-C
Q22	Onde reside atualmente?	Oeste goiano - Eixo GO-060.	Q22-C
Q22	Onde reside atualmente?	Interior de outros estados	Q22-C
Q22	Onde reside atualmente?	Outro estado	Q22-C
Q22	Onde reside atualmente?	Sudoeste goiano	Q22-C
Q22	Onde reside atualmente?	Região Centro Goiano (Eixo BR-153)	Q22-C
Q22	Onde reside atualmente?	Região Centro Goiano (Eixo BR-153).	Q22-C
Q22	Onde reside atualmente?	Região Entorno do Distrito Federal	Q22-C
Q22	Onde reside atualmente?	Região Entorno do Distrito Federal.	Q22-C
Q22	Onde reside atualmente?	Região Norte Goiano	Q22-C
Q22	Onde reside atualmente?	Região Oeste Goiano (Eixo GO-060).	Q22-C
Q22	Onde reside atualmente?	Região Sudeste Goiano (Estrada de Ferro).	Q22-C
Q22	Onde reside atualmente?	Região Sudoeste Goiano.	Q22-C
Q22	Onde reside atualmente?	Sudeste goiano - Estrada de Ferro	Q22-C
Q22	Onde reside atualmente?	Sudeste goiano - Estrada de Ferro.	Q22-C
Q22	Onde reside atualmente?	Outra cidade do interior do Estado de Goiás	Q22-C
Q22	Onde reside atualmente?	Capital de outros estados	Q22-C
Q22	Onde reside atualmente?	Centro goiano - Eixo BR-153	Q22-C
Q22	Onde reside atualmente?	Cidades de outros estados.	Q22-C
Q22	Onde reside atualmente?	Em branco	Q22-N
Q22	Onde reside atualmente?	Resposta em branco	Q22-N
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	2º grau compacto	Q23-A
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	Atual Ensino Médio	Q23-A
Q23	Qual o curso de Ensino Médio que você concluiu ou está concluindo?	Colegio, antigo científico ou clássico	Q23-A
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Regular.	Q23-A
Q23	Qual o curso de Ensino Médio que você concluiu ou está concluindo?	Regular	Q23-A
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Regular	Q23-A
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Profissional (técnico)	Q23-B
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	Magistério	Q23-B
Q23	Qual o curso de Ensino Médio que você concluiu ou está	Magistério	Q23-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
	concluindo?		
Q23	Qual o curso de Ensino Médio que você concluiu ou está concluindo?	Profissional (Técnico)	Q23-B
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Profissional (Técnico).	Q23-B
Q23	Qual o curso de Ensino Médio que você concluiu ou está concluindo?	Técnico (Eletrônica, Química, Contabilidade, Agrícola, etc)	Q23-B
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	Técnico-profissionalizante	Q23-B
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	Supletivo	Q23-C
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Supletivo	Q23-C
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Supletivo.	Q23-C
Q23	Qual o curso de Ensino Médio que você concluiu ou está concluindo?	Supletivo	Q23-C
Q23	Qual a modalidade de Ensino Médio você concluiu ou concluirá?	Resposta em branco	Q23-N
Q23	Qual a modalidade de Ensino Médio que você concluiu ou está concluindo?	Em branco	Q23-N
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pela qualidade dos cursos oferecidos.	Q24-A
Q24	Qual a razão principal de você ter escolhido a UEG é?	Pela qualidade dos cursos oferecidos.	Q24-A
Q24	Qual a razão principal de você ter escolhido a UEG é?	Pela qualidade dos cursos oferecidos	Q24-A
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pela qualidade dos cursos oferecidos	Q24-A
Q24	Qual a razão principal pela qual você escolheu a UEG?	Por ser gratuita	Q24-B
Q24	Qual a razão principal de você ter escolhido a UEG é?	Por ser gratuita	Q24-B
Q24	Qual a razão principal de você ter escolhido a UEG é?	Por ser gratuita.	Q24-B
Q24	Qual a razão principal pela qual você escolheu a UEG?	Por ser gratuita.	Q24-B
Q24	Qual a razão principal de você ter escolhido a UEG é?	Pela proximidade de minha residência	Q24-C
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pelo número de candidatos/vaga	Q24-C
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pela proximidade de minha residência.	Q24-C
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pelo número de candidatos-vaga.	Q24-C
Q24	Qual a razão principal de você ter escolhido a UEG é?	Pela proximidade de minha residência.	Q24-C
Q24	Qual a razão principal pela qual você escolheu a UEG?	Pela proximidade de minha residência	Q24-C
Q24	Qual a razão principal pela qual você escolheu a UEG?	Resposta em branco	Q24-N
Q24	Qual a razão principal pela qual você escolheu a UEG?	Em branco	Q24-N
Q24	Qual a razão principal de você ter escolhido a UEG é?	Em branco	Q24-N
Q25	Você já iniciou algum curso superior?	Não.	Q25-A
Q25	Você já iniciou algum curso superior?	Não	Q25-A
Q25	Você já iniciou algum curso superior?	Sim, mas abandonei.	Q25-B
Q25	Você já iniciou algum curso superior?	Sim, e já o conclui	Q25-B
Q25	Você já iniciou algum curso superior?	Sim, já conclui um e estou cursando outro	Q25-B
Q25	Você já iniciou algum curso superior?	Sim, mas abandonei	Q25-B
Q25	Você já iniciou algum curso superior?	Resposta em branco	Q25-N
Q25	Você já iniciou algum curso superior?	Em branco	Q25-N
Q26	Qual é a renda mensal de seu grupo familiar?	Um salário mínimo	Q26-A
Q26	Qual é a renda mensal de seu grupo familiar?	Um salário mínimo.	Q26-A
Q26	Qual é a renda mensal de seu grupo familiar?	Quatro a cinco salários mínimos.	Q26-B
Q26	Qual é a renda mensal de seu grupo familiar?	Quatro a cinco salários mínimos	Q26-B
Q26	Qual é a renda mensal de seu grupo familiar?	Dois a três salários mínimos.	Q26-B
Q26	Qual é a renda mensal de seu grupo familiar?	Dois a três salários mínimos	Q26-B
Q26	Qual é a renda mensal de seu grupo familiar?	Seis a sete salários mínimos.	Q26-C
Q26	Qual é a renda mensal de seu grupo familiar?	Seis a sete salários mínimos	Q26-C
Q26	Qual é a renda mensal de seu grupo familiar?	Oito a dez salários mínimos.	Q26-C
Q26	Qual é a renda mensal de seu grupo familiar?	Oito a dez salários mínimos	Q26-C

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q26	Qual é a renda mensal de seu grupo familiar?	Acima de dez salários mínimos.	Q26-C
Q26	Qual é a renda mensal de seu grupo familiar?	Acima de dez salários mínimos	Q26-C
Q26	Qual é a renda mensal de seu grupo familiar?	Em branco	Q26-N
Q26	Qual é a renda mensal de seu grupo familiar?	Resposta em branco	Q26-N
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Não trabalho.	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, mas recebo ajuda financeira da família ou de outras pessoas.	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, mas recebo ajuda financeira da família ou de outras pessoas	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, mas recebo ajuda financeira	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, e contribuo para sustento da família	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Não trabalho	Q27-A
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho e sou o principal responsável pelo sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho e sou o principal responsável pelo sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho e sou o principal responsável pelo sustento da família.	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho e sou responsável pelo sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho e sou responsável pelo sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho para o meu próprio sustento.	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, sou responsável pelo meu sustento e contribuo parcialmente para o sustento da família.	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho para o meu próprio sustento	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, sou responsável pelo meu sustento, e contribuo parcialmente para o sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Trabalho, sou responsável pelo meu sustento, e contribuo parcialmente para o sustento da família	Q27-B
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Resposta em branco	Q27-N
Q27	Qual é a sua participação na vida econômica do seu grupo familiar?	Em branco	Q27-N
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	TV.	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Jornais	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Jornais.	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Internet	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Folhetos.	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Televisão.	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	TV	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Cartazes.	Q28-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
	intermédio de?		
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Televisão	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Folhetos	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Internet.	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Cartazes	Q28-A
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Escolas de Ensino Médio.	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Escolas de Ensino Médio	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Outros	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Amigos, vizinhos ou parentes	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Amigos, vizinhos ou parentes.	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Cursinho	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Cursinho.	Q28-B
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Em branco	Q28-N
Q28	Você ficou sabendo do Processo Seletivo da UEG por intermédio de?	Resposta em branco	Q28-N
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Internet.	Q29-A
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Internet	Q29-A
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Revistas	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Revistas.	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal impresso	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal impresso.	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal escrito.	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal escrito	Q29-B
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal falado (TV)	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal falado (rádio)	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Outras fontes	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Televisão	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Televisão.	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Rádio	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Jornal falado (TV).	Q29-C
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Em branco	Q29-N
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Resposta em branco	Q29-N

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q29	Qual o meio de comunicação que você mais utiliza para se manter informado?	Marcação Multipla	Q29-N
Q3	Como pretende se manter durante o curso universitário?	Trabalhando	Q3-A
Q3	Como pretende se manter durante o curso universitário?	Trabalhando.	Q3-A
Q3	Como pretende se manter durante o curso universitário?	Com recursos da família.	Q3-B
Q3	Como pretende se manter durante o curso universitário?	Com recursos da família	Q3-B
Q3	Como pretende se manter durante o curso universitário?	Com bolsa de estudos.	Q3-B
Q3	Como pretende se manter durante o curso universitário?	Com bolsa de estudos	Q3-B
Q3	Como pretende se manter durante o curso universitário?	Com ajuda dos outros	Q3-B
Q3	Como pretende se manter durante o curso universitário?	Resposta em branco	Q3-N
Q3	Como pretende se manter durante o curso universitário?	Em branco	Q3-N
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Superior completo.	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Pós-graduação.	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Superior completo	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Pós-graduação	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Superior incompleto.	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Superior incompleto	Q30-A
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Nenhuma instrução.	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Fundamental completo	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Fundamental completo.	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Fundamental incompleto	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Fundamental incompleto.	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Médio completo	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Médio completo.	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Médio incompleto	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Ensino Médio incompleto.	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Nenhuma instrução	Q30-B
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Resposta em branco	Q30-N
Q30	Qual o nível de escolaridade de seu pai ou responsável?	Em branco	Q30-N
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Superior completo	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Superior completo.	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Superior incompleto	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Superior incompleto.	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Pós-graduação	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Pós-graduação.	Q31-A
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Fundamental completo.	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Nenhuma instrução	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Fundamental completo	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Nenhuma instrução.	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Médio incompleto.	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Médio incompleto	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Médio completo.	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Médio completo	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Fundamental incompleto.	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Ensino Fundamental incompleto	Q31-B
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Em branco	Q31-N
Q31	Qual o nível de escolaridade de sua mãe ou responsável?	Resposta em branco	Q31-N
Q32	Qual o seu tipo lazer predileto?	Cinema	Q32-A
Q32	Qual o seu lazer predileto?	Música.	Q32-A
Q32	Qual seu lazer predileto?	Leituras	Q32-A
Q32	Qual o seu lazer predileto?	Leituras.	Q32-A
Q32	Qual o seu tipo lazer predileto?	Leituras.	Q32-A
Q32	Qual o seu tipo lazer predileto?	Música	Q32-A
Q32	Qual seu lazer predileto?	Música	Q32-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q32	Qual o seu tipo lazer predileto?	Música.	Q32-A
Q32	Qual o seu tipo lazer predileto?	Cinema.	Q32-A
Q32	Qual o seu tipo lazer predileto?	Leituras	Q32-A
Q32	Qual seu lazer predileto?	Cinema	Q32-A
Q32	Qual o seu lazer predileto?	Cinema.	Q32-A
Q32	Qual o seu tipo lazer predileto?	Shopping	Q32-B
Q32	Qual o seu tipo lazer predileto?	Festas	Q32-B
Q32	Qual o seu tipo lazer predileto?	Encontrar os amigos.	Q32-B
Q32	Qual o seu tipo lazer predileto?	Encontrar os amigos	Q32-B
Q32	Qual o seu lazer predileto?	Festas.	Q32-B
Q32	Qual o seu lazer predileto?	Encontrar os amigos.	Q32-B
Q32	Qual seu lazer predileto?	Festas	Q32-B
Q32	Qual o seu tipo lazer predileto?	Festas.	Q32-B
Q32	Qual seu lazer predileto?	Encontrar os amigos	Q32-B
Q32	Qual o seu tipo lazer predileto?	Jogos	Q32-C
Q32	Qual o seu tipo lazer predileto?	Jogos.	Q32-C
Q32	Qual o seu lazer predileto?	Jogos.	Q32-C
Q32	Qual seu lazer predileto?	Jogos	Q32-C
Q32	Qual o seu tipo lazer predileto?	Outros	Q32-N
Q32	Qual seu lazer predileto?	Outros	Q32-N
Q32	Qual o seu lazer predileto?	Outros.	Q32-N
Q32	Qual o seu tipo lazer predileto?	Outros.	Q32-N
Q32	Qual o seu tipo lazer predileto?	Em branco	Q32-N
Q32	Qual o seu lazer predileto?	Em branco	Q32-N
Q33	Sexo?	Feminino.	Q33-A
Q33	Sexo:	Feminino	Q33-A
Q33	Qual o seu sexo?	Feminino	Q33-A
Q33	Qual o seu sexo?	Masculino	Q33-B
Q33	Sexo:	Masculino	Q33-B
Q33	Sexo?	Masculino.	Q33-B
Q33	Qual o seu sexo?	Marcação Multipla	Q33-N
Q33	Sexo?	Em branco	Q33-N
Q33	Sexo:	Em branco	Q33-N
Q33	Qual o seu sexo?	Resposta em branco	Q33-N
Q34	Tipo de residência de sua família:	Própria	Q34-A
Q34	Tipo de residência de sua família?	Própria.	Q34-A
Q34	Qual o tipo de residência de sua família?	Própria	Q34-A
Q34	Qual o tipo de residência de sua família?	Alugada	Q34-B
Q34	Tipo de residência de sua família?	Cedida.	Q34-B
Q34	Tipo de residência de sua família:	Cedida	Q34-B
Q34	Qual o tipo de residência de sua família?	Cedida	Q34-B
Q34	Tipo de residência de sua família:	Alugada	Q34-B
Q34	Tipo de residência de sua família?	Alugada.	Q34-B
Q34	Qual o tipo de residência de sua família?	Resposta em branco	Q34-N
Q34	Tipo de residência de sua família:	Em branco	Q34-N
Q34	Tipo de residência de sua família?	Outros.	Q34-N
Q34	Tipo de residência de sua família:	Outros	Q34-N
Q34	Tipo de residência de sua família?	Em branco	Q34-N
Q34	Qual o tipo de residência de sua família?	Outros	Q34-N
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Duas.	Q35-A
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Uma	Q35-A
Q35	Quantas pessoas vivem da renda mensal de seu grupo	Uma.	Q35-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
	familiar?		
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Três a quatro.	Q35-A
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Três a quatro	Q35-A
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Duas	Q35-A
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Acima de seis	Q35-B
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Cinco a seis.	Q35-B
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Cinco a seis	Q35-B
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Acima de seis.	Q35-B
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Acima de sete	Q35-B
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Resposta em branco	Q35-N
Q35	Quantas pessoas vivem da renda mensal de seu grupo familiar?	Em branco	Q35-N
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Geográfica	Q36-A
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Geográfica.	Q36-A
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Econômica.	Q36-A
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Econômica	Q36-A
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Racial	Q36-B
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Racial.	Q36-B
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Regional.	Q36-B
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Social	Q36-B
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Social.	Q36-B
Q36	Que condição você aponta como obstáculo de acesso ao Ensino Superior?	Em branco	Q36-N
Q37	Que meio de transporte você utiliza?	Moto	Q37-A
Q37	Que meio de transporte você utiliza?	Carro da família	Q37-A
Q37	Que meio de transporte você utiliza?	Carro próprio	Q37-A
Q37	Que meio de transporte você utiliza?	Coletivo	Q37-B
Q37	Que meio de transporte você utiliza?	Outros meios	Q37-N
Q37	Que meio de transporte você utiliza?	Resposta em branco	Q37-N
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Com a própria família	Q38-A
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Com a própria família.	Q38-A
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Com os parentes	Q38-A
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Com os parentes.	Q38-A
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Com sua própria família	Q38-A
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Sozinho(a)	Q38-B
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Sozinho(a).	Q38-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
	mais provável situação de moradia?		
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Pensão, pensionato ou república.	Q38-B
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Pensão, pensionato ou república	Q38-B
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Outra.	Q38-N
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Em branco	Q38-N
Q38	Se aprovado(a) neste Processo Seletivo, qual será a sua mais provável situação de moradia?	Outra	Q38-N
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	Até 20 km.	Q39-A
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a UEG?	Até 20 Km	Q39-A
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	Até 20 km.	Q39-A
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	Até 20 km	Q39-A
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	Até 20 Km	Q39-A
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	De 51 a 100 km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	Mais de 100 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	Mais de 100 km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	De 20 a 50 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	Mais de 100 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a UEG?	Mais de 100 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	De 51 a 100 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	Mais de 100 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a UEG?	De 20 a 50 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	De 50 a 100 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a UEG?	De 50 a 100 Km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	De 21 a 50 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	De 20 a 50 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	De 50 a 100 km.	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	De 21 a 50 km	Q39-B
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde irá estudar?	Em branco	Q39-N
Q39	Se aprovado(a), qual será a distância entre a sua provável moradia e a Unidade da UEG onde vai estudar?	Em branco	Q39-N
Q4	Como você mora atualmente?	Com a própria família.	Q4-A
Q4	Como você mora atualmente?	Com a própria família	Q4-A
Q4	Como você mora atualmente:	Com a própria família	Q4-A
Q4	Como você mora atualmente?	Sozinho, de aluguel.	Q4-B
Q4	Como você mora atualmente?	Sozinho, de aluguel	Q4-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q4	Como você mora atualmente?	Sozinho, em residência própria	Q4-B
Q4	Como você mora atualmente?	Em república	Q4-B
Q4	Como você mora atualmente:	Em república	Q4-B
Q4	Como você mora atualmente?	Em pensionato ou pensão	Q4-B
Q4	Como você mora atualmente?	Em entidade assistencial.	Q4-B
Q4	Como você mora atualmente?	Em entidade assistencial	Q4-B
Q4	Como você mora atualmente:	Outros	Q4-N
Q4	Como você mora atualmente:	Em branco	Q4-N
Q4	Como você mora atualmente?	Em branco	Q4-N
Q4	Como você mora atualmente?	Resposta em branco	Q4-N
Q4	Como você mora atualmente?	Outros.	Q4-N
Q4	Como você mora atualmente?	Outros	Q4-N
Q40	Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?	Não	Q40-A
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Não.	Q40-A
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Não	Q40-A
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Não.	Q40-A
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Sim, em três.	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Sim, em três.	Q40-B
Q40	Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?	Sim, em três	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Sim, em mais de três.	Q40-B
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Sim, em uma	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Sim, em mais de três.	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Sim, em uma.	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Sim, em uma.	Q40-B
Q40	Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?	Sim, em uma	Q40-B
Q40	Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?	Sim, em mais de três	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Sim, em duas.	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Sim, em duas.	Q40-B
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Sim, em duas	Q40-B
Q40	Você pretende prestar Processo Seletivo para ingressar em outra universidade no 1º semestre de 2007?	Sim, em duas	Q40-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Sim, em três	Q40-B
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Sim, em mais de três	Q40-B
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no primeiro semestre de 2009?	Em branco	Q40-N
Q40	Você pretende submeter-se a processo seletivo para ingressar em outra universidade no segundo semestre de 2010?	Em branco	Q40-N
Q40	Você pretende submeter-se a Processo Seletivo para ingressar em outra universidade no segundo semestre de 2009?	Em branco	Q40-N
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Dos Conteúdos Programáticos	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Do sistema de cotas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Das provas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Das Áreas de Ensino	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Dos Cursos e Vagas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Da inscrição	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Sugiro alteração em vários itens.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Da Matrícula	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	O sistema de cotas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	O sistema de cotas.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Os cursos e as vagas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Os objetos de avaliação	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Sugiro alteração em vários itens	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Da inscrição.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	As provas.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Os cursos e vagas.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste	Do sistema de cotas.	Q41-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
	Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado		
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	As provas	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Das áreas de ensino.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Da matrícula.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Das provas.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Dos cursos e vagas.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Sugiro alteração em vários itens.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	A matrícula.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	As áreas de ensino	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	A inscrição	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	A matrícula	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	A inscrição.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	As áreas de ensino.	Q41-A
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Não sugiro alteração alguma.	Q41-B
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Não sugiro alteração alguma.	Q41-B
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Não sugiro alteração alguma	Q41-B
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Em branco	Q41-N
Q41	Se você tem sugestões para o aperfeiçoamento deste processo seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Resposta em branco	Q41-N
Q41	Se você tem sugestões para o aperfeiçoamento deste Processo Seletivo, indique qual dos itens abaixo deve ser prioritariamente melhorado	Em branco	Q41-N
Q42	Você usa microcomputador?	Sim, em outros locais.	Q42-A
Q42	Você usa microcomputador?	Sim, de outros locais	Q42-A
Q42	Você usa microcomputador?	Sim, de casa.	Q42-A
Q42	Você usa microcomputador?	Sim, de casa	Q42-A
Q42	Você usa microcomputador?	Sim, de amigos.	Q42-A
Q42	Você usa microcomputador?	Sim, de amigos	Q42-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q42	Você usa microcomputador?	Sim, do trabalho	Q42-A
Q42	Você usa microcomputador?	Sim, em casa.	Q42-A
Q42	Você usa microcomputador?	Sim, de outros locais.	Q42-A
Q42	Você usa microcomputador?	Sim, do trabalho.	Q42-A
Q42	Você usa microcomputador?	Não	Q42-B
Q42	Você usa microcomputador?	Não.	Q42-B
Q42	Você usa microcomputador?	Marcação Multipla	Q42-N
Q42	Você usa microcomputador?	Em branco	Q42-N
Q42	Você usa microcomputador?	Resposta em branco	Q42-N
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Tempo integral	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Duas horas	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Duas horas.	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Uma hora.	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Uma hora	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Três horas.	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Três horas	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Quatro horas	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Quatro horas.	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Tempo integral.	Q43-A
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Nenhuma	Q43-B
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Nenhuma hora	Q43-B
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Nenhuma hora.	Q43-B
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Marcação Multipla	Q43-N
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Em branco	Q43-N
Q43	Qual o tempo que você reserva diariamente para seus estudos?	Resposta em branco	Q43-N
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Uma.	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Uma	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de seu grupo familiar?	Duas	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de seu grupo familiar?	Uma	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Duas.	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Duas	Q44-A
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Três a quatro	Q44-B
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Três a quatro.	Q44-B
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Cinco a seis	Q44-B
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Três a quatro	Q44-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
	seu grupo familiar?		
Q44	Quantas pessoas contribuem para a obtenção da renda de seu grupo familiar?	Cinco a seis	Q44-B
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Cinco a seis.	Q44-B
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Em branco	Q44-N
Q44	Quantas pessoas contribuem para a obtenção da renda de seu grupo familiar?	Resposta em branco	Q44-N
Q44	Quantas pessoas contribuem para a obtenção da renda de sua família?	Resposta em branco	Q44-N
Q45	Você aprova as mudanças no provão do MEC?	Sim	Q45-A
Q45	Você aprova as mudanças no provão do MEC?	Não	Q45-B
Q45	Você aprova as mudanças no provão do MEC?	Resposta em branco	Q45-N
Q46	Você é a favor da reserva de cotas para o ingresso na Universidade Estadual de Goiás - UEG?	Sim	Q46-A
Q46	Você é a favor da reserva de cotas para o ingresso na Universidade Estadual de Goiás - UEG?	Não	Q46-B
Q46	Você é a favor da reserva de cotas para o ingresso na Universidade Estadual de Goiás - UEG?	Resposta inválida	Q46-N
Q46	Você é a favor da reserva de cotas para o ingresso na Universidade Estadual de Goiás - UEG?	Resposta em branco	Q46-N
Q47	Você é a favor da reserva de cotas para o ingresso na Universidades Brasileiras?	Sim	Q47-A
Q47	Você é a favor da reserva de cotas para o ingresso na Universidades Brasileiras?	Não	Q47-B
Q47	Você é a favor da reserva de cotas para o ingresso na Universidades Brasileiras?	Resposta em branco	Q47-N
Q48	Além de livros didáticos, quantos livros você lê por mês?	Um a dois	Q48-A
Q48	Além de livros didáticos, quantos livros você lê por mês?	Três a cinco	Q48-B
Q48	Além de livros didáticos, quantos livros você lê por mês?	Acima de cinco	Q48-B
Q48	Além de livros didáticos, quantos livros você lê por mês?	Nenhum	Q48-C
Q48	Além de livros didáticos, quantos livros você lê por mês?	Resposta em branco	Q48-N
Q5	Como você se considera:	Branco	Q5-A
Q5	Como você se considera:	Branco (a)	Q5-A
Q5	Como você se considera:	Branco(a).	Q5-A
Q5	Como você se considera:	Branco.	Q5-A
Q5	Como você se considera:	Negro (a)	Q5-B
Q5	Como você se considera:	Pardo(a) mulato(a).	Q5-B
Q5	Como você se considera:	Pardo mulato.	Q5-B
Q5	Como você se considera:	Pardo mulato	Q5-B
Q5	Como você se considera:	Pardo (a) mulato (a)	Q5-B
Q5	Como você se considera:	Negro(a).	Q5-B
Q5	Como você se considera:	Negro	Q5-B
Q5	Como você se considera:	Negro.	Q5-B
Q5	Como você se considera:	Amarelo (a) (de origem oriental)	Q5-C
Q5	Como você se considera:	Amarelo (de origem oriental)	Q5-C
Q5	Como você se considera:	Indígena.	Q5-D
Q5	Como você se considera:	Em branco	Q5-N
Q5	Como você se considera:	Resposta em branco	Q5-N
Q6	Como você se preparou para este processo seletivo?	Estudei apenas durante a 3ª série regular	Q6-A
Q6	Como você se preparou para este processo seletivo?	Estudei sozinho	Q6-A
Q6	Como você se preparou para o Processo Seletivo?	Estudei sozinho.	Q6-A
Q6	Como você se preparou para este Processo Seletivo?	Estudei sozinho.	Q6-A
Q6	Como você se preparou para o vestibular?	Estudei sozinho	Q6-A
Q6	Como você se preparou para este Processo Seletivo?	Estudei apenas durante a 3ª série regular.	Q6-A

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q6	Como você se preparou para o vestibular?	Estudei apenas durante a 3ª série regular	Q6-A
Q6	Como você se preparou para o Processo Seletivo?	Estudei apenas durante a 3ª série regular.	Q6-A
Q6	Como você se preparou para o Processo Seletivo?	Fiz cursinho por menos de um ano.	Q6-B
Q6	Como você se preparou para o Processo Seletivo?	Fiz cursinho por um ano ou mais.	Q6-B
Q6	Como você se preparou para este Processo Seletivo?	Fiz cursinho por um ano ou mais.	Q6-B
Q6	Como você se preparou para o vestibular?	Fiz cursinho por um ano	Q6-B
Q6	Como você se preparou para este Processo Seletivo?	Fiz cursinho por menos de um ano.	Q6-B
Q6	Como você se preparou para o vestibular?	Fiz cursinho por menos de um ano	Q6-B
Q6	Como você se preparou para este processo seletivo?	Fiz cursinho por menos de um ano	Q6-B
Q6	Como você se preparou para este processo seletivo?	Fiz cursinho por um ano ou mais	Q6-B
Q6	Como você se preparou para o Processo Seletivo?	Em branco	Q6-N
Q6	Como você se preparou para o vestibular?	Resposta em branco	Q6-N
Q6	Como você se preparou para este Processo Seletivo?	Em branco	Q6-N
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Da inscrição	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Dos Conteúdos Programáticos	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Dos Cursos e Vagas	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Sugiro alteração em vários itens	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Da Matrícula	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Das Áreas de Ensino	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Das provas	Q7-A
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Não sugiro alteração alguma	Q7-B
Q7	Se você tem sugestões para o aperfeiçoamento deste manual, indique qual dos itens abaixo deve ser prioritariamente melhorado:	Resposta em branco	Q7-N
Q8	Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?	Apenas de um Processo Seletivo	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Apenas de um Processo Seletivo	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Apenas de um Processo Seletivo.	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Não.	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Não optei por cotas.	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Não optei por cotas	Q8-A
Q8	Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?	Não	Q8-A
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Não	Q8-A
Q8	Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?	Mais de um Processo Seletivo	Q8-B
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Mais de um Processo Seletivo.	Q8-B

Questão	Questão Original	Alternativa Original	Alternativa Representada
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Mais de um Processo Seletivo	Q8-B
Q8	Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?	Resposta em branco	Q8-N
Q8	Você que optou pelo Sistema de Cotas já participou de outro Processo Seletivo?	Resposta inválida	Q8-N
Q8	Você, que optou pelo Sistema de Cotas, já participou de outro Processo Seletivo?	Em branco	Q8-N
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Matemática.	Q9-A
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Matemática	Q9-A
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Física	Q9-A
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Física.	Q9-A
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	História	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Língua Portuguesa.	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Língua Portuguesa	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Língua Estrangeira.	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	História.	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Geografia.	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Geografia	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Língua Estrangeira	Q9-B
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Biologia	Q9-C
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Biologia.	Q9-C
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Química.	Q9-C
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Química	Q9-C
Q9	Da relação abaixo, qual a disciplina com que você tem maior afinidade?	Em branco	Q9-N