



**UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS DE QUIXADÁ
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

GUILHERME DOS SANTOS CUNHA

**DESCOBERTA E ANÁLISE DOS PERFIS METODOLÓGICOS DOS PROFESSORES
DAS DISCIPLINAS INTRODUTÓRIAS A PROGRAMAÇÃO NO CAMPUS DA UFC
EM QUIXADÁ**

QUIXADÁ

2018

GUILHERME DOS SANTOS CUNHA

DESCOBERTA E ANÁLISE DOS PERFIS METODOLÓGICOS DOS PROFESSORES
DAS DISCIPLINAS INTRODUTÓRIAS A PROGRAMAÇÃO NO CAMPUS DA UFC EM
QUIXADÁ

Monografia apresentada no curso de Sistemas de Informação da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em Sistemas de Informação. Área de concentração: Computação.

Orientador: Prof. Dr. João Ferreira de Lavor

QUIXADÁ

2018

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C978d Cunha, Guilherme dos Santos.

Descoberta e análise dos perfis metodológicos dos professores das disciplinas introdutórias a programação no campus da UFC em Quixadá / Guilherme dos Santos Cunha. – 2018.
30 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Sistemas de Informação, Quixadá, 2018.
Orientação: Prof. Dr. João Ferreira de Lavor.

1. Metodologia de ensino. 2. Perfil do docente. 3. Mineração de dados (Computação). I. Título.

CDD 005

GUILHERME DOS SANTOS CUNHA

DESCOBERTA E ANÁLISE DOS PERFIS METODOLÓGICOS DOS PROFESSORES
DAS DISCIPLINAS INTRODUTÓRIAS A PROGRAMAÇÃO NO CAMPUS DA UFC EM
QUIXADÁ

Monografia apresentada no curso de Sistemas de Informação da Universidade Federal do Ceará, como requisito parcial à obtenção do título de bacharel em Sistemas de Informação. Área de concentração: Computação.

Aprovada em: ___/___/_____.

BANCA EXAMINADORA

Prof. Dr. João Ferreira de Lavor. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Marcos Antônio de Oliveira
Universidade Federal do Ceará (UFC)

Prof. Dr. David Sena Oliveira
Universidade Federal do Ceará (UFC)

Aos meus pais, e toda minha família.

AGRADECIMENTOS

Aos meus pais, pelo apoio e a educação que me possibilitou estar em uma universidade.

A minha família, irmão, tios e primos pelo apoio e torcida.

Ao Prof. Dr. João Lavor, pela disponibilidade e excelente orientação.

Aos professores David Sena e Marcos Oliveira, membros da banca, pelas valiosas contribuições.

Aos meus colegas de graduação, pelo companheirismo, ajuda e amizade.

A todos, alunos, funcionários, professores e amigos, que me ajudaram no decorrer do curso.

“Não existe triunfo sem perda, não há vitória sem sofrimento, não há liberdade sem sacrifício” (Tolkien)

RESUMO

A tecnologia vem se tornando cada vez mais presente no dia a dia de pessoas e organizações. Isso faz com que aumente a procura por cursos da área de Tecnologia da Informação (TI) nas instituições de ensino superior. Porém, já no seu início, estes cursos apresentam grandes desafios aos estudantes, uma vez que a grade curricular é composta por várias disciplinas complexas e importantes para o desenvolvimento do conhecimento necessário para sua área de atuação. Neste trabalho, buscamos identificar os perfis dos professores da disciplina introdutória a programação. Foram utilizados dados de questionários aplicados aos alunos destas disciplinas. O foco foi na abordagem e na metodologia do professor. Com o auxílio de ferramentas de mineração de dados, realizamos a clusterização destes dados, juntos com os dados de testes uniformes aplicados a todas as turmas. Destes dados, retiramos os perfis dos professores relacionados a médias de notas e a taxa de evasão das turmas. Com esses dados é possível maximizar as variáveis relacionadas a um melhor desempenho dos alunos, e assim, melhorar o desempenho das turmas posteriores. O Campus da UFC em Quixadá é um ótimo campo para este estudo, já que possui seis cursos com essa disciplina, com carga horária, estrutura e ementas semelhantes.

Palavras-chave: Mineração de dados. Perfil do docente. Metodologia de Ensino.

ABSTRACT

Technology has become more and more present in the daily lives of people and organizations. This increases the demand for courses in the area of Information Technology (IT) in higher education institutions. However, in the beginning, these courses present great challenges to students, since the curriculum is composed of several complex disciplines and important for the development of the knowledge necessary for their area of activity. In this work, we seek to identify the profiles of teachers from the introductory discipline to programming. Questionnaires applied to the students of these subjects were used. The focus was mainly on the teacher's approach and methodology. With the aid of data mining tools, we performed a clustering of this data, together with the uniform test data applied to all the classes. From these data, we extracted the profiles of the teachers related to averages of grades and the rate of evasion of the classes. With this data it is possible to maximize the variables related to a better performance of the students, and thus, to improve the performance of the subsequent classes. The UFC Campus in Quixadá is a great field for this study, since it has six courses with this discipline, with hours, structure and similar menus.

Keywords: Data Mining. Teacher profile. Teaching Methodology.

LISTA DE FIGURAS

Figura 1	Demonstração da fórmula da distância euclidiana	20
Figura 2	Dados em CSV	21
Figura 3	Atributos da clusterização	24
Figura 4	Interface do Weka com a clusterização dos perfis dos professores	30
Figura 5	Interface do Weka com a clusterização dos perfis dos professores com as médias	31
Figura 6	Interface do Weka com a clusterização dos perfis dos professores com a taxa de evasão	31

LISTA DE TABELAS

Tabela 1 –	Atributos escolhidos para clusterização	22
Tabela 2 –	Perfis dos professores	23
Tabela 3 –	Perfis dos professores com as médias	23
Tabela 4 –	Perfis dos professores com a taxa de evasão	24

LISTA DE SÍMBOLOS

% Porcentagem

SUMÁRIO

1	INTRODUÇÃO	14
2	TRABALHOS RELACIONADOS	17
3	FUNDAMENTAÇÃO TEÓRICA	18
3.1	Ensino de programação.....	18
3.2	Mineração de dados	19
3.3	Clusterização	20
4	PROCEDIMENTOS	22
4.1	Recebimento e Limpeza dos dados.....	22
4.2	Definições de análise.....	23
4.3	Aplicação da técnica de clusterização.....	23
5	RESULTADOS	26
5.1	Perfis dos professores.....	26
5.2	Análise dos perfis relacionados com as médias das turmas.....	27
5.3	Análise dos perfis relacionados com as taxas de evasão.....	27
6	CONCLUSÃO	29
7	CONSIDERAÇÕES FINAIS.....	30
	REFERÊNCIAS	31
	APÊNDICE	33

1 INTRODUÇÃO

Neste século, a tecnologia vem se tornando cada vez mais presente no dia a dia de pessoas e organizações. Isso faz com que aumente a procura por cursos da área de Tecnologia da Informação (TI) nas instituições de ensino superior. Porém, já no seu início, estes cursos apresentam grandes desafios aos estudantes, uma vez que a grade curricular é composta por várias disciplinas complexas e importantes para o desenvolvimento do conhecimento necessário para sua área de atuação. Estas disciplinas costumam apresentar um elevado índice de desistência, pois o aluno muitas vezes não tem noção do que encontrará no ambiente universitário, sendo que muitos iniciam sem ter um curso técnico na área ou habilidades em raciocínio lógico e matemático suficientes. Essa dificuldade é abordada por Souto e Duduchi (2008), que apontam que apesar das disciplinas que iniciam o ensino de programação serem apontadas como a base da formação acadêmica dos alunos dos cursos de computação, pode-se verificar que é justamente nelas que muitos alunos encontram sua primeira e maior dificuldade.

As disciplinas de Introdução à Programação são a base para o desenvolvimento lógico e algorítmico de graduandos da área de Computação. Estas disciplinas são essenciais para a construção da fundamentação necessária para o entendimento de tópicos mais avançados de disciplinas posteriores (França e Tedesco, 2015). Contudo, ensinar programação é dos grandes desafios da educação em Computação (Caspersen e Kolling, 2009) e pesquisas apontam que são várias as dificuldades enfrentadas pelos estudantes ao iniciarem seus estudos na área. Programar exige elevado nível de generalização, abstração e pensamento crítico. Motivação, por exemplo, também têm impacto na aprendizagem de Introdução à Programação (Rountree et al. 2004). Apesar de não haver um consenso quanto a melhor metodologia ou linguagem de programação a ser usada, observa-se que essa disciplina necessita de que o aluno resolva uma grande quantidade de problemas (Vier et al. 2015).

É importante lembrar que a apropriação ou não dos conceitos iniciais de programação tem relação direta com o desempenho do aluno no decorrer de todo o curso, já que disciplinas seguintes dependem fortemente desses conceitos.

No campus da UFC em Quixadá, a disciplina introdutória tem o nome de Fundamentos de Programação e Introdução a Programação para Design, ela é parte da grade curricular dos seis cursos do campus. Na disciplina são abordados os seguintes assuntos: Algoritmos, Conceitos Fundamentais de Programação, Expressões, Controles de Fluxo, Funções e Procedimentos, Ponteiros, Vetores e Matrizes, Cadeias de Caracteres, Alocação

Dinâmica, Tipos Estruturados e Arquivos (UFC Quixadá, 2008).

Esse trabalho buscou, com base nos dados de questionários respondidos por diversos alunos, identificar os perfis dos professores da disciplina de Fundamentos de Programação, focando principalmente em sua abordagem e metodologia, nos tipos de exercícios passados e nas ferramentas usadas para avaliação e atividades. Aliado a isso, analisamos resultados de testes aplicados a todas as turmas, para tentar identificar variáveis que podem influenciar as médias e a taxa de evasão. Com esses dados, é possível maximizar as variáveis relacionadas a um melhor desempenho dos alunos, e assim, melhorar o desempenho das turmas posteriores. O Campus da UFC em Quixadá é um ótimo campo para este estudo, já que possui seis cursos com essa disciplina, com carga horária, estrutura e ementas semelhantes.

Macedo (2016) analisou fatores sociais, econômicos e motivacionais importantes na distinção de aluno evadidos ou não. Seu trabalho utilizou respostas de um questionário aplicado aos ingressantes de 2015 para fazer uma análise de perfil e tentar identificar grupos de alunos com perfis similares.

Já Silva (2014), realizou um estudo utilizando mineração de dados a partir de técnicas de clusterização, para conhecer o perfil dos alunos que ingressam na universidade. O autor buscou responder a questionamentos sobre medidas adotadas pela UFC (Universidade Federal do Ceará) ao longo dos anos, como por exemplo, a nova regra de frequência de 2008.2 e a adesão ao SISU (Sistema de Seleção Unificado) em 2011.

Vieira, De Lima Júnior e De Paula, Vieira (2015) realizaram uma análise das dificuldades apresentadas pelos alunos ao cursar a disciplina de Algoritmo e Linguagem de Programação 1, em um curso de Sistemas de Informação. Para isso, foi feito um levantamento dos índices de aprovação/reprovação na disciplina no triênio 2010-2012 e aplicado um questionário junto aos alunos sobre as dificuldades de assimilação dos conteúdos da disciplina.

Neste trabalho, a técnica para mineração de dados escolhida é a clusterização, que consiste em encontrar grupos de objetos, levando em conta que objetos pertencentes a um determinado grupo devem ser similares entre si e diferentes dos objetos de outros grupos. O algoritmo de clusterização escolhido foi o K-Means, que funciona escolhendo aleatoriamente k registros para representar os centroides (médias) dos clusters. Todos os registros são colocados em determinado cluster com base na distância entre o registro e o centroide. A ferramenta utilizada como auxílio foi o Weka.

Este estudo está dividido da seguinte forma: no próximo capítulo, serão

apresentados os trabalhos relacionados, no capítulo seguinte será apresentada a fundamentação teórica, posteriormente, os procedimentos utilizados no desenvolvimento do trabalho. No quinto capítulo, são apresentados os resultados, no seguinte, as conclusões, e, por fim, as considerações finais.

2 TRABALHOS RELACIONADOS

Neste tópico serão apresentados os trabalhos relacionados, suas semelhanças e diferenças com este estudo.

Macedo (2016), analisou fatores sociais, econômicos e motivacionais importantes na distinção de aluno evadidos ou não, nos três primeiros semestres dos cursos da UFC-Quixadá. Ele utilizou respostas de um questionário aplicado aos ingressantes de 2015. Sua amostra foi de 237 alunos. Para fazer uma análise de perfil e tentar identificar grupos de alunos com perfis similares. Após a análise de perfil, o autor utilizou análise fatorial para sumarizar os dados. O estudo buscou trazer contribuições de possíveis motivos para a evasão discente. Este estudo, difere do trabalho de Macedo ao utilizar clusterização, como método de análise de dados, também difere ao buscar identificar os perfis dos professores, não dos alunos.

Silva (2014) realizou um estudo, utilizando mineração de dados, a partir de técnicas de clusterização, para conhecer o perfil dos alunos que ingressam na universidade. No trabalho foram analisados os perfis de alunos encontrados por disciplina e por semestre, a fim de saber como os perfis foram formados seguindo uma trilha de dependência de disciplina. A partir dessa análise, o autor procurou responder a questionamentos sobre medidas adotadas pela Universidade Federal do Ceará ao longo dos anos. Assim como neste trabalho, Silva (2014) utilizou o algoritmo de clusterização K-means, aliado a distância euclidiana, para notar repetições de um mesmo perfil ao longo das disciplinas estudadas. Porém, assim como o trabalho anterior, este busca identificar perfis de alunos, para atingir seus objetivos.

Vieira, De Lima Junior e De Paula, Vieira (2015) realizaram uma análise das dificuldades apresentadas pelos alunos ao cursar a disciplina de Algoritmo e Linguagem de Programação 1, em um curso de Sistemas de Informação. Para isso, foi feito um levantamento dos índices de aprovação/reprovação na disciplina no triênio 2010-2012 e aplicado um questionário junto aos alunos sobre as dificuldades de assimilação dos conteúdos da disciplina. Os autores, com os dados coletados, buscaram mapear os pontos que geram maiores dificuldades de aprendizado. Mas, diferentemente deste estudo, os autores não utilizaram nenhuma técnica de análise de dados. Eles também focaram a pesquisa apenas no conteúdo ministrado, e não na metodologia em si.

3 FUNDAMENTAÇÃO TEÓRICA

Neste terceiro tópico, serão fundamentados os conceitos principais utilizados neste trabalho. No primeiro tópico, os conceitos de ensino de programação, em seguida, um detalhamento maior sobre mineração de dados. No tópico final um aprofundamento sobre a técnica de clusterização, assim como do algoritmo usado nesta técnica, além da medida de similaridade utilizada: a distância euclidiana.

3.1 Ensino de Programação

Segundo a estrutura curricular constante nas Diretrizes Curriculares do MEC (Azeredo, 2000), a matéria de Programação faz parte da área de formação básica em Ciência da Computação, junto as matérias Computação e Algoritmos, e Arquitetura de Computadores. Seu conteúdo abrange, além do ensino de linguagens de programação propriamente ditas, os conceitos, os princípios e os modelos de programação e o estudo de estruturas de dados e de métodos de classificação e pesquisa de dados.

O ensino de programação gera grandes benefícios para a sociedade, interagindo com diversas áreas. Também proporciona benefícios para alunos dessa área, como por exemplo, o desenvolvimento de habilidades para resolução de problemas gerais (Fessakis et al. 2013). Em todo o mundo, vários países desenvolvidos já mobilizaram-se para implantar o ensino de programação à estudantes mais jovens, incluindo até mesmo crianças (Kalelioğlu, 2015).

Porém, a geração atual que ingressa nas universidades, é uma geração muito ligada a jogos e sistemas interativos no dia a dia, e interessada em aprender rapidamente a desenvolvê-los. E muitos, segundo Hu (2008), tem a ideia que desenvolverá sistemas similares já no início de seu aprendizado. Coutinho et al. (2008), completa afirmando que a expectativa dos novos alunos é que seus programas devem se comportar de maneira gráfica e interativa, assim como os sistemas e jogos a qual eles estão habituados.

Pereira et al. (2005), elenca cinco questões relacionadas ao processo de ensino e aprendizagem: 1. Dificuldade para interpretar o problema antes mesmo da dificuldade para representá-lo em algum modelo formal. 2. Dificuldade para identificar os conhecimentos prévios para que os alunos tenham condições de aprender e desenvolver soluções algorítmicas de forma satisfatória. 3. A necessidade de definir ou não o paradigma de programação para a estruturação das soluções dos problemas. 4. Dificuldade para se definir o nível e a maneira de

formalização desejada, que pode variar entre o uso rudimentar ou de forma mais estruturada (pseudocódigo) de linguagem natural, uso de formas gráficas de visualização tais como fluxogramas e a utilização de linguagens de programação específicas. 5. Habilidades e experiências distintas dos alunos que determinam diferentes ritmos de aprendizagem.

Concluindo, o estudo de programação não se restringe ao estudo de linguagens de programação. As linguagens de programação constituem-se em uma ferramenta de concretização de produto de software, que representa o resultado da aplicação de uma série de conhecimentos que transformam a especificação da solução de um problema em um programa de computador que efetivamente resolve aquele problema (Santos; Costa, 2005, p. 3).

3.2 Mineração de Dados

A Mineração de Dados é uma área multidisciplinar, com elementos de computação e estatística. Por ser uma área multidisciplinar, as definições variam de acordo com sua aplicação. A seguir, vemos três definições variadas.

Cabena et al. (1998), define voltado a área de banco de dados. Ele a apresenta como uma área de pesquisa que atua em reconhecimento de padrões, estatísticas, banco de dados e visualização. Hand et al. (2006) define como um conjunto de técnicas úteis para analisar grandes quantidades de dados, descobrir padrões e transformar essa informação em algo compreensível aos tomadores de decisão. Esse é um viés estatístico. Já Fayyad et al (1996), define sob uma ótica de aprendizagem de máquina. Ele a aponta como a principal etapa da descoberta de conhecimento, entre o pré-processamento dos dados e a análise dos padrões descobertos pela mineração.

Como podemos ver nas definições, um ponto que se repete, apesar das variações, é a descoberta de padrões. As diferentes técnicas utilizadas no processo de minerar os dados, são classificadas em duas categorias: As preditivas, que tentam prever valores desconhecidos ou valores futuros dos dados, como por exemplo, a regressão e a classificação, e as descritivas, que tentam encontrar padrões que ajudam em um melhor entendimento dos dados, como a técnica de associação e clusterização. Um exemplo de técnica preditiva, a classificação consiste em prever a que grupo pertence um determinado dado de um conjunto, para isso é fornecido um conjunto de registro já com as classes aos quais o registro pertence. Depois, o algoritmo aprende os padrões presentes nos dados. Então, quando um novo registro é adicionado aos dados, o algoritmo é capaz de prever a que classe ele pertence.

Já no lado das descritivas, a associação é utilizada para analisar a influência de um

dado sobre outro, dentro de um conjunto de dados.

A clusterização, como foi mencionado no início, será a técnica descritiva utilizada neste trabalho, e será abordada no tópico seguinte.

3.3 Clusterização

A Clusterização é uma tarefa descritiva, com o objetivo de grupos de objetos, ou padrões nos dados tais que, objetos pertencentes a um grupo sejam similares entre si e diferentes de objetos que pertencem a outro grupo.

De acordo com JAIN et al (1999), a atividade de clusterização envolve 5 passos:

1. Seleção de atributos: O processo de identificação de atributos distintos de um conjunto de atributos candidatos.

2. Medida de similaridade: A medida de similaridade determina o quanto um objeto é similar a outro, possibilitando assim, a identificação de qual cluster o objeto deve pertencer.

3. Algoritmo de clusterização: Atualmente existe uma grande variedade de algoritmos de clusterização, mas não há um algoritmo que possa ser usado para resolver todo e qualquer problema. Dessa forma, é fundamental analisar o problema para selecionar ou desenvolver um algoritmo adequado.

4. Validação dos clusters: Todo algoritmo de clusterização consegue gerar partições, independente de estas possuírem significado semântico. E, além disso, algoritmos diferentes normalmente agrupam os objetos de forma diferente. Por isso, é necessário que os resultados obtidos pelos algoritmos sejam validados por meio de critérios de avaliação eficientes.

5. Interpretação dos resultados: Por fim, o resultado dos algoritmos de clusterização deve ser interpretado por especialistas para que sejam atribuídos significados aos clusters de forma que o usuário possa compreendê-los.

O K-means foi o algoritmo utilizado neste trabalho. É um algoritmo simples. Seu objetivo é encontrar a melhor divisão de P dados em K grupos de maneira que a distância entre os dados de um grupo e seu centro, somado por todos os grupos, seja minimizada.

O algoritmo inicia escolhendo aleatoriamente k registros para representar os centroides (médias) dos clusters. Todos os registros são colocados em um determinado cluster com base na distância entre o registro e o centroide. Quando todos os registros já estiverem armazenados em um cluster, o centroide é recalculado. O processo se repete, examinando

cada registro novamente e o colocando no cluster cujo centroide é mais próximo (Silva, 2014).

A medida de similaridade estabelecida foi a distância euclidiana, que considera a distância d entre dois elementos X_i e X_j no espaço p -dimensional:

$$DE(x, y) = \sqrt{\sum_i^p (x_i - y_i)^2} \quad (1)$$

Fonte: Trovão e Dedecca (2014)

Algumas ferramentas podem auxiliar na utilização dessas técnicas: Podemos citar o Projeto R, o Rapid Miner e o WEKA (Waikato Environment for Knowledge Analysis), que foi a ferramenta utilizada neste trabalho.

4 PROCEDIMENTOS

Neste tópico, serão relatadas as etapas deste trabalho. Primeiro, o recebimento e limpeza dos dados. Após, a definição de quais aspectos das informações seriam analisados, e, finalmente, a aplicação da técnica de clusterização.

4.1 Recebimento e limpeza de dados

Os dados foram repassados ao autor deste trabalho pelo professor David Sena, da Universidade Federal do Ceará – Campus Quixadá. São dados colhidos em todas as turmas de fundamentos de programação e introdução a programação para design no mencionado Campus. Estes dados consistem em um questionário, aplicado a 103 alunos, abordando diversos aspectos sobre a metodologia do professor da disciplina, como por exemplo, Foco do Ensino (Se é voltado mais a lógica de programação ou aprendizado da linguagem), Abordagem do professor (Se o conhecimento é passado de forma que o professor é o único agente ativo ou se há uma construção participativa do conhecimento) e Tipo de exercícios (Se os exercícios possuem contexto ou são mais diretos, com características lógico-matemáticas). Também há dados sobre as notas dos alunos e médias em 9 testes idênticos passado a todas as turmas mencionadas.

Os dados do questionário, provenientes do GoogleForms, já estavam padronizados, e a plataforma permite download em CSV. Poucas respostas duplicadas foram encontradas (Com o mesmo nome de Aluno, e da mesma turma), então foi deixada a mais recente. Algumas modificações foram feitas nas tabelas, para melhor se adequar na mineração de dados. Por exemplo, o dado “Mini-Teste” originalmente estava em uma pergunta sobre os métodos de avaliação dos professores, que permitia uma ou mais respostas, na modificação, foi criada uma tabela exclusiva para respostas que continham o resultado “Mini-teste”, já que era um dado considerado importante para a análise. Com os dados dos questionários, foi criado um arquivo em formato ARFF, para se adequar ao requisito da ferramenta Weka.

Figura 1 – Dados do questionário em CSV

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Lista de exercicios	Escrita	Escrita	Não	5	5	2	1	5	2	4	5	3	5	4			
2	Trabalhos individuais/Juiz Online	Avaliação no computador;Mini testes	Avaliação no computador	Sim	5	5	5	5	5	4	5	5	2	3	3			
3	Lista de exercicios/Juiz Online	Avaliação no computador;Mini testes	Avaliação no computador	Sim	4	2	3	5	5	1	2	4	3	2	1			
4	Lista de exercicios	Avaliação no computador	Avaliação no computador	Não	5	1	2	5	3	1	1	3	5	5				
5	Lista de exercicios	Escrita	Escrita	Não	5	3	3	5	5	3	4	3	5	3				
6	Lista de exercicios/Juiz Online	Avaliação no computador;Mini testes	Avaliação no computador	Sim	5	4	5	4	4	3	4	5	1	2	4			
7	Lista de exercicios	Avaliação no computador	Avaliação no computador	Não	5	5	4	5	5	1	2	4	2	4	3			
8	Lista de exercicios/Juiz Online	Escrita;Avaliação no computador;Mini testes	Ambos	Sim	5	5	4	5	5	4	5	5	4	3				
9	Juiz Online	moodle	Vari	Não	5	5	5	5	4	4	4	4	3	3				
10	Lista de exercicios/Trabalhos individuais/Trabalhos em grupo	Escrita;Avaliação no computador	Ambos	Não	5	4	1	4	5	1	4	3	1	1	1			
11	Lista de exercicios/Trabalhos individuais/Trabalhos em grupo/Juiz Online	Avaliação no computador	Avaliação no computador	Não	5	5	3	5	5	1	4	4	3	3	2			
12	Lista de exercicios	Escrita	Escrita	Não	5	3	1	4	3	1	4	2	2	5	3			
13	Lista de exercicios	Escrita	Escrita	Não	5	4	1	3	4	1	2	3	2	3				
14	Lista de exercicios/Juiz Online	Escrita;Mini testes	Escrita	Sim	5	2	1	5	5	3	4	5	2	2	1			
15	Trabalhos individuais/Juiz Online	Trabalho e Prova	Vari	Não	5	5	2	5	5	1	5	5	4	5	3			
16	Lista de exercicios/Juiz Online	Avaliação no computador	Avaliação no computador	Não	5	4	3	5	5	4	4	5	4	5	1			
17	Lista de exercicios	Avaliação no computador	Avaliação no computador	Não	5	3	3	5	5	2	1	3	3	3				
18	Lista de exercicios	Avaliação no computador	Avaliação no computador	Não	2	1	1	5	3	2	1	2	3	1	4			
19	Lista de exercicios	Avaliação no computador	Avaliação no computador	Não	5	1	3	5	3	2	1	1	4	4	3			
20	Lista de exercicios/Trabalhos em grupo/Juiz Online	Mini testes	Vari	Sim	5	4	5	5	4	1	5	5	4	2	2			
21	Lista de exercicios	Escrita	Escrita	Não	5	4	3	5	5	4	4	5	2	3	4			
22	Lista de exercicios	Escrita	Escrita	Não	3	4	4	3	2	3	4	3	3	3				
23	Trabalhos individuais/Trabalhos em grupo/Juiz Online	Escrita;Avaliação no computador;Mini testes	Escrita	Sim	5	5	5	5	5	5	5	4	4	5				
24	Lista de exercicios/Trabalhos individuais	Escrita	Escrita	Não	5	4	3	5	5	3	5	3	4	4				
25	Lista de exercicios	Avaliação no computador;Mini testes	Avaliação no computador	Sim	5	3	5	2	3	1	3	5	3	1	5			
26	Trabalhos individuais/Juiz Online	Avaliação no computador;Mini testes	Avaliação no computador	Sim	5	4	4	5	5	4	4	5	3	5	5			
27	Trabalhos em grupo/Juiz Online	Escrita;Avaliação no computador	Ambos	Não	5	4	1	4	5	3	3	4	3	1	5			
28	Lista de exercicios/Trabalhos individuais/Trabalhos em grupo	Escrita;Mini testes	Escrita	Sim	5	4	2	5	5	4	5	5	5	5				
29	Juiz Online	Avaliação no computador	Avaliação no computador	Não	4	4	2	2	3	2	1	4	3	5	3			
30	Lista de exercicios/Trabalhos em grupo	Escrita	Escrita	Não	4	2	1	3	5	4	4	4	1	3	4			

Fonte: elaborada pelo autor.

4.2 Definições das análises

Após verificar e preparar os dados, decidimos analisar 2 pontos:

1. Os perfis dos professores das turmas com as melhores médias: Nesse ponto foi identificado as características (destacadas nos questionários) que estão presente com mais frequência nos perfis que possuem melhores médias nos testes em suas turmas. Nos dados originais, existem 9 testes, mas como os 9 não foram aplicados em todas as turmas, decidimos estender a análise apenas até o sétimo teste.

2. Perfis dos professores com as menores taxas de evasão: Baseado na frequência individual dos alunos nos testes, verificamos quais professores obtiveram mais êxito em ter uma maior quantidade de alunos nos testes finais. Assim como o primeiro ponto, foram identificados características comuns a esses professores. Neste ponto foram analisados todos os 9 testes, com os professores que os aplicaram, e 7 com os professores que aplicaram apenas essa quantidade. Para contar como aluno evadido, elegemos os alunos que tinham, no mínimo, três faltas seguidas nos últimos testes realizados nas turmas, tanto com 9, quanto com 7 testes.

4.3 Aplicação da técnica de clusterização

Na etapa final dos procedimentos, utilizamos a ferramenta para realizar a clusterização. Dos dados dos questionários, escolhemos alguns pontos, para incluirmos na mineração. Esses pontos foram escolhidos por sua importância, e pela sua variação maior entre os professores. Os pontos incluídos na mineração foram:

Tabela 1 – Atributos incluídos na clusterização

Atributo	Descrição
aulas_práticas	Se é realizado aulas práticas na maioria das aulas.
atividades_retribuidas	Se o professor bonifica diretamente os alunos pelas atividades realizadas.
recursos_complementares	Se o professor disponibiliza recursos para estudos complementares.
tira_duvidas	Se o professor tira dúvidas durante a semana.
cobra_participacao	Se o professor cobra participação ativa do aluno, até não dando presença para o aluno que ele julgar não participativo.
praticas_grupo	Se o professor realiza ou passa atividade práticas em grupo para os alunos.
dojos	“Dojo é um método de aprendizagem para desenvolvimento de projetos em computação em que todas as pessoas constroem juntas a solução, o professor oferece um desafio ao grupo o piloto, sentado ao computador, é a única pessoa que pode utilizá-lo para concluir o desafio. O copiloto permanece ao seu lado, mas somente para observar o piloto em ação e oferecer indicações. Os demais observam e podem discutir entre si e com o copiloto e o piloto.” - Descrição retirada diretamente do questionário.
anti_plagio	Se o professor possui uma política anti-plágio eficiente em atividades e avaliações.
abordagem_prof	Se a abordagem do professor é mais participativa, construindo o conhecimento com os alunos (1 - Totalmente construtiva), ou se ele o único agente ativo dentro da sala, que transmite o conhecimento, enquanto o aluno, passivo, apenas o recebe (5 – Totalmente expositiva).
exercicios_propostos	Se os exercícios passados são contextualizados (1-Totalmente contextualizados) ou são exercícios focados mais em conceitos

	matemáticos (5 - Totalmente Lógico-matemáticos)
foco_ensino	Se o professor dá maior foco a conceitos mais característicos da linguagem utilizada (1 - Linguagem) ou se ele foca nos conceitos mais lógicos e matemáticos (5 – Totalmente lógico).
juiz_online	Se o professor utiliza o Juiz Online (Moodle, por exemplo) como ferramenta de avaliação das atividades.
tipo_prova	O tipo de prova que o professor passa a turma, pode ser escrita, pelo computador, ambos ou outras (No caso, o “outras”, se refere a o que não for as duas primeiras. Por terem quantidades baixas decidimos agrupá-las em uma única variável)
mini_testes	Se o professor passa mini-testes avaliativos durante a disciplina, de modo a compor a nota do aluno.

Figura 2 – Atributos definidos no arquivo .ARFF

```

bmw-browsers.arff (-/Documentos/TCC/TCC2/Final) - gedit
@relation respostas
@attribute aulas_praticas numeric
@attribute atividades_retribuidas numeric
@attribute cobra_participacao numeric
@attribute recursos_complementares numeric
@attribute tira_duvidas numeric
@attribute dojos numeric
@attribute praticas_grupo numeric
@attribute anti_plagio numeric
@attribute abordagem_prof numeric
@attribute exercicios_propostos numeric
@attribute foco_ensino numeric
@attribute juiz_online {Sim, Não, Só}
@attribute tipo_prova {Escrita, Computador, Ambos, Vari}
@attribute mini_testes {Sim, Não}

@data
5,5,2,1,5,2,4,5,3,5,4,Não,Escrita,Não
5,5,5,5,4,5,5,2,3,3,Sim,Computador,Sim
4,2,3,5,5,1,2,4,3,2,1,Sim,Computador,Sim
5,1,2,5,3,1,1,3,5,5,5,Não,Computador,Não
5,3,3,5,5,3,4,3,5,3,3,Não,Escrita,Não
5,4,5,4,4,3,4,5,1,2,4,Sim,Computador,Sim
5,5,4,5,5,1,2,4,2,4,3,Não,Computador,Não
5,5,4,5,5,5,4,5,5,4,3,Sim,Ambos,Sim
5,5,5,5,4,4,4,4,3,3,Só,Var1,Não
5,4,1,4,5,1,4,3,1,1,1,Não,Ambos,Não
5,5,3,5,5,1,4,4,3,3,2,Sim,Computador,Não
5,3,1,4,3,1,4,2,2,5,3,Não,Escrita,Não
5,4,1,3,4,1,2,3,2,3,3,Não,Escrita,Não
5,2,1,5,5,3,4,5,2,2,1,Sim,Escrita,Sim
5,5,2,5,5,1,5,5,4,5,3,Sim,Var1,Não
5,4,3,5,5,4,4,5,4,5,1,Sim,Computador,Não
5,3,3,5,5,2,1,3,3,3,3,Não,Computador,Não
2,1,1,5,3,2,1,2,3,1,4,Não,Computador,Não
5,1,3,5,3,2,1,1,4,4,3,Não,Computador,Não
5,4,5,5,4,1,5,5,4,2,2,Sim,Var1,Sim
5,4,3,5,5,4,4,5,2,3,4,Não,Escrita,Não
3,4,4,3,2,3,3,4,3,3,3,Não,Escrita,Não

```

Fonte: elaborada pelo autor.

5 RESULTADOS

Neste tópico, serão apresentados os resultados da mineração e suas respectivas análises. Primeiro, será mostrado os perfis predominantes de professores. Em seguida, uma análise com as variáveis dos perfis dos professores com a melhor média em sua turma, e, por fim, uma análise levando em conta as taxas de evasão de cada turma.

5.1 Perfil dos professores

Tabela 1 – Perfil dos professores

Atributos	Perfis		Média
	1 (45%)	2 (55%)	
Aulas Práticas	4.7609	4.5789	4.6602
Atividades Retribuídas	4	3.5088	3.7282
Cobrança de Participação	3.2826	2.5439	2.8738
Recursos Complementares	4.4783	4.0526	4.2427
Tira Dúvidas	4.6304	4.2632	4.4272
Dojos	2.7391	2.3684	2.534
Práticas de Grupo	3.913	2.8246	3.3107
Anti-Plágio	4.5435	3.9649	4.2233
Abordagem do Professor	2.7391	2.9649	2.8641*
Exercícios Propostos	2.8696	3.2105	3.0583*
Foco do Ensino	2.8478	3.1228	3 *
Juiz Online	Sim	Não	Não
Tipo de Prova	Computador	Computador	Computador

Fonte: elaborada pelo autor.

Legenda: * - (1 - Totalmente construtiva) - (5 – Totalmente expositiva)

** - (1-Totalmente contextualizados) - (5 - Totalmente Lógico-matemáticos)

*** - (1 – Focado na Linguagem) - (5 – Focado na parte lógico)

Os resultados da clusterização mostraram alguns perfis de professores, ao analisar esses perfis, percebemos dois opostos, que ao diminuir a quantidade de clusters, ficaram predominantes.

1. Professor mais ativo e participativo. Cobra mais participação dos alunos. Realiza Dojos e atividades em grupo. Tem uma política anti-plágio eficiente. Exercícios mais contextualizados, abordagem mais participativa e maior foco nos aspectos da linguagem.

2. Professor menos participativo. Com foco voltado mais a parte matemática, é, na visão dos alunos, o único agente ativo na construção do conhecimento, repassando conhecimento aos alunos, que passivamente, recebem. Aplica exercícios menos

contextualizados, focados na parte lógico-matemática.

5.2 Análise dos perfis relacionados com as médias das turmas

Tabela 2 – Perfis dos professores com as médias das turmas

Atributos	Perfis		Média
	1	2	
Aulas Práticas	4.55	4.734	4.6336
Atividades Retribuídas	3.715	3.69	3.7036
Cobrança de Participação	2.765	2.878	2.8164
Recursos Complementares	4.2933	4.252	4.2745
Práticas de Grupo	2.725	3.982	3.2964
Anti-Plágio	4.1267	4.612	4.3473
Abordagem do Professor	2.8917	2.684	2.7973
Exercícios Propostos	3.045	3.118	3.0782
Foco do Ensino	3.01	3.196	3.0945
Juiz Online	Sim	Sim	Sim
Tipo de Prova	Computador	Computador	Computador
Mini-Testes	Não	Sim	Não
Média	2.7	2.84	2.7636

Fonte: elaborada pelo autor.

Aqui, relacionamos os perfis dos professores com as médias das turmas nos testes aplicados. Apesar das médias bem próximas dos professores nos testes, algumas variáveis se destacaram nesta análise, principalmente três: Realização de práticas de grupo, Política anti-plágio eficiente e abordagem mais participativa do professor. Outros pontos que podem ser citados é a maior realização de aulas práticas e o uso de minitestes durante o semestre.

5.3 Análise dos perfis relacionados com as taxas de evasão das turmas

Tabela 3 – Perfis dos professores com as taxas de evasão

Atributos	Perfis		Média
	1	2	
Aulas Práticas	4.67	4.57	4.6336
Atividades Retribuídas	3.9843	3.2125	3.7036
Cobrança de Participação	3.1286	2.27	2.8164
Recursos Complementares	4.2214	4.3675	4.2745
Práticas de Grupo	3.3586	3.1875	3.2964
Anti-Plágio	4.5671	3.9625	4.3473
Abordagem do Professor	2.5771	3.1825	2.7973
Exercícios Propostos	2.9657	3.275	3.0782
Foco do Ensino	3.08	3.12	3.0945

Juiz Online	Sim	Não	Sim
Tipo de Prova	Computador	Escrita	Computador
Mini-Testes	Sim	Não	Sim
Taxa de Evasão	23.17	26.027	24.2091

Fonte: elaborada pelo autor.

Neste último tópico, relacionamos com a taxa de evasão das turmas. Essas taxas foram definidas pelo cálculo da porcentagem dos alunos que se encaixam no requisito definido no tópico 4.2 com a quantidade total de alunos inicialmente. Esse valor foi incluído na clusterização, junto com os perfis de cada professor.

Após a clusterização, verificamos alguns pontos que se destacaram no perfil com a menor taxa de evasão: Retribuição pelas atividades realizadas, cobrança pela participação em sala de aula e abordagem mais participativa na aprendizagem dos alunos. Pode ser pontuado também a maior quantidade de aulas práticas, uma política anti-plágio eficiente e a maior contextualização dos problemas apresentados. O uso de computador nas avaliações e o uso de minitestes e juiz online também merecem ser citados.

6 CONCLUSÃO

Após a extração e análise de dados, podemos concluir alguns pontos sobre os dados analisados. Algumas características dos professores se destacam, tanto na obtenção de boas médias, quanto na baixa taxa de evasão. As principais foram:

1. A abordagem mais participativa na aprendizagem dos alunos: Uma abordagem diferente da relação ativo/passivo entre professor e aluno normalmente vista nas salas de aula, onde o professor participa no ciclo de aprendizagem de cada aluno, ajudando e promovendo o desenvolvimento de suas habilidades e conhecimento.

2. Política anti-plágio eficiente: Um método eficiente de detectar plágio, tanto em provas como em atividades. Isso inibe os alunos a usarem respostas de outros colegas ou da internet, motivando-os a fazer suas próprias atividades.

3. Práticas de grupo: A realização de atividades em grupo durante as aulas. Principalmente no primeiro semestre, essa ação é bem importante, porque promove a interação entre novos alunos, que na maioria das vezes, não se conhecem. Também ajuda na construção do conhecimento, ao discutirem as possíveis soluções para um problema, cada aluno aprende um pouco durante isso.

Outros pontos, também merecem serem destacados: Como a influência das bonificações para atividades passadas e a cobrança de participação em sala de aula (com condicionamento do registro de presença nesta participação) no baixo índice de evasão. A utilização de meios tecnológicos de avaliação de atividades e provas também pode influenciar na baixa evasão dos alunos, como o uso de computador em vez da prova escrita e o juiz online nas atividades durante o semestre. Como mencionado por Coutinho et al. (2008), a expectativa dos novos alunos é que seus programas se comportem de maneira semelhante aos sistemas e jogos ao qual estão habituados, então, quanto mais próximo disto, mais motivado o aluno estará para continuar na disciplina, e, conseqüentemente, no curso.

Já na avaliação das médias, a utilização de minitestes durante o semestre se mostra uma boa influência nas médias dos alunos. Deve-se a isso provavelmente a avaliação contínua, que motiva os alunos a estudarem rotineiramente, em vez de alguns dias antes da avaliação “principal”.

Pode-se concluir destas análises, que a participação mais intensa do professor em no decorrer da disciplina, desde a cobrança e participação até o incentivo a atividades, metodologias novas que fogem do estilo ativo/passivo habitual e uso de tecnologias nas atividades e provas têm uma influência importante no sucesso da turma.

7 CONSIDERAÇÕES FINAIS

Neste trabalho, foi utilizado mineração de dados e técnica de clusterização nos dados de todas as turmas introdutórias a programação no campus da Universidade Federal do Ceará em Quixadá. Foram levados em conta dados de testes realizados uniformemente por todas as turmas e de um questionário, aplicado a uma parte dos alunos, sobre a ação e metodologia dos professores no decorrer do semestre.

Obviamente, outros pontos, alheios ao professor e a universidade, como por exemplo, motivação do aluno ao entrar no curso, condições socioeconômicas e se o aluno trabalha ou não, também tem influência no resultado individual, e, conseqüentemente, da turma. Porém, ficou claro o papel importante que o professor, sua participação e sua metodologia tem sobre o sucesso da mesma. Essas variáveis devem ser potencializadas e a participação mais ativa do professor incentivada, buscando amenizar um dos grandes problemas apontados por Lopes (2016), em um estudo no próprio campus: A evasão e o alto índice de reprovações na disciplina de fundamentos de programação.

Como trabalhos futuros, sugerimos a inclusão de algumas questões no questionário: se o aluno trabalha ou não, se o aluno veio de escola profissional e se ele possui conhecimento prévio em programação. Uma análise futura em dados padronizados de diferentes semestres, talvez incluindo as disciplinas seguintes de Orientação a Objetos e Estrutura de Dados (para as turmas que as possuem), a fim de verificar explicitamente as mudanças de metodologia e dos índices analisados, também seria proveitosa.

REFERÊNCIAS

- AZEREDO, P. A. **Uma proposta de Plano Pedagógico para a Matéria de Programação.** Anais do II Curso: Qualidade de Cursos de Graduação da Área de Computação e Informática (WEI 2000). Editora Universitária Champagnat.
- CABENA, P; HADJINIAN, P; STADLER, R; JAAPVERHEES; ZANASI, A. **Discovering Data Mining: From Concept to Implementation.** Prentice Hall, 1998.
- CASPERSEN, M. E., & KOLLING, M. **STREAM: A first programming process.** ACM Transactions on Computing Education (TOCE), 9(1), 4. 2009
- COUTINHO, F. R. S.; ALMEIDA, J.; PRATES, R.. O; CHAIMOWICZ, L.. **Belesminha: Um jogo educacional para apoio ao aprendizado de recursividade.** In: SBGAMES – VII Simpósio Brasileiro de Jogos para Computador e Entretenimento Digital, Belo Horizonte. 2008.
- FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory.; SMYTH, Padhraic. **From Data Mining to Knowledge Discovery in Databases.** AI MAGAZINE, Providence, p.37-54, 1997. Disponível em: <<http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf>>. Acesso em: 20 mar. 2018
- FESSAKIS, G.; GOULI, E; MAVROUDI, E. **Problem solving by 5–6 years old kindergarten children in a computer programming environment: A case study.** Computers & Education, v. 63, p. 87-97, 2013.
- FRANÇA, R; TEDESCO, P. **Caracterizando a pesquisa sobre autoavaliação na aprendizagem de programação para iniciantes.** In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2015. p. 549.
- GOMES, A., AREIAS, C., HENRIQUES, J., and MENDES, A. J. (2008). **Aprendizagem de programação de computadores: dificuldades e ferramentas de suporte.** Revista Portuguesa de Pedagogia, 42(2).
- HAND, D; MANNILA, H; SMYTH, P. **Principles of Data Mining.** MIT Press, 2001.
- HU, M. **Using Game Scenarios for Teaching Novice Programmers.** Institutes of Technology and Polytechnics (ITP): in New Zeland. Journal of Applied Computing & Information Technology. 2008.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. **Data clustering: A review.** ACM computing surveys (CSUR) , v. 31, 1999.
- KALELIOĞLU, F. **A new way of teaching programming skills to K-12 students: Code.org.** Computers in Human Behavior, v. 52, p. 200-210, 2015.
- LOPES, Luclécia Correia. **Análise de dados acadêmicos que indicam evasão de alunos em um curso de graduação tecnológica.** TCC. Universidade Federal do Ceará, Campus Quixadá, Quixadá, CE, 2016.

MACEDO, Crislânio de Souza. **Análise estatística da relação entre evasão e as respostas do questionário para ingressantes da UFC-Quixadá. 2016.** TCC (Graduação em Ciência da Computação) - Universidade Federal do Ceará, Campus de Quixadá, Quixadá, 2016.

PEREIRA, J. C. R., RAPKIEWICZ, C. E., DELGADO, C., XEXEO, J. A. M. **Ensino de Algoritmos e Programação: Uma Experiência no Nível Médio.** XIII Workshop de Educação em Computação. São Leopoldo, RS. 2005.

ROUNTREE, N., ROUNTREE, J., ROBINS, A., and HANNAH, R. (2004). **Interacting factors that predict success and failure in a cs1 course.** ACM SIGCSE Bulletin, 36(4):101–104.

SANTOS, R. P. e COSTA, H. A. X. **TBC-AED e TBC-AED/WEB: Um Desafio no Ensino de Algoritmos, Estruturas de Dados e Programação.** In: IV Workshop de Educação em Computação e Informática do estado de Minas Gerais (WEIMIG). Varginha, MG. 2005.

SILVA, Tércio Jorge da. **Extração de conhecimento nos dados da Universidade Federal do Ceará via mineração de dados: descoberta e análise dos perfis dos alunos.** 2014. 64 f. TCC (graduação em Sistemas de Informação) - Universidade Federal do Ceará, Quixadá, 2014.

SOUTO, Aletéia Vanessa Moreira; DUDUCHI, Marcelo, 2008. **Um processo de avaliação baseado em ferramenta computadorizada para o apoio ao ensino de programação de computadores.** São Paulo, Brasil, pp. 611.

Trovão, C. J. B. M., & Dedecca, C. S. (2014). **Export conditions and formal employment in brazilian manufacturing industries in the period 2003-2008.** Revista de Economia Contemporânea, 18(2), 157-184.

UNIVERSIDADE FEDERAL DO CEARÁ. **Projeto político-pedagógico do curso de Sistemas de informação.** Quixadá. Maio, 2008.

VIER, J., GLUZ, J., and JAQUES, P. A. (2015). **Empregando redes bayesianas para modelar automaticamente o conhecimento dos alunos em lógica de programação.** Revista Brasileira de Informática na Educação, 23(02):45.

VIEIRA, Carlos Eduardo Costa; DE LIMA JUNIOR, José Augusto Teixeira; DE PAULA VIEIRA, Priscila. **Dificuldades no Processo de Aprendizagem de Algoritmos: uma Análise dos Resultados na Disciplina de AL1 do Curso de Sistemas de Informação da FAETERJ–Campus Paracambi.** Cadernos UniFOA, v. 10, n. 27, p. 5-15, 2015.

APÊNDICES

APÊNDICE A – Utilização da ferramenta Weka na clusterização de dados

Figura 4 – Clusterização dos perfis dos professores

The screenshot shows the Weka Explorer application window. The 'Clusterer' tab is active, displaying the SimpleKMeans algorithm configuration and its output. The 'Cluster mode' section is set to 'Use training set'. The 'Clusterer output' section shows the following details:

- Number of iterations: 4
- Within cluster sum of squared errors: 161.10348334604623
- Initial starting points (random):
 - Cluster 0: 5, 5, 3, 5, 5, 1, 4, 4, 3, 3, 2, Sim, Computador
 - Cluster 1: 5, 3, 5, 2, 3, 1, 3, 5, 3, 1, 5, Não, Computador
- Missing values globally replaced with mean/mode
- Final cluster centroids:

Attribute	Full Data (103.0)	Cluster#	
		0 (46.0)	1 (57.0)
aulas_praticas	4.6602	4.7609	4.5789
atividades_retribuidas	3.7282	4	3.5088
cobra_participacao	2.8738	3.2826	2.5439
recursos_complementares	4.2427	4.4783	4.0526
tira_duvidas	4.4272	4.6304	4.2632
dojos	2.534	2.7391	2.3684
praticas_grupo	3.3107	3.913	2.8246
anti_plagio	4.2233	4.5435	3.9649
abordagem_prof	2.8641	2.7391	2.9649
exercicios_propostos	3.0583	2.8696	3.2105
foco_ensino	3	2.8478	3.1228
juiiz_online	Não	Sim	Não
tipo_prova	Computador	Computador	Computador

Time taken to build model (full training data) : 0.01 seconds

=== Model and evaluation on training set ===

The 'Result list' on the left shows several runs of SimpleKMeans, with the most recent one at 17:00:02 selected. The 'Status' bar at the bottom indicates 'OK'.

Figura 5 – Clusterização dos perfis dos professores com as médias

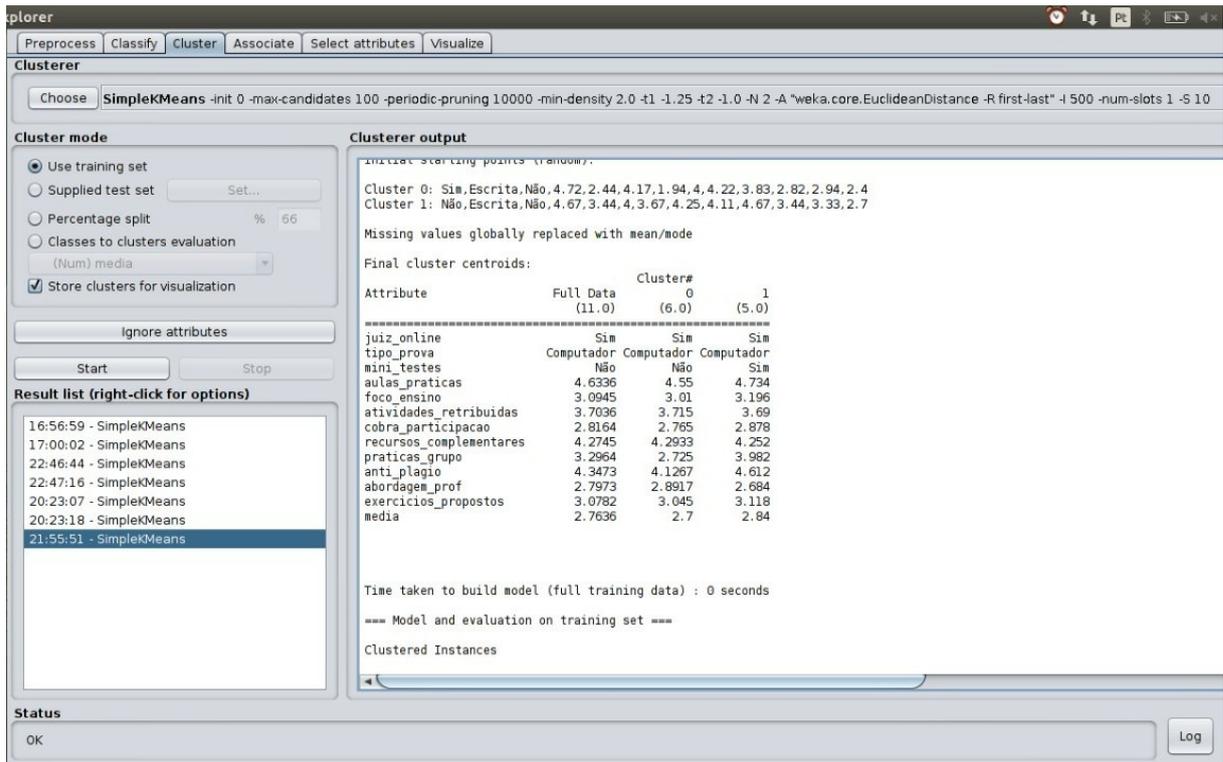


Figura 6 – Clusterização dos perfis dos professores com a taxa de evasão

