



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
CURSO DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

GILDARD CHLODULFE ADEBAYO MYEHOUENOU

**ESTUDO COMPARATIVO DO DESEMPENHO DOS ALUNOS DE CIÊNCIA DA
COMPUTAÇÃO NA UFC QUIXADÁ USANDO MINERAÇÃO DE DADOS**

QUIXADÁ
2019

GILDARD CHLODULFE ADEBAYO MYEHOUENOU

ESTUDO COMPARATIVO DO DESEMPENHO DOS ALUNOS DE CIÊNCIA DA
COMPUTAÇÃO NA UFC QUIXADÁ USANDO MINERAÇÃO DE DADOS

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Orientador: Prof. Dr. Marcos Antônio
de Oliveira

Coorientador: Prof. Tércio Jorge da Silva

QUIXADÁ

2019

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)

M995e Myehouenou, Gildard Chlodulfe Adebayo.

Estudo comparativo do desempenho dos alunos de Ciência da Computação na UFC Quixadá usando Mineração de Dados / Gildard Chlodulfe Adebayo Myehouenou. – 2019.
66 f. : il.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Quixadá, Curso de Sistemas de Informação, Quixadá, 2019.

Orientação: Prof. Dr. Marcos Antônio de Oliveira .

Coorientação: Prof. Tércio Jorge da Silva.

1. Mineração de dados (computação). 2. Cluster (Sistema de computador). 3. Algoritmos. 4. Desempenho Acadêmico. I. Título.

CDD 005

GILDARD CHLODULFE ADEBAYO MYEHOUENOU

ESTUDO COMPARATIVO DO DESEMPENHO DOS ALUNOS DE CIÊNCIA DA
COMPUTAÇÃO NA UFC QUIXADÁ USANDO MINERAÇÃO DE DADOS

Trabalho de Conclusão de Curso apresentado ao
Curso de Graduação em Sistemas de Informação
do Campus Quixadá da Universidade Federal
do Ceará, como requisito parcial à obtenção do
grau de bacharel em Sistemas de Informação.

Aprovada em: ____/____/____

BANCA EXAMINADORA

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

Prof. Tércio Jorge da Silva (Coorientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Regis Pires Magalhães
Universidade Federal do Ceará (UFC)

Profa. Dra. Maria Viviane de Menezes
Universidade Federal do Ceará (UFC)

Aos meus pais que sempre fizeram tudo para que não me faltasse nada, e que mesmo estando do outro lado do mundo, sempre estiveram perto de mim no meu coração.

AGRADECIMENTOS

A minha família que sempre acreditou em mim. A minha companheira Maria Jessica Paulo Alves que sempre me apoiou nos momentos difíceis desta jornada.

Ao Prof. Tércio Jorge da Silva por me orientar e me coorientar durante o trabalho de conclusão de curso, e por todos os conselhos e ensinamentos que foram repassados.

Ao Prof. Dr. Marcos Antônio de Oliveira por aceitar me orientar na segunda etapa da elaboração deste trabalho e por todos os conselhos repassados.

Ao Doutorando em Engenharia Elétrica, Ednardo Moreira Rodrigues, e seu assistente, Alan Batista de Oliveira, aluno de graduação em Engenharia Elétrica, pela adequação do *template* utilizado neste trabalho para que o mesmo ficasse de acordo com as normas da biblioteca da Universidade Federal do Ceará (UFC).

Agradeço a todos os professores por me proporcionarem o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me ajudado a aprender.

“Que do pior você possa extrair o melhor e que no melhor, nunca adormeça. Pois os prazeres de amanhã se vivem nas promessas da véspera.”

(Marek Abi)

RESUMO

Este trabalho tem como foco o estudo comparativo do desempenho dos alunos do curso de Ciência da Computação usando a mineração de dados. A fim de obter um rendimento melhor dos alunos, a partir do ano de 2015, os professores da disciplina de Fundamentos de Programação fizeram reuniões para adotar uma metodologia que consiste em métodos de avaliação, uma base de trabalhos e de atividades. Os métodos de avaliação consistem na aplicação de testes contínuos e avaliações parciais pela ferramenta online *Moodle*. Nas disciplinas de Estrutura de Dados e Programação Orientada a Objetos a solução proposta foi a colaboração entre os professores e os monitores das disciplinas com o objetivo de aplicar listas de exercícios que ajudem os alunos a praticar e entender melhor as aulas. O objetivo principal deste trabalho é comparar o desempenho dos alunos antes e depois do uso da metodologia de avaliação na disciplina de Fundamentos de Programação. Uma comparação entre os resultados nas diferentes disciplinas é feita para avaliar o impacto de uma tal metodologia no desempenho dos alunos ao longo dos anos 2013 a 2017. O método de mineração de dados chamado agrupamento foi escolhido para ser usado nos dados da universidade. O algoritmo *K-means* foi aplicado a um conjunto de dados dos alunos da Universidade Federal do Ceará com o auxílio da ferramenta *Scikit-Learn*, que é uma biblioteca de aprendizado de máquina de código aberto em *Python*. Com os resultados obtidos da mineração de dados, conclusões puderam ser tiradas sobre o efeito das medidas adotadas.

Palavras-chave: Mineração de dados. Agrupamento. *K-means*. Desempenho

ABSTRACT

This paper focuses on the comparative study of the performance of Computer Science students using data mining. In order to achieve better student performance, from 2015 onwards, the teachers of Programming Fundamentals held meetings to adopt a methodology consisting of assessment methods, a work and activity base. The evaluation methods consist of the application of continuous tests and partial evaluations by the online tool textit Moodle. In the Data Structure and Object Oriented Programming course, the solution was a collaboration between teachers and instructor students to apply exercise lists that help students practice and better understand the classes. The main objective of this paper is to compare student performance before and after using the assessment methodology in the Programming Fundamentals discipline. A comparison of results across different courses is made to assess the impact of such a methodology on student performance over the years 2013 to 2017. The data mining method called clustering was chosen to be used on university data. The textit K-means algorithm was applied to a data set of students from the Federal University of Ceará with the aid of the textit Scikit-Learn tool, which is an open source machine learning library in textit Python. With the results obtained from data mining, conclusions could be drawn about the effect of the adopted measures.

Keywords: Data mining. Clustering. K-means. Performance

LISTA DE FIGURAS

Figura 1 – Etapas da Descoberta de Conhecimento	21
Figura 2 – Dados recebidos (Disciplinas e notas)	27
Figura 3 – Dados recebidos (Cidades e situação dos alunos)	27
Figura 4 – Arquivo CSV de Fup 2013.1	27
Figura 5 – Dados transformados	28
Figura 6 – Método do cotovelo	29
Figura 7 – O centroide 0 de Fundamentos de Programação (FUP) 2013.1	31
Figura 8 – Centroide 1 e 2 de FUP 2014.1	32
Figura 9 – Centroide 0 e 1 de FUP 2015.1	33
Figura 10 – Centroide 3 de FUP 2015.1	34
Figura 11 – Centroide 0 e 1 de FUP 2016.1	34
Figura 12 – Centroide 0 e 1 de FUP 2017.1	35
Figura 13 – Grafo de evolução do número de alunos com uma nota acima da média de 2013 a 2017	36
Figura 14 – Centroide 0 de Estrutura de Dados (ED) 2013.2	37
Figura 15 – Centroide 0 de ED 2014.2	38
Figura 16 – Centroide 0 de ED 2015.2	38
Figura 17 – Centroide 0 de ED 2016.2	39
Figura 18 – Centroide 0 de ED 2017.2	39
Figura 19 – Grafo de evolução do número de alunos com uma nota acima da média de 2013 a 2017	41
Figura 20 – Centroide 0 de Programação Orientada a Objetos (POO) 2013.2	42
Figura 21 – Centroide 0 de POO 2014.2	42
Figura 22 – Centroide 0 de POO 2015.2	43
Figura 23 – Centroide 0 de POO 2016.2	43
Figura 24 – Centroide 0 de POO 2017.2	44
Figura 25 – Grafo de evolução da quantidade de alunos com nota acima da média	45
Figura 26 – Tabela descritiva da situação do aluno.	50
Figura 27 – O centroide 1 de FUP 2013.1	50
Figura 28 – Os centroides 2 e 3 de FUP 2013.1	51
Figura 29 – Grupos de alunos de FUP 2013.1	51

Figura 30 – Grupos de alunos de FUP 2014.1	51
Figura 31 – Centroide 0 de FUP 2014.1	52
Figura 32 – Grupos de alunos de FUP 2015.1	52
Figura 33 – Grupos de alunos de FUP 2016.1	52
Figura 34 – Centroide 2 e 3 de FUP 2016.1	53
Figura 35 – Centroide 4 e 5 de FUP 2016.1	53
Figura 36 – Grupos de alunos de FUP 2017.1	53
Figura 37 – Centroide 2 e 3 de FUP 2017.1	54
Figura 38 – Grupos de alunos de ED 2013.2	55
Figura 39 – Centroide 1 e 2 de ED 2013.2	55
Figura 40 – Centroide 3 de ED 2013.2	56
Figura 41 – Grupos de alunos de ED 2014.2	56
Figura 42 – Centroide 1 de ED 2014.2	56
Figura 43 – Centroide 2 e 3 de ED 2014.2	57
Figura 44 – Centroide 4 e 5 de ED 2014.2	57
Figura 45 – Grupos de alunos de ED 2015.2	57
Figura 46 – Centroide 1 de ED 2015.2	58
Figura 47 – Centroide 2 e 3 de ED 2015.2	58
Figura 48 – Grupos de alunos de ED 2016.2	58
Figura 49 – Centroide 1 e 2 de ED 2016.2	59
Figura 50 – Grupos de alunos de ED 2017.2	59
Figura 51 – Centroide 1 de ED 2017.2	59
Figura 52 – Centroide 2 e 3 de ED 2017.2	60
Figura 53 – Centroide 4 e 5 de ED 2017.2	60
Figura 54 – Grupos de alunos de POO 2013.2	61
Figura 55 – Centroide 1 de POO 2013.2	61
Figura 56 – Centroide 2 e 3 de POO 2013.2	62
Figura 57 – Grupos de alunos de POO 2014.2	62
Figura 58 – Centroide 1 de POO 2014.2	62
Figura 59 – Centroide 2 e 3 de POO 2014.2	63
Figura 60 – Grupos de alunos de POO 2015.2	63
Figura 61 – Centroide 1 e 2 de POO 2015.2	63

Figura 62 – Centroide 3 e 4 de POO 2015.2	64
Figura 63 – Grupos de alunos de POO 2016.2	64
Figura 64 – Centroide 1 e 2 de POO 2016.2	64
Figura 65 – Centroide 3 de POO 2016.2	65
Figura 66 – Centroide 4 de POO 2016.2	65
Figura 67 – Grupos de alunos de POO 2017.2	65
Figura 68 – Centroide 1 e 2 de POO 2017.2	66
Figura 69 – Centroide 3 de POO 2017.2	66

LISTA DE ABREVIATURAS E SIGLAS

CSV	Comma Separated Values
ED	Estrutura de Dados
FUP	Fundamentos de Programação
KDD	<i>Knowledge Discovery in Databases</i>
POO	Programação Orientada a Objetos
SGBD	Sistema de Gerenciamento de Banco de Dados
SIGAA	Sistema Integrado de Gestão de Atividades Acadêmicas

SUMÁRIO

1	INTRODUÇÃO	15
2	TRABALHOS RELACIONADOS	18
2.1	Mineração de dados educacionais para analisar o desempenho acadêmico dos alunos.	18
2.2	Extração de conhecimento nos dados da Universidade Federal do Ceará via mineração de dados: descoberta e análise dos perfis dos alunos.	18
2.3	Recomendação no ensino superior usando técnicas de mineração de dados.	19
2.4	Tabela comparativa dos trabalhos relacionados e do estudo realizado. . .	19
3	FUNDAMENTAÇÃO TEÓRICA	21
3.1	Knowledge Discovery in Databases (<i>Knowledge Discovery in Databases</i> (KDD))	21
3.2	Mineração de Dados	22
3.3	Agrupamento	24
3.4	Desempenho acadêmico.	24
4	METODOLOGIA	25
4.1	Escolha do método de mineração de dados	25
4.2	Recebimento e limpeza e transformação dos dados	25
4.3	Definição da análise a ser feita nos dados	28
4.4	Aplicação da técnica de mineração de dados	28
5	RESULTADOS	31
5.1	Comparação do desempenho dos alunos em Fundamentos de Programação	31
5.1.1	<i>Resultados de 2013</i>	31
5.1.2	<i>Resultados de 2014</i>	32
5.1.3	<i>Resultados de 2015</i>	33
5.1.4	<i>Resultados de 2016</i>	33
5.1.5	<i>Resultados de 2017</i>	34
5.2	Comparação do desempenho dos alunos em Estrutura de Dados	36
5.2.1	<i>Resultados de 2013</i>	36
5.2.2	<i>Resultados de 2014</i>	37

5.2.3	<i>Resultados de 2015</i>	38
5.2.4	<i>Resultados de 2016</i>	39
5.2.5	<i>Resultados de 2017</i>	39
5.3	Comparação do desempenho dos alunos em Programação Orientada a Objetos	41
5.3.1	<i>Resultados de 2013</i>	41
5.3.2	<i>Resultados de 2014</i>	42
5.3.3	<i>Resultados de 2015</i>	43
5.3.4	<i>Resultados de 2016</i>	43
5.3.5	<i>Resultados de 2017</i>	44
6	CONCLUSÕES E TRABALHOS FUTUROS	47
	REFERÊNCIAS	49
	APÊNDICE A – FIGURAS RELACIONADAS A FUNDAMENTAÇÃO DE PROGRAMAÇÃO DE 2013 ATÉ 2017	50
	APÊNDICE B – FIGURAS RELACIONADAS A ESTRUTURA DE DA- DOS DE 2013 ATÉ 2017	55
	APÊNDICE C – FIGURAS RELACIONADAS A PROGRAMAÇÃO ORI- ENTADA A OBJETOS DE 2013 ATÉ 2017	61

1 INTRODUÇÃO

O campus da UFC em Quixadá é um dos mais novos campus da instituição. O campus oferece 5 cursos do nível Bacharelado (Sistemas de informação, Engenharia de Software, Engenharia da Computação, Ciência da Computação, Design Digital) e um curso Tecnólogo (Redes de Computadores).

Todo ano o campus recebe uma quantidade considerável de novos alunos, mas a quantidade de formados no final do ano sempre é relativamente inferior. De fato, a experiência mostrou que para cada turma de alunos, poucos se formam no prazo regulamentar. Nesse contexto, é interessante analisar o desempenho acadêmico dos alunos que ingressam todo ano na universidade. Para isso, podem ser usados os dados gerados pela própria universidade. Assim como toda empresa e grande organização, a Universidade Federal do Ceará gera dados que sem um processo de análise adequado não traz muita informação.

A Descoberta de Conhecimento consiste na extração de conhecimento útil de uma grande quantidade de dados. Nesse processo de descoberta, existem várias etapas em que cada uma tem a sua importância primordial. Dentre essas etapas, a mineração de dados é uma das mais importantes e consiste em um conjunto de técnicas para analisar grandes quantidades de dados com o objetivo de encontrar um padrão consistente, resumir os dados, extrair conhecimento ou realizar previsões. A mineração de dados pode ser feita através de várias técnicas como a associação, a classificação, o agrupamento e várias outras como foi detalhado por Camilo e Silva (2009) no trabalho intitulado "Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas".

No contexto deste trabalho, no passo de mineração de dados do processo de Descoberta do Conhecimento, a técnica escolhida é o agrupamento que se fundamenta nas informações encontradas nos dados que descrevem os objetos e seus relacionamentos. De acordo com Pang-Ning Michael Steinbach (2009), o objetivo dessa técnica é formar grupos de objetos que sejam semelhantes entre si em um mesmo grupo e diferentes dos objetos de outros grupos. Quanto maior a semelhança dentro de um grupo, e maior a diferença entre grupos, melhor ou mais distinto será o agrupamento.

A fim de obter um rendimento melhor dos alunos, a partir do ano de 2015, os professores da disciplina de Fundamentos de Programação fizeram reuniões para adotar uma metodologia que consiste em um método de avaliação, uma base de trabalhos e de atividades centralizada. Os métodos de avaliação consistem em aplicação de testes contínuos e avaliações parciais pela ferramenta online *Moodle*. Os testes contínuos consistem na avaliação dos alunos

sobre cada conteúdo da disciplina que foram ministrados pelos professores. Esses testes servem para avaliar o grau de entendimento do conteúdo repassado e manter os alunos preparados para as avaliações. Desse modo, os alunos das diferentes turmas e diferentes cursos teriam a mesma base, independentemente do professor que esteja ministrando a disciplina.

A aplicação do algoritmo *K-means* nos dados da universidade possibilitou encontrar vários grupos de alunos com características próprias. Por exemplo, encontrou-se um grupo de alunos com as seguintes características: boas notas, boas frequências, sem reprovação, ira individual e geral bons e que moram em Quixadá. Este é um grupo formado dos melhores alunos de uma determinada disciplina. Os dados de maneira geral são compostos por: matrícula do aluno, a frequência do aluno na disciplina, a nota na disciplina, a situação do aluno (trancado, cancelado, aproveitamento interno, aproveitamento externo, matriculado, aprovado por média, aprovado, reprovado, reprovado por falta) entre outros atributos. Como falado anteriormente, as disciplinas que foram usadas são: Fundamentos de Programação, Estrutura de Dados e Programação Orientada a Objetos. O motivo dessa escolha é pelo fato que essas disciplinas são ligadas por uma relação de dependência. Por exemplo para poder cursar a disciplina de Programação Orientada a Objetos, é necessário ter cursado a disciplina de Fundamentos de Programação e ter sido aprovado.

Neste trabalho, foram analisados os diferentes grupos de alunos por semestre e por disciplina resultantes da aplicação do método de mineração dados. A análise feita foi uma comparação dos diferentes grupos encontrados, mas principalmente uma comparação antes (2013 a 2014) e depois (2015 a 2017) do método de avaliação adotado pelos professores na disciplina de Fundamentos de Programação. Outra comparação foi feita também entre a disciplina de Fundamentos de Programação, e as duas outras disciplinas que são Programação Orientada a Objetos e Estrutura de dados em que não foi adotado o mesmo método de avaliação para observar o comportamento dos alunos dessas disciplinas. Por exemplo, em Fundamentos de Programação, uma comparação entre os diferentes grupos dos semestre de 2014.1 e 2015.1 é feita para analisar o desempenho dos alunos de um ano para outro. Os passos para a elaboração deste trabalho foram: primeiro, identificar um método de mineração de dados a ser aplicado aos dados da universidade, em seguida fazer uma comparação entre os diferentes semestres de cada disciplina dos resultados obtidos do método de mineração de dados e por fim apresentar uma conclusão dessa comparação para auxiliar os gestores.

No Capítulo 2 são apresentados os trabalhos relacionados. No Capítulo 3, são

apresentados os diferentes conceitos relevantes para o bom entendimento desse trabalho, no Capítulo 4 são apresentados os procedimentos e metodologias, no Capítulo 5 são apresentados os resultados e por fim no Capítulo 6 a conclusão e os trabalhos futuros.

2 TRABALHOS RELACIONADOS

A seguir são apresentados os trabalhos que foram usados como base para a criação deste estudo. Uma breve descrição de cada trabalho é realizada com o propósito de destacar as semelhanças e diferenças em relação a este trabalho.

2.1 Mineração de dados educacionais para analisar o desempenho acadêmico dos alunos.

No trabalho de Baradwaj e Pal (2012) o objetivo principal é por meio da aplicação de técnicas de mineração de dados, classificar os alunos de acordo com o desempenho acadêmico, e identificar os alunos que precisam de uma atenção particular para que medidas fossem tomadas para os semestres posteriores. Foram usados dados da Universidade Veer Bahadur Singh Purvanchal da cidade de Jaunpur na Índia, do curso de *Master of Computer Applications* de 2007 a 2010. Inicialmente, tinham 50 tabelas. Depois das fases de limpeza e pré-processamento, os dados foram reunidos em uma tabela só. Oito variáveis foram consideradas: a nota no semestre anterior, a performance no seminário, trabalho entregue (sim ou não), trabalho de laboratório (concluiu ou não), presença, nota no final do semestre, nota de proficiência (participou ou não participou) e a nota de classificação de nível. O trabalho de Baradwaj e Pal (2012) assemelha-se a este trabalho na análise dos dados de alunos do ensino superior, mas se diferencia pelo fato de ser alunos de pós-graduação. Outra diferença é a técnica de árvore de decisão que foi utilizada por Baradwaj e Pal (2012). Essas decisões geram regras de classificação dos conjuntos de dados.

2.2 Extração de conhecimento nos dados da Universidade Federal do Ceará via mineração de dados: descoberta e análise dos perfis dos alunos.

O trabalho de Silva (2014) tinha como objetivo encontrar diferentes perfis de alunos, aplicando técnicas de mineração de dados. Esse trabalho analisou também o impacto de algumas novas medidas como a implantação de novas regras de frequência, mas também o impacto da adesão do SISU nos perfis dos alunos ingressando na Universidade. Na disciplina de Fundamentos de Programação, em relação às notas e a quantidade de alunos, os perfis de alunos se mantiveram constante durante os anos antes da adesão do SISU. Em relação às médias de faltas, nos anos antes do SISU, o maior percentual de falta era 55%, depois do SISU, foram observados perfis com média de falta entre 65% e 75%. Os dados que serviram de base são da Universidade Federal do Ceará (UFC) Campus do Pici entre os anos 2006 até 2012. Os dados são

constituídos dos atributos seguintes: o identificador do aluno, a quantidade de faltas, o número de matrículas, o número de insucessos entre outros. O que há de semelhante com este trabalho é o fato que os alunos são de graduação, são da mesma instituição e que as disciplinas escolhidas são aquelas com maior índice de reprovação. Outro ponto de semelhança é método de mineração de dados usados que é o agrupamento com o algoritmo K-means. O que pode destacar-se é que nesse trabalho os dados são referentes ao curso de Computação enquanto neste trabalho serão analisados os dados do curso de Ciência da Computação. Um ponto de diferença entre os trabalhos é a quantidade de disciplinas que foram selecionadas.

2.3 Recomendação no ensino superior usando técnicas de mineração de dados.

No estudo de Vialardi *et al.* (2009), o objetivo principal é encontrar padrões que serão usados para dar recomendações positivas ou negativas a um aluno que se inscreveu no processo seletivo do curso de Engenharia de Sistemas da Universidad de Lima levando em conta o seu perfil. Na maioria dos casos, os alunos escolhem os cursos sem saber se o curso convém às suas esperanças. Os dados usados nesse trabalho são os dados dos alunos do curso de Engenharia de Sistemas do ano 2002 até o ano 2008. Afim de fazer essa pesquisa, os dados foram organizados em uma tabela. Os atributos usados são: a lista de cursos em que o aluno se inscreveu, nome do curso, a média geral do aluno e sua classificação (A, B, C) entre outros. Foram feitos vários ensaios nos dados. Os resultados obtidos permitiram fazer recomendações positivas ou negativas com uma precisão de 77,3% aos alunos que pretendem se inscrever no curso de Engenharia de Sistemas. Essa pesquisa tem semelhança com o trabalho que está sendo desenvolvido pelo fato dele usar também as técnicas de mineração de dados, também por ser desenvolvido por fins acadêmicos. Pode-se também ressaltar alguns pontos de diferença como o fato dessa pesquisa ser feita para o curso de Engenharia de Sistemas e a técnica usada foi a árvore de decisão que foi gerado pelo uso do algoritmo C4.5.

2.4 Tabela comparativa dos trabalhos relacionados e do estudo realizado.

Nesta tabela, está resumida as diferenças e semelhanças dos diferentes trabalhos que serviram de apoio para o desenvolvimento desta estudo.

Tabela 1 – Comparativo dos trabalhos relacionados

	Baradwaj e Pal (2012)	Silva (2014)	Vialardi <i>et al.</i> (2009)	Trabalho proposto
Técnica usada	Árvore de decisão	Agrupamento	Árvore de decisão	Agrupamento
Dataset	Pós-graduação	Graduação	Ingressante em graduação	Graduação
Número de cursos	Um	Um	Todos os cursos	Um
Número de disciplinas	Uma	Quatro	Todas as disciplinas do curso	Três
Atributos	Nota do semestre anterior, performance no seminário, trabalho entregue, trabalho de laboratório, presença, nota final no semestre, nota de proficiência, nota de classificação de nível	Identificador do aluno, quantidade de faltas, número de matrículas, número de insucessos, média notas, reprovação por falta, trancamento	Lista de cursos desejados, média geral do aluno, notas individuais	Identificador do aluno, frequência, nota, insucessos, reprovação por falta, trancamento, horas integralizadas, ira individual, ira geral, cidade, situação, quantidade de matrículas

Fonte: Elaborada pelo autor.

3 FUNDAMENTAÇÃO TEÓRICA

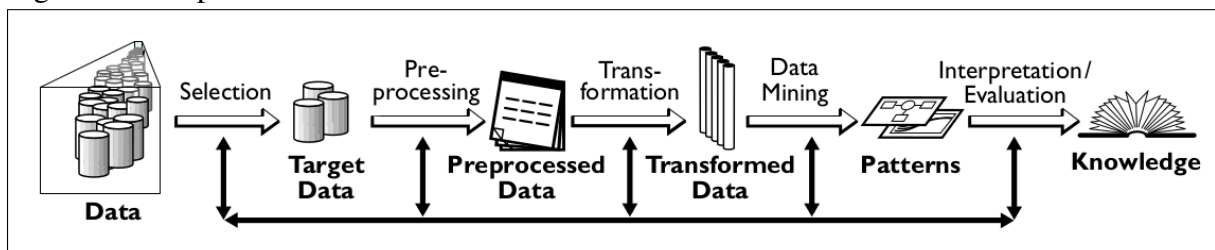
Neste capítulo, são apresentadas as definições e conceitos importantes para o entendimento deste trabalho. Na Seção 3.1 é apresentada a definição de KDD (*Knowledge Discovery in Databases* ou Descoberta de Conhecimento nas Bases de Dados), e as diferentes fases do processo. Em seguida, na Seção 3.2 são apresentados conceitos sobre a Mineração de Dados e a demonstração das etapas importantes durante um processo de Descoberta de Conhecimento. Logo depois, na Seção 3.3 é apresentado o conceito de método de agrupamento. A Seção 3.4 trata-se da definição da noção de desempenho acadêmico de aluno.

3.1 Knowledge Discovery in Databases (KDD)

A Descoberta de Conhecimento em Bases de Dados (*Knowledge Discovery in Databases*), segundo a definição de Fayyad *et al.* (1996b) é o processo não trivial de identificação nos dados de padrões válidos, novos, potencialmente úteis e sobretudo compreensíveis. Podemos considerar também uma outra definição de Fayyad *et al.* (1996b) que considera a descoberta de conhecimento em bases de dados o processo de usar um banco de dados ao qual é aplicado qualquer tipo de operação que seja uma seleção, um pré-processamento, uma subamostragem ou qualquer transformação com o objetivo de aplicar métodos (algoritmos) de mineração de dados para encontrar padrões e avaliar aqueles padrões considerados conhecimento.

As etapas do KDD são definidas no trabalho de Silva (2004) como um processo de descoberta de conhecimento em banco de dados é feito de forma iterativa, cognitiva e exploratória, nos quais estão envolvidos vários passos e as decisões são feitas pelo analista do domínio ou analista de dados. A Figura 1 demonstra as etapas que são realizadas durante um processo de descoberta de conhecimento em bases de dados.

Figura 1 – Etapas da Descoberta de Conhecimento



Fonte: (FAYYAD *et al.*, 1996a)

Resumidamente, como demonstrado na figura 1, pode-se descrever os passos mais

importantes de um processo de descoberta de conhecimento em um banco de dados da seguinte maneira:

1. Definição de qual o conhecimento se quer descobrir se entendendo o domínio, e o tipo de decisão que o conhecimento proporcionará em termos de melhora;
2. Seleção: É elegido o conjunto de dados com o qual se quer trabalhar, desses dados são tirados aqueles mais relevantes para a extração do conhecimento;
3. Limpeza de dados e pré-processamento: nesta fase são retirados dados fora do padrão (*outliers*), dados inconsistentes (dados incorretos ou redundantes), e é escolhida uma abordagem para tratar dados ausentes;
4. Transformação ou Formatação: consiste em redução e projeção dos dados para tornar mais fácil a análise dos mesmos;
5. Mineração de dados: fase na qual se escolhe e se aplica a técnica e o algoritmo a ser utilizado nos dados para a obtenção do conhecimento;
6. Interpretação/Avaliação: os resultados obtidos com a mineração dos dados são avaliados por especialistas e serão julgados úteis ou não. O conhecimento útil deve ser interpretado de maneira que os usuários possam entender.

3.2 Mineração de Dados

Segundo Pang-Ning Michael Steinbach (2009), a mineração de dados é o processo de descoberta automático de informações úteis em grandes depósitos de dados.

De uma perspectiva estatística, a mineração de dados pode ser definida como: “a análise de grandes conjuntos de dados a fim de encontrar relacionamentos inesperados e de resumir os dados de uma forma que eles sejam tanto úteis quanto compreensíveis ao dono dos dados” (HAND *et al.*, 2001).

Por outro lado, a definição a partir da perspectiva de banco de dados propõe que a Mineração de Dados é um campo interdisciplinar que junta técnicas de máquinas de conhecimentos, reconhecimento de padrões, estatísticas, banco de dados e visualização, para conseguir extrair informações de grandes bases de dados (CABENA *et al.*, 1997).

A partir da perspectiva da área de aprendizado de máquina a Mineração de Dados é um passo no processo de Descoberta de Conhecimento que consiste na realização da análise dos dados e na aplicação de algoritmos de descoberta que, sob certas limitações computacionais, produzem um conjunto de padrões de certos dados (FAYYAD *et al.*, 1996a).

As tarefas de mineração de dados são geralmente divididas em duas categorias: tarefas de previsão e tarefas descritivas.

Segundo Pang-Ning Michael Steinbach (2009), o objetivo das tarefas de previsão é prever o valor de um determinado atributo baseado nos valores de outros atributos. O atributo a ser previsto é comumente conhecido com a variável dependente ou alvo, enquanto que os atributos usados para fazer a previsão são conhecidas como as variáveis independentes ou explicativas.

De acordo com Pang-Ning Michael Steinbach (2009), há dois tipos de tarefas: a classificação, a qual é usada para variáveis alvos discretas, e a regressão, que é usada para variáveis alvo contínuas. Por exemplo, prever se um usuário Web fará uma compra em uma livraria online é uma tarefa de classificação, porque a variável alvo é de valor binário. Por outro lado, prever o preço de uma ação é uma tarefa de regressão, porque o preço é um atributo de valor contínuo.

O objetivo de ambas as tarefas é aprender um modelo que minimize o erro entre os valores previsto e real da variável alvo. A modelagem de previsão pode ser usada para identificar clientes que responderão a uma campanha de vendas, prever perturbações no ecossistema da Terra ou julgar se um paciente possui uma determinada doença baseado nos resultados de exames médicos.

As tarefas descritivas, segundo Pang-Ning Michael Steinbach (2009), tem como objetivo derivar padrões (correlações, tendências, grupos, trajetórias e anomalias) que resumem os relacionamentos subjacentes nos dados. As tarefas descritivas da mineração de dados são muitas vezes exploratórias em sua natureza e frequentemente requerem técnicas de pós-processamento para validar e explicar os resultados. Dentre as tarefas descritivas, temos a análise de associação e o agrupamento.

A análise de associação é usada para descobrir padrões que descrevam características altamente associadas dentro dos dados. Os padrões descobertos são normalmente representados na forma de regras de implicação ou subconjuntos de características. Devido ao tamanho exponencial do seu espaço de busca, o objetivo da análise de associação é extrair os padrões mais interessantes de uma forma eficiente. Aplicações úteis de análise de associação incluem a descoberta de genes que apresentam funcionalidade associada, a identificação de páginas Web que sejam acessadas juntas ou a compreensão dos relacionamentos entre diferentes elementos do sistema climático da Terra (PANG-NING MICHAEL STEINBACH, 2009).

A técnica de mineração de dados que foi aplicada ao dados dos alunos da Universidade Federal do Ceará (UFC) Campus de Quixadá é uma técnica descritiva, o agrupamento.

3.3 Agrupamento

Segundo Pang-Ning Michael Steinbach (2009), o agrupamento junta objetos baseando-se apenas em informações encontradas nos dados que descrevem os objetos e seus relacionamentos. O objetivo é que os objetos dentro de um grupo sejam semelhantes entre si e diferentes dos objetos de outros grupos. Quanto maior a semelhança (ou homogeneidade) dentro de um grupo, maior a diferença entre grupos e melhor ou mais distinto será o agrupamento. Neste trabalho, o algoritmo de agrupamento usado é o *K-means*.

K-means é um algoritmo que define um protótipo em termos de um centroide, que é geralmente a média de um grupo de pontos, e é geralmente aplicada a objetos em um espaço n -dimensional contínuo. A técnica de agrupamento *K-means* é simples, primeiro são escolhidos K centroides iniciais, onde K é um parâmetro especificado pelo usuário, a saber, o número de grupos desejado. Cada ponto é atribuído a um centroide mais próximo, e cada coleção de pontos atribuídos a um centroide é um grupo. O centroide de cada grupo é então atualizado baseado nos pontos atribuídos ao grupo. São repetidos os passos de atribuição e atualização até que nenhum ponto mude de grupo ou, equivalente até que os centroides permaneçam os mesmos (PANG-NING MICHAEL STEINBACH, 2009).

3.4 Desempenho acadêmico.

Segundo o trabalho de Mondardo e Pedon (2005), o desempenho acadêmico refere-se ao grau final relativo a todas as disciplinas cursadas pelos alunos ao longo do semestre ou do curso. Nesse trabalho, o desempenho acadêmico equivale às medidas das notas cursadas pelo estudante durante o semestre (para alunos ainda matriculados) e durante o curso (para alunos formados).

4 METODOLOGIA

Este capítulo é organizado da maneira seguinte: na seção 4.1 é apresentado o motivo da escolha do método de mineração de dados, em seguida na seção 4.2 são apresentadas as etapas de recebimento, limpeza e transformação do dados, na seção 4.3 é definida a análise a ser feita e por fim na seção 4.4 são mostrados os parâmetros escolhidos para aplicar o algoritmo de agrupamento.

4.1 Escolha do método de mineração de dados

O método de mineração de dados escolhido para ser aplicado aos dados da universidade é o agrupamento. Esse método foi escolhido porque ele oferece a possibilidade de procurar objetos semelhantes e colocá-los dentro de um mesmo grupo. Ele permite de ver como os elementos se relacionam entre si dentro de um grupo. No caso deste estudo, o agrupamento foi usado para identificar os grupos de alunos pela notas acima da média, pela cidade, pelas quantidade de vezes que se matricularam na disciplina e também para identificar os alunos que conseguiram integralizar o máximo número de horas. Com esse tipo de agrupamento, os resultados da aplicação do algoritmo podem ser úteis para outras análises sobre o mesmo conjunto de dados como identificar dentro dos alunos com um bom desempenho nas diferentes disciplinas que foram usadas para este estudo, aqueles que conseguiram manter um bom desempenho ao longo do curso.

4.2 Recebimento e limpeza e transformação dos dados

Depois de algumas reuniões que foram feitas com o coorientador deste trabalho e o coordenador do curso de Ciência da Computação foi decidido usar o conjunto de dados dos alunos de Ciência da Computação do ano 2013 a 2017. Para reduzir o risco de identificação os alunos presentes na base dados, o coordenador extraiu os dados do Sistema Integrado de Gestão de Atividades Acadêmicas (SIGAA) e substituiu as matrículas dos alunos por números sequencias. Esses dados foram salvos em uma planilha Excel como pode ser observado nas Figuras 2 e 3.

Para poder obter os dados prontos à aplicação do algoritmo de mineração, várias operações foram feitas neles. Depois do armazenamento no Sistema de Gerenciamento de Banco de Dados (SGBD), foram realizadas consultas no banco de dados para gerar arquivos Comma Separated Values (CSV), que são arquivos cujo conteúdo é um conjunto de informações separado

por vírgulas como mostrado na Figura 4. Esses arquivos foram gerados para as disciplinas seguintes: Fundamentos de Programação (FUP), Programação Orientada a Objetos (POO) e Estrutura de Dados (ED). Os arquivos gerados são referentes aos semestres onde a maioria dos alunos estão cursando as disciplinas pela primeira vez, ou seja, para FUP, os semestres escolhidos são ímpares e para POO e ED, são os semestres pares.

Foram implementadas funções para substituir os dados de carácter alfabético por números. Um exemplo de transformação dos dados é o caso do cálculo do número de insucessos, de matrícula e de trancamento. Para isso uma função foi implementada para verificar se um aluno tem se matriculado nos semestres anteriores e ainda está matriculado no semestre em questão, se for o caso, o contador é acrescentado. A maioria dos valores dos atributos ficaram no intervalo de 0 (zero) a 10 (dez). Para que todos ficassem nesse intervalo, os atributos como a frequência e as horas integralizadas foram normalizados para ficarem na mesma escala. Na Figura 5 podem ser observados os dados resultantes da transformação do semestre 2013.1 na disciplina de FUP.

Os atributos considerados os mais importantes, foram selecionados e foram armazenados no Sistema de Gerenciamento de Banco de Dados (SGBD) Postgres 9.4.25 em uma tabela. Cada disciplina em cada semestre representa uma tabela única como mostra a Figura 5. Os atributos escolhidos são os seguintes:

1. *Período*: semestre em que foi cursada a disciplina;
2. *Situação*: é a situação do aluno (ex: Aprovado, Reprovado, Reprovado por falta, Trancado, Cancelado, Aprovado por média)
3. *Nota*: a média que o aluno obteve na disciplina;
4. *Cidade*: a cidade onde mora o aluno, as cidades foram substituídas pelos números de 1 e 2, onde 1 corresponde aos alunos que moram em Quixadá e o 2 corresponde aos alunos que não moram em Quixadá;
5. *Trancamentos*: é a quantidade de vezes que o aluno já trancou a disciplina;
6. *Rep faltas*: é o número de faltas que o aluno já teve na disciplina;
7. *Insucessos*: é considerado como insucesso a reprovação por falta, por nota ou trancamento. Esse atributo é uma contagem dessas diferentes situações;
8. *Qtd vezes matriculas*: é a quantidade de vezes que o aluno já se matriculou nessa disciplina;
9. *Ira individual*: é o ira individual;
10. *Ira geral*: é o ira geral;

11. *Horas integralizadas*: é a quantidade de horas que o aluno integralizou durante o curso.

Essa etapa do trabalho corresponde às 3 etapas do processo de Descoberta do Conhecimento: a seleção, a limpeza e pré-processamento dos dados e a transformação ou formatação. Depois dessas etapas, os dados obtidos estão prontos para que o algoritmo de agrupamento seja aplicado neles.

Figura 2 – Dados recebidos (Disciplinas e notas)

	A	B	C	D	E	F
1	Período Letivo	Componente Curricular	Frequência	Nota	Situação	Matrícula
322	2014,2	FUNDAMENTOS DE BANCO DE DADOS	96,87	8,9	APROVADO MÉDIA	10
323	2014,2	ANALISE E PROJETO DE SISTEMAS	96,87	8,9	APROVADO MÉDIA	10
324	2014,2	LINGUAGENS DE PROGRAMAÇÃO	96,87	8,7	APROVADO MÉDIA	10
325	2014,2	TOPICOS ESPECIAIS IV	100	8,8	APROVADO MÉDIA	10
326	2014,2	ÁLGEBRA LINEAR	96,87	8,5	APROVADO MÉDIA	10
327	2015,1	DESENVOLVIMENTO DE SOFTWARE PARA WEB	100	10	APROVADO MÉDIA	10
328	2015,1	REDES DE COMPUTADORES	96,87	9,6	APROVADO MÉDIA	10
329	2015,1	COMPILADORES	98,43	9,3	APROVADO MÉDIA	10
330	2015,1	COMPUTAÇÃO GRÁFICA	90,62	9,6	APROVADO MÉDIA	10
331	2015,1	MATEMÁTICA COMPUTACIONAL	100	7,8	APROVADO MÉDIA	10
332	2015,2	ENGENHARIA DE SOFTWARE	100	10	APROVADO MÉDIA	10

Fonte: Elaborada pelo autor

Figura 3 – Dados recebidos (Cidades e situação dos alunos)

	A	B	C	D	E	F	G	H	I	J
1	Matrícula	Cidade/Estado Nascimento	Bairro	Cidade Atual	Estado atual	Período Inicial	IRA individual	IRA Geral	Status	Tipo de Saída
2	242	FORTALEZA/CE	ZONA RURAL	FORTALEZA	CE	2013.1	0,4472	2168	CANCELADO	CANCELAMENTO
3	110	JUCÁS/CE	CENTRO	JUCÁS	CE	2013.1	6,0053	6963	CANCELADO	CANCELAMENTO
4	133	QUIXADÁ/CE	Alto	ACARI	RN	2013.1	0,2913	2991	CANCELADO	ABANDONO
5	244	IGUATU/CE	--	--	--	2013.1	--	--	CANCELADO	DESISTÊNCIA
6	100	QUIXADÁ/CE	--	--	--	2013.1	0	2724	CANCELADO	ABANDONO
7	45	FORTALEZA/CE	Vila União	FORTALEZA	CE	2013.1	1,8128	2837	CANCELADO	ABANDONO
8	174	BATURITÉ/CE	--	--	--	2013.1	--	--	CANCELADO	DESISTÊNCIA
9	221	IBICUITINGA/CE	Alto São Francisco	QUIXADÁ	CE	2013.1	3372	3686	CANCELADO	ABANDONO
10	22	BANABUIÚ/CE	--	--	--	2013.1	0	3176	CANCELADO	DESISTÊNCIA
11	16	FORTALEZA/CE	Parque Santa Rosa	FORTALEZA	CE	2013.1	1,8902	2902	CANCELADO	MUDANÇA DE CURSO
12	188	FORTALEZA/CE	Campo Novo	QUIXADÁ	CE	2013.1	0,8158	3379	CANCELADO	ABANDONO
13	167	FORTALEZA/CE	CENTRO	QUIXADÁ	CE	2013.1	1,0103	3522	CANCELADO	ABANDONO
14	92	SENADOR POMPEU/CE	--	--	--	2013.1	0	3176	CANCELADO	ABANDONO

Fonte: Elaborada pelo autor

Figura 4 – Arquivo CSV de Fup 2013.1

	A	B	C	D	E	F	G	H	I	J
1	matricula,situacao,nota,periodo,frequencia,cidade,trancamentos,rep_faltas,insucessos,qtd_vezes_matriculas,ira_individual,ira_geral,horas_integralizadas									
2	228,3,5,7,2016.1,0,2,0,0,0,1,0,3726,0,78,864									
3	218,9,8,2,2016.1,100,1,0,0,0,1,6,9771,6,84,1248									
4	64,6,0,5,2016.1,87,5,1,0,0,1,1,0,1364,1,0									
5	62,9,9,2,2016.1,100,1,0,0,0,1,7,01,6,87,1248									
6	143,6,0,2016.1,100,2,0,0,2,2,1,1124,1,341,256									
7	213,7,0,7,2016.1,72,91,2,0,2,3,3,1,3441,1,853,160									
8	24,6,0,7,2016.1,87,5,2,0,0,1,1,2,1671,2,597,160									
9	29,6,0,6,2016.1,87,5,2,0,2,3,3,1,685,2,298,224									
10	180,9,9,2016.1,100,1,0,0,0,1,4,7962,4,839,864									
11	207,9,7,7,2016.1,100,1,0,0,0,1,5,6533,5,625,992									
12	144,9,9,7,2016.1,100,1,0,0,0,1,5,6521,5,624,1056									

Fonte: Elaborada pelo autor

Figura 5 – Dados transformados

situacao	nota	frequencia	cidade	trancamentos	rep_faltas	insucessos	qtd_vezes_matriculas	ira_individual	ira_geral	horas_integralizadas	
0	8	5.2	10.000	2	0	0	0	1	2.9526	4.870	0.8
1	2	0.0	0.000	1	0	0	0	1	0.0000	0.000	0.0
2	6	0.0	7.916	1	0	0	1	1	0.4472	2.168	0.3
3	8	6.0	9.791	1	0	0	0	1	6.0053	6.963	3.3
4	6	0.8	7.500	1	0	0	1	1	0.2913	2.991	0.0

Fonte: Elaborada pelo autor

4.3 Definição da análise a ser feita nos dados

Nesta fase do trabalho foram feitas algumas discussões sobre a análise que seria feita sobre os dados. Para isso, foi necessário marcar uma reunião entre o autor e o professor responsável da organização das reuniões entre professores a fim de obter informações específicas sobre o período em que começou a ser abordada a questão da adoção da nova metodologia de avaliação que consiste na aplicação de testes contínuos, atividades, trabalhos e provas com o auxílio da ferramenta online Moodle do campus de Quixadá na disciplina de Fundamentos de Programação. Com o objetivo de decidir quais questões deveriam ser respondidas sobre os dados, foram levantados os seguintes questionamentos:

1. A adoção do método de avaliação baseado em testes contínuos, realização de atividades, trabalhos e provas pela ferramenta Moodle a partir do ano de 2015 teve um impacto positivo sobre o desempenho dos alunos nas disciplinas de Fundamentos de Programação?
2. A não utilização do mesmo método aplicado pelos professores de Fundamentos de Programação nas disciplinas de Programação Orientada a Objetos e Estrutura de Dados teve um impacto sobre o desempenho dos alunos?

4.4 Aplicação da técnica de mineração de dados

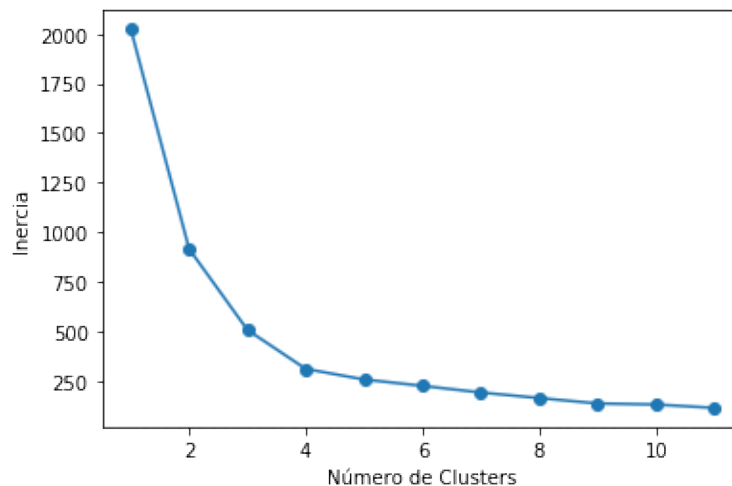
Nesta seção foi aplicada a técnica de mineração de dados aos dados da universidade com o auxílio da ferramenta *open source* Scikit-learn. Como foi mencionado anteriormente, os dados foram divididos em diferentes arquivos CSV para cada semestre e para cada uma das disciplinas escolhidas, no caso Fundamentos de Programação, Programação Orientada a Objetos e Estrutura de Dados. O algoritmo *K-means* foi aplicado aos dados com os parâmetros seguintes:

- *n_clusters*: O número de grupos a serem formados e o número de centroides a serem gerados;

- `init`: k-means ++ (Método para inicialização, o padrão é "k-means ++" onde os centroides são escolhidos de uma maneira inteligente para acelerar a convergência; outra opção é "random" onde os centroides são escolhidos de maneira aleatória);
- `n_init`: 10 (Número de vezes que o algoritmo será executado com diferentes centroides);
- `max_iter`: 300 (Número máximo de iterações do algoritmo para uma única execução);
- `tol`: 0.0001 (Tolerância relativa em relação à inércia para declarar convergência);
- `algorithm`: full (O algoritmo clássico do k-means);

Para auxiliar na escolha desse número, usamos o método do cotovelo que segundo Kodinariya e Makwana (2013) que consiste em executar o *K-means* várias vezes fazendo uma iteração em um intervalo em que cada número representa a quantidade de grupos. A partir de cada iteração, podemos usar uma propriedade do resultado do algoritmo que é a inércia. A inércia é uma medida que informa sobre o quanto os grupos gerados são internamente coerentes entre si. Menor a inércia é, melhor é a qualidade dos grupos. Na Figura 6 pode-se ver o método do cotovelo aplicado aos dados dos alunos de Fundamentos de Programação de 2013.1. Nesse caso, foi escolhido o número de 4 grupos, pois a partir da quarta iteração a inércia decresce aproximadamente de maneira constante.

Figura 6 – Método do cotovelo



Fonte: Elaborada pelo autor

Com a aplicação do método de mineração de dados, encontrou-se vários grupos de alunos com características distintas. Como mostrado na Figura 7, podemos observar os centroides de alunos na disciplina de Fundamentos de Programação no semestre de 2013.1. É importante lembrar que com a finalidade de facilitar a leitura dos resultados da mineração de dados, os atributos frequência e horas integralizadas que foram normalizados anteriormente,

foram sujeitos ao mesmo processo inverso, para poder mostrar os resultados em uma escala mais inteligível.

Cada centroide é a representação da média dos atributos de cada grupo. Cada grupo corresponde a um perfil de aluno. Como mostrado na Figura 7, podemos observar os grupos de alunos na disciplina de Fundamentos de Programação no primeiro semestre do ano 2013.

5 RESULTADOS

Neste capítulo são apresentados os resultados da análise sobre os dados da Universidade Federal do Ceará Campus de Quixadá com o auxílio da técnica de mineração de dados chamado agrupamento. Os grupos de alunos que foram considerados para fazer a comparação foram os grupos dos alunos que tiveram uma nota igual ou acima da média ou seja uma nota maior ou igual a 7. Na Seção 5.1 são apresentados os resultados referentes ao desempenho dos alunos na disciplina de Fundamentos de Programação. Em seguida na Seção 5.2 pode-se observar os resultados dos alunos de Estrutura de Dados. Por fim, na Seção 5.3 são apresentados os resultados dos alunos da disciplina de Programação Orientada a Objetos.

5.1 Comparação do desempenho dos alunos em Fundamentos de Programação

Nesta etapa do trabalho, são analisados os resultados da mineração de dados aplicada aos dados dos alunos de Fundamentos de Programação do ano de 2013 a 2017. O agrupamento com o auxílio do algoritmo *K-means* foi aplicado somente às turmas do primeiro semestre ou seja, a turma onde a maioria são alunos ingressantes no mesmo ano. Dos resultados da mineração de dados, foram identificados três grupos de alunos que se repetem em todos os anos sem muita alteração.

5.1.1 Resultados de 2013

Figura 7 – O centroide 0 de FUP 2013.1

Nota	8.600
Frequência	97.656
Situacao	8.750
Cidade	1.125
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.477
Ira Geral	8.234
Horas Integr	3088.000

Fonte: Elaborada pelo autor

Na Figura 7, os alunos do centroide 1 têm uma nota média de 8,6 com uma frequência no mesmo registro de 97,65%. Esses alunos são na maioria moradores de Quixadá também

como o grupo precedente. Pelos números de trancamento, insucessos e reprovação por falta que são todos iguais a 0 (zero), pode ser deduzido que nunca tiveram nenhum tipo de insucesso na disciplina de Fundamentos de Programação. Conseqüentemente, o número de matrículas é também igual a 1. Esses alunos têm um IRA individual de 8,47 e um IRA Geral de 8,23. Eles totalizaram um número de horas integralizadas de 3088 e o número de alunos constituindo esse grupo é de 8 alunos. O número de alunos matriculados nesse semestre é de 46.

No centroide 2 da Figura 28, pode ser observado que os alunos obtiveram uma nota média de 0,34 com uma frequência de 80,08%. A situação nesse grupo avizinha o número 6 que corresponde a situação de reprovação como pode ser observado na Figura 26. Alguns desses alunos tiveram reprovação por falta como indica o atributo reprovação por falta pelo número 0,24. O atributo Insucessos indica 1, o que significa que realmente todos os alunos representando este centroide reprovaram. Eles estavam na primeira tentativa deles como indica o número de matrículas igual a 1. Eles têm respectivamente uma IRA Individual e Geral de 0,99 e 3,24 e totalizaram 133,12 de horas integralizadas. Os alunos formando este grupo são 25 ou seja a maioria da turma.

5.1.2 Resultados de 2014

Na Figura 8, o centroide 2 ilustra os alunos com uma nota média de 9,37 e uma frequência igual a 91,66 %. Esses são os melhores alunos da turma e são 8 alunos no total. Eles totalizaram 2936 horas integralizadas e têm respectivamente um IRA individual e geral de 7,56 e 7,38. Eles não possuem nenhum tipo de insucessos e são todos alunos ingressantes no semestre em questão. Eles são no total 8 alunos em um total de 37 alunos matriculados.

Figura 8 – Centroide 1 e 2 de FUP 2014.1

Nota	0.891	Nota	9.375
Frequência	82.193	Frequência	91.664
Situacao	6.091	Situacao	9.000
Cidade	1.273	Cidade	1.125
Trancamento	0.000	Trancamento	0.000
Rep falta	0.091	Rep falta	0.000
Insucessos	1.364	Insucessos	0.000
Qtd matriculas	1.364	Qtd matriculas	1.000
Ira individual	2.489	Ira individual	7.566
Ira Geral	3.226	Ira Geral	7.380
Horas Integr	413.091	Horas Integr	2936.000

Fonte: Elaborada pelo autor

5.1.3 Resultados de 2015

Na Figura 9, podemos observar o centroide 1 que tem as características seguintes: a frequência é de 90,56% e a média da nota dos alunos é de 8,94. A grande maioria dos alunos foram aprovados por média. Todos os alunos desse grupo são moradores de Quixadá ou seja eles passam a semana toda em Quixadá. Esses alunos não têm nenhum tipo de insucesso que seja reprovação por falta, trancamento ou reprovação por nota. Todos esses alunos são novos ingressantes pois a quantidade de matrícula é igual a 1. Pode ser contado um total de 19 alunos nesse grupo para 54 alunos matriculados.

Figura 9 – Centroide 0 e 1 de FUP 2015.1

Nota	1.484	Nota	8.942
Frequência	82.891	Frequência	90.567
Situacao	6.211	Situacao	8.684
Cidade	1.316	Cidade	1.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.211	Rep falta	0.000
Insucessos	1.526	Insucessos	-0.000
Qtd matriculas	1.526	Qtd matriculas	1.000
Ira individual	2.949	Ira individual	7.305
Ira Geral	3.557	Ira Geral	7.141
Horas Integr	448.000	Horas Integr	1898.105

Fonte: Elaborada pelo autor

Na Figura 10, pode ser observado centroide representativo desse grupo de alunos. Os alunos obtiveram uma média de 8,3 e uma frequência de 89,84%. Neste grupo, pode ser observado que nem todos os alunos são moradores de Quixadá mas, a maioria é. Nenhum trancamento foi observado e alguns dos alunos já reprovaram a disciplina pelo menos uma vez pela quantidade de insucessos ser igual a 0,16 e a quantidade de matrícula ser 1,16. Por fim, esse grupo contém 12 alunos.

5.1.4 Resultados de 2016

O grupo de alunos seguinte representado pelo centroide 0 na Figura 11 é grupo dos alunos da disciplina nesse semestre. Eles são 9 alunos, têm uma média de 8,95 e uma frequência de 100%. Todos os alunos desse grupo tem a situação igual a 9, o que corresponde a uma aprovação por média como mostra a Figura 26. Eles não apresentam nenhum tipo de insucesso que seja trancamento, reprovação por falta e são todos alunos do primeiro semestre.

Figura 10 – Centroide 3 de FUP 2015.1

Nota	8.300
Frequência	89.841
Situacao	8.833
Cidade	1.083
Trancamento	0.000
Rep falta	0.083
Insucessos	0.167
Qtd matriculas	1.167
Ira individual	3.297
Ira Geral	3.499
Horas Integr	586.667

Fonte: Elaborada pelo autor

Figura 11 – Centroide 0 e 1 de FUP 2016.1

Nota	8.944
Frequência	100.000
Situacao	9.000
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	7.776
Ira Geral	7.573
Horas Integr	1311.556

Nota	8.590
Frequência	99.583
Situacao	8.900
Cidade	1.100
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	5.464
Ira Geral	5.451
Horas Integr	985.600

Fonte: Elaborada pelo autor

O outro grupo dos alunos com um bom rendimento é representado pelo centroide 1 como pode ser observado na Figura 34. Eles tem um total de horas integralizadas de 985,60, um IRA individual de 5,46 e um IRA geral de 5,45. Esse grupo contém alguns alunos que não moram em Quixadá mas é constituído na maioria de alunos moradores de Quixadá. A média da nota é de 8,59 e a frequência é de 99,58% de presença. A maioria dos alunos desse grupo foram aprovados por média e não apresentam nenhum tipo de insucesso. Todos são do primeiro semestre. Esses grupo é feito de 10 alunos. No total foram 39 alunos matriculados.

5.1.5 Resultados de 2017

No centroide 0 ilustrado na Figura 12, pode ser observado uma representação dos alunos desse grupo. Esses alunos obtiveram uma nota média de 8,91 e uma frequência de 96,75%. Os alunos são todos novos ingressantes pois a quantidade de matrícula é igual a 1 e eles não apresentam nenhum tipo de insucessos que seja trancamento ou reprovação por falta. Outro grupo de com um bom resultado pode ser observados na mesma figura no centroide 1. No total,

tiveram 30 alunos aprovados em uma quantidade total de 54 alunos matriculados.

Figura 12 – Centroide 0 e 1 de FUP 2017.1

Nota	8.919	Nota	9.600
Frequência	96.756	Frequência	0.000
Situacao	8.963	Situacao	3.667
Cidade	1.111	Cidade	1.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	-0.000
Insucessos	0.000	Insucessos	0.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	7.707	Ira individual	5.159
Ira Geral	7.510	Ira Geral	5.172
Horas Integr	612.741	Horas Integr	1051.333

Fonte: Elaborada pelo autor

Em 2013.1 o centroide 0 formado por alunos de média acima de 8 era composto de 8 alunos com uma frequência acima de 97%. Nesse ano, eram matriculados 46 alunos. A situação nesse ano foi preocupante, pois houve 25 reprovações, uma medida seria uma investigação pelos gestores para identificar o problema que ocorreu nesse ano para que um tal acontecimento não se reproduza. No ano 2014, foi observado um centroide que é bastante similar ao centroide 0 do ano 2013. Esse centroide do ano 2014, apresentou um grupo de 8 alunos também com uma média acima de 9. O total de alunos matriculados nesse ano era de 37 alunos. Pode ser deduzido dessa comparação que não houve um aumento considerável no desempenhos dos alunos de 2013 para 2014.

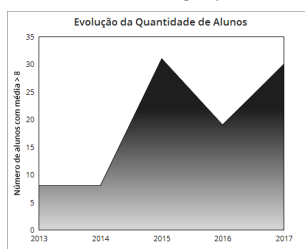
Em 2015, foram observados dois centroides de alunos com um bom desempenho. O primeiro é composto de 19 e o outro de 12 alunos todos com uma média acima de 8. Somando, foram 31 alunos em um total de 54 alunos. Comparando esse resultado com 2014 que só apresentou 8 alunos, pode-se acreditar que o desempenho melhorou de uma maneira exponencial. Vale ressaltar que a adoção da base centralizada ocorreu nesse ano de 2015. Por tanto, ela teve um impacto no desempenho dos alunos.

Em 2016, os centroides 0 e 1 formados são similares ao 1 e 3 do ano 2015 pelo fato dos alunos nesses grupos tiveram média acima de 8 e uma frequência acima de 99%. Somando, foram observados 19 alunos com essa característica do total de 39 alunos matriculados. Nesse ano, os alunos tiveram uma média de frequência melhor do que em 2015 mas por outro lado a quantidade de alunos aprovados é menor que em 2015. A quantidade de alunos matriculados em 2015 é superior à quantidade matriculados em 2016.

Em 2017 os centroides parecidos aos centroides observados em 2016 totalizam um número maior de alunos. Em 2016, a quantidade de alunos nesses centroides era de 19 enquanto em 2017, foram 30 alunos. O total de alunos matriculados em 2017 é de 54 e 2016 foram 39 alunos. Em 2016, os alunos tiveram melhores médias na nota e na frequência, mas em 2017 o número de alunos aprovados foi maior como falado anteriormente.

Nos anos de 2013 e 2014 foi observado que a quantidade de alunos com uma nota acima da média era relativamente inferior ao número de alunos matriculados nas turmas. A partir do ano de 2015, foi adotada pelos professores da disciplina de Fundamentos de Programação uma metodologia que consiste em conjunto de métodos de avaliação, uma base de trabalhos e de atividades centralizada e uma aplicação de testes contínuos para que os alunos pudessem apresentar um desempenho melhor. Com os resultados obtidos a partir do ano de 2015, foi observado que realmente essa metodologia proporcionou ao alunos da turma seguintes um desempenho melhor aos alunos dos anos anteriores. Foram observados alunos melhores médias e melhores frequência e um diminuição considerável dos alunos reprovados. Com tudo, pode-se deduzir que o uso desse conjunto de métodos de avaliações, teste contínuos e aplicação de provas pela ferramenta Moodle teve um impacto positivo no desempenho dos alunos a partir do ano 2015. Na Figura 13 pode ser observado o grafo de evolução do número de alunos com uma nota acima da média de 2013 a 2017 na disciplina de Fundamentos de Programação

Figura 13 – Grafo de evolução do número de alunos com uma nota acima da média de 2013 a 2017



Fonte: Elaborada pelo autor

5.2 Comparação do desempenho dos alunos em Estrutura de Dados

5.2.1 Resultados de 2013

Nesta etapa são analisados os resultados da mineração de dados aplicada aos dados dos alunos de Estrutura de Dados do ano de 2013 a 2017. O agrupamento com o auxílio do

algoritmo *K-means* foi aplicado às turmas do segundo semestre de 2013, as turmas onde a maioria são alunos ingressantes no mesmo ano. Dos resultados da mineração de dados, foram identificados grupos de alunos que se repetem em todos os anos sem muita alteração.

Figura 14 – Centroide 0 de ED 2013.2

Nota	8.117
Frequência	98.957
Situacao	8.833
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.412
Ira Geral	8.209
Horas Integr	3050.667

Fonte: Elaborada pelo autor

O que pode ser retido do centroide 0 é que ele é constituído de 6 alunos que ficaram uma nota média de 8,11 e uma frequência de 98,95%. É importante destacar que todos os alunos desse grupo são moradores de Quixadá e não possuem nenhum tipo de insucesso. Esses alunos são alunos do segundo semestre e passaram de primeira na disciplina pois a quantidade de matrículas é igual a 1. Eles têm um IRA individual e geral bastante importante de respectivamente 8,41 e 8,20. Ao alunos desse grupo podem ser adicionados os alunos representados pelo centroide 2 da Figura 39 composto de 2 alunos com o perfil parecido. No total são contados 8 alunos formando esses grupo com um total 14 alunos matriculados.

5.2.2 Resultados de 2014

Como observado na Figura 15, o centroide 0 grupo é um grupo em que os alunos têm uma nota média de 7,56 e uma frequência igual 97,91%. Os alunos desse grupo têm uma situação igual a 8,83 o que tende a ser correspondente à situação aprovador por média. Alguns alunos são moradores de outra cidade. Eles nunca trancaram, não tiveram reprovação por falta e se matricularam somente uma vez na disciplina. O IRA individual desses alunos é de 8,05 e o geral é igual 7,83. Eles totalizaram 3162,66 horas integralizadas no decorrer do curso e são no total 6 alunos.

Figura 15 – Centroide 0 de ED 2014.2

Nota	7.567
Frequência	97.915
Situacao	8.833
Cidade	1.167
Trancamento	-0.000
Rep falta	-0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.056
Ira Geral	7.831
Horas Integr	3162.667

Fonte: Elaborada pelo autor

5.2.3 Resultados de 2015

O centroide 0 da Figura 16 mostra um grupo de 12 alunos que tiveram um nota média de 7,65 e um frequência de 97,65%. A situação dos alunos nesse grupo é de 8,75 que tende a ser equivalente a a situação de aprovação por média como está descrito Figura 26. Todos os alunos desse grupo moram em Quixadá. Eles não tem nenhum trancamento, nenhuma reprovação por falta e enfim nenhum de tipo de insucesso. Todos os alunos são novos ingressantes no mesmo ano, pode ser afirmado isso pela quantidade de matrícula ser igual a 1. Os outros centroides nas Figuras 46 e 47 são os alunos reprovados e eles formam a maioria nessa turma. Em um total de 34 alunos, 22 reprovaram e 12 passaram.

Figura 16 – Centroide 0 de ED 2015.2

Nota	7.658
Frequência	97.653
Situacao	8.750
Cidade	1.000
Trancamento	0.000
Rep falta	-0.000
Insucessos	-0.000
Qtd matriculas	1.000
Ira individual	7.783
Ira Geral	7.579
Horas Integr	2029.333

Fonte: Elaborada pelo autor

Figura 17 – Centroide 0 de ED 2016.2

Nota	8.958
Frequência	98.046
Situacao	8.958
Cidade	1.042
Trancamento	-0.000
Rep falta	-0.000
Insucessos	-0.000
Qtd matriculas	1.000
Ira individual	6.163
Ira Geral	6.093
Horas Integr	1197.167

Fonte: Elaborada pelo autor

5.2.4 Resultados de 2016

O centroide 0 ilustrado na Figura 17 representa o grupo dos alunos com uma média acima de 8. Esse grupo é composto de 24 alunos que tiveram uma nota média igual a 8,95 e uma frequência de 98,04%. Todos os alunos desse grupos são novos ingressantes. Esses alunos em consequência não têm nenhuma reprovação por falta e nunca trancaram a disciplina. O IRA individual e geral deles é de respectivamente 6,16 e 6,09 e eles totalizaram um quantidade de 1197,16 horas integralizadas. O número de alunos no grupo é de 24 para 39 matriculados.

5.2.5 Resultados de 2017

Figura 18 – Centroide 0 de ED 2017.2

Nota	9.790
Frequência	96.874
Situacao	9.000
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	9.237
Ira Geral	8.913
Horas Integr	678.400

Fonte: Elaborada pelo autor

O centroide 0 ilustrado na Figura 18, representa o grupo dos alunos com média acima de 9. Nesse grupo constam 10 alunos que obtiveram uma nota média de 9,79 e uma frequência de 96,87%. Todos os alunos desse grupo foram aprovado por média como indica a situação igual

a 9. Todos eles são moradores de Quixadá e não apresentam nenhum tipo de insucesso que seja o trancamento, a reprovação por falta ou por nota nessa disciplina. Todos eles são do segundo semestre e totalizam 678,4 horas integralizadas. O IRA individual e geral são respectivamente de 8,23 e 8,91.

Em 2013, o centroide formado pelos alunos com uma média maior a 8 é composto de 6 alunos com uma frequência de 98,95. Quando a comparação é feita com o grupo similar em 2014, observamos que o número de alunos permaneceu o mesmo, mas houve uma diminuição na média dos alunos em 2014. Em geral os dois grupo de alunos se parecem bastante. Pode ser concluído que de 2013 para 2014 não houve uma mudança no desempenho dos alunos que ficou péssimo nesse período.

De 2014 para 2015, foram observados mudanças que valem a pena mencionar. Em 2014 o número total de alunos é de 17 enquanto em 2015 foram contados 34 alunos. Comparando os grupos de alunos que são semelhantes, pode ser observado que no ano de 2015, foram 12 alunos que apresentaram características semelhantes aos 6 alunos do mesmo grupo de 2014. Do ano 2014 para 2015, a média dos alunos baixou de 8,83 para 7,65. É importante lembrar que a quantidade total de alunos matriculados dobrou de 2014 para 2015. Para concluir esta etapa, pode ser deduzido que a quantidade dos alunos nesses grupos aumentou em 2015.

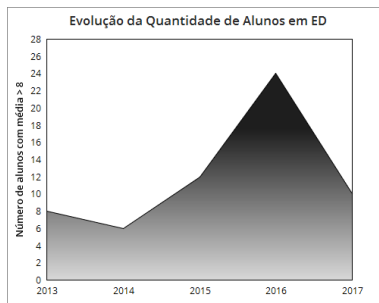
Entre os ano 2015 e 2016, o número total de alunos inscritos em 2015 é de 34 em 2016 se inscreveram 39 alunos. Em 2015, o grupo dos alunos era constituído de 12 alunos que tiveram um média de 7,65. No ano de 2016, foi escolhido o grupo de alunos que se aproximava desse grupo de 2015 e foi observado que em 2016, o grupo escolhido era constituído de 24 alunos que tinham uma nota média de 8,95. No grupo de 2016 foi observada uma quantidade maior ao grupo de 2015. Assim, o ano de 2016 apresentou um grupo maior de alunos com média melhor que o ano de 2015 com aproximadamente o mesmo número de alunos matriculados.

Nos anos de 2016 para 2017 a comparação foi feita como anteriormente entre os grupos de alunos que são semelhantes. Em 2016 se inscreveram na disciplina 39 alunos contra 42 em 2017. Em 2016, o centroide formado por alunos de média acima de 8 contava 24 alunos com uma média de 8,95. Em 2017, o centroide 0 é similar ao centroide de 2013 mas os alunos desse grupo apresentam uma média melhor que em 2016. Um aspecto que os diferencia também é o número de alunos em 2017 que é de 10. Feita essa comparação pode ser afirmado que os alunos desse grupo em 2016 apresentou uma quantidade maior que em 2017.

Feita a comparação entre os resultados obtidos da mineração de dados na disciplina

de Estrutura de Dados e a disciplina de Fundamentos de Programação, foi observado que o desempenho dos alunos não evolui da mesma maneira depois do ano de 2015 como foi o caso em Fundamentos de Programação. É importante lembrar que na disciplina de Estrutura de Dados, não foram utilizadas as mesmas metodologias de testes contínuos e aplicação de atividades, trabalhos e provas pelo Moodle como foi feito em Fundamentos de Programação. Isso pode ser um dos motivos do desempenho observado pelos alunos na disciplina de Programação Orientada a Objetos. Na Figura 19 pode ser observado uma evolução da quantidade dos alunos com uma nota acima da média de 2013 a 2017.

Figura 19 – Grafo de evolução do número de alunos com uma nota acima da média de 2013 a 2017



Fonte: Elaborada pelo autor

5.3 Comparação do desempenho dos alunos em Programação Orientada a Objetos

Nesta etapa são analisados os resultados da mineração de dados aplicada aos dados dos alunos de Programação Orientada a Objetos do ano de 2013 a 2017. O agrupamentos com o auxílio do algoritmo *K-means* foi aplicado às turmas do segundo semestre de 2013, as turmas onde a maioria são alunos ingressantes no mesmo ano. Dos resultados da mineração de dados, foram identificados grupos de alunos que se repetem em todos os anos sem muita alteração.

5.3.1 Resultados de 2013

Como ilustrado na Figura 20, foi observado o centroide 0 que representa os alunos de um certo grupo. Os alunos nesse primeiro grupo são na maioria moradores da cidade de Quixadá pelo fato que o atributo cidade indica 1,12. Os alunos ficaram com uma nota média de 7,38 e uma frequência de 96,87%. Outro ponto importante a salientar é que os alunos não apresentam nenhum trancamento, nenhuma reprovação por falta e nenhum outro tipo de insucesso. Todos os alunos da turma são do segundo semestre pois eles se matricularam uma única vez na disciplina.

O IRA individual e geral é de respectivamente 8,47 e 8,23 e a quantidade total integralizada de horas é 3088. Nesse grupo foi contado 8 alunos para um total de 16 alunos matriculados.

Figura 20 – Centroide 0 de POO 2013.2

Nota	7.387
Frequência	96.874
Situacao	8.625
Cidade	1.125
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.477
Ira Geral	8.234
Horas Integr	3088.000

Fonte: Elaborada pelo autor

5.3.2 Resultados de 2014

O centroide 0 da Figura 21 representa o grupo dos alunos com um média acima de 8. Eles obtiveram um nota média de 8,67 com uma frequência de 100%. A situação deles é de 8,74 que tende a representar uma aprovação por nota como pode confirmar a média deles. Esse grupo é constituído na maioria de alunos moradores de Quixadá. A quantidade de matrícula igual a 1 indica que todos os alunos desse grupo são do segundo semestre. Nenhum deles trancou nem reprovou por falta. Nesse grupo foram contados 7 alunos. Em 2014, foram também 16 alunos matriculados.

Figura 21 – Centroide 0 de POO 2014.2

Nota	8.671
Frequência	100.000
Situacao	8.714
Cidade	1.143
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	7.805
Ira Geral	7.600
Horas Integr	3044.571

Fonte: Elaborada pelo autor

5.3.3 Resultados de 2015

Figura 22 – Centroide 0 de POO 2015.2

Nota	7.991
Frequência	98.863
Situacao	8.818
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.001
Ira Geral	7.780
Horas Integr	2071.273

Fonte: Elaborada pelo autor

O centroide analisado na Figura 22 representa o grupo dos alunos que tiveram uma média de nota aproximando 8. Nesse grupo, os alunos ficaram com uma nota média de 7,99 e uma frequência quase perfeita de 98,66%. O atributo cidade indica o número 1, o que equivale a dizer que todos os alunos são moradores da cidade de Quixadá. Eles não apresentam nenhum tipo de insucesso que seja um trancamento, uma reprovação por falta. A quantidade de matrícula é igual a 1. O IRA individual desses alunos é de 8 e o geral 7,78. Nesse grupo são presentes 11 alunos.

5.3.4 Resultados de 2016

Figura 23 – Centroide 0 de POO 2016.2

Nota	8.511
Frequência	93.748
Situacao	9.000
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	7.776
Ira Geral	7.573
Horas Integr	1311.556

Fonte: Elaborada pelo autor

O centroide 0 ilustrado na Figura 23 representa o grupo dos alunos com uma média

acima de 8. O atributo situação indica o número 9 que significa uma aprovação por média como é ilustrado na Figura 26. Esses alunos tiveram uma nota média de 8,51 e uma frequência de 93,74%. Pelo atributo cidade ser igual a 1, pode ser afirmado que a totalidade dos alunos desse grupo são moradores de Quixadá. A quantidade de matrícula igual a 1 indica que os alunos são todos do segundo semestre ou seja são novos ingressante no mesmo ano. O IRA individual nesse grupo é de 7,77 enquanto o geral indica 7,57. A quantidade de alunos presentes nesse grupo é de 9.

5.3.5 Resultados de 2017

No centroide 0 da Figura 24, pode ser observado a representação do grupos dos alunos com uma média acima de 8. Esse alunos tiveram um nota média de 9,23 e uma frequência de 96,40%. O atributo cidade igual a 1 indica que todos os alunos dessa turma são moradores da cidade de Quixadá. Podemos observar também que o trancamento, o insucessos e a reprovação por falta são todos iguais a 0 (zero). Pela quantidade de matrículas ser igual a 1, pode ser deduzido que os alunos desse grupo cursaram a disciplina uma única vez. O IRA individual desses alunos é de 8,33 e o IRA geral é igual a 8,08. O número de alunos nesse grupo é igual a 20.

Figura 24 – Centroide 0 de POO 2017.2

Nota	9.235
Frequência	96.403
Situacao	9.000
Cidade	1.100
Trancamento	-0.000
Rep falta	-0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	8.336
Ira Geral	8.087
Horas Integr	684.800

Fonte: Elaborada pelo autor

Nesta seção é feita uma comparação dos resultados obtidos da mineração de dados aplicada aos dados dos alunos da universidade de 2013 a 2017 para a disciplina de Programação Orientada a Objetos.

Em 2013, o centroide 0 formado pelos alunos com uma média superior a 7 é constituído de 8 alunos em um total de 16 alunos matriculados, ou seja a metade. No ano de

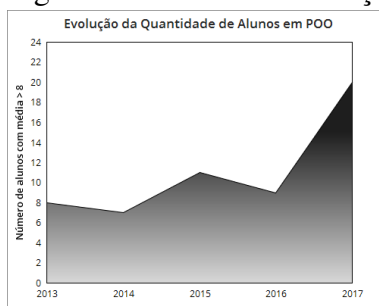
2014, o centroide 0 observado é similar ao mesmo de 2013 com algumas diferenças. Em 2014 os alunos tiveram uma média de nota maior a 2013 o seja 8,67 e uma frequência de 100%. Por outro lado a quantidade de alunos nesse grupo diminuiu para 7 e foram matriculados 16 alunos também.

Do ano 2014 para o ano 2015, a quantidade dos alunos matriculados na disciplina de POO dobrou. Passamos de 16 alunos inscritos em 2014 para 31 alunos em 2015. Em 2015, foram observados um total de 11 alunos com uma média de 7,99 no centroide similar ao centroide de 2014 que apresentou 8 alunos. Em 2015, a média dos alunos baixou junto com frequência. Os demais atributos são similares aos atributos do centroide de 2014. De 2014 para 2015 que foi o ano em que os professores fizeram a colaboração com os monitores para passar listas, foi observado um aumento do grupo dos alunos com média acima de 8.

Em 2016, o centroide 0 observado é formado por 9 alunos com uma média de 8,51 e uma frequência de 93,74%. Esse centroide é similar ao centroide de 2015, onde foi observada uma quantidade de 11 alunos com uma frequência de 98,86% melhor do que a frequência de 2016. O número total de alunos matriculados em 2016 foi de 32 e 31 em 2015. Pode ser observada uma leve diminuição na quantidade dos alunos nesse grupo de 2016.

O maior aumento da quantidade dos alunos nos grupos com média acima de 9 foi observado em 2017. O centroide 0 observado no ano de 2017 é formado de 20 alunos com uma média acima de 9 e uma frequência de 96,40%. Em comparação ao centroide parecido observado em 2016, foi contado 9 alunos com uma média acima de 8 e uma frequência 93%. Em 2017, foram matriculados 36 alunos contra 32 em 2016. O maior desempenho dos alunos foi observado nesse ano 2017.

Figura 25 – Gráfico de evolução da quantidade de alunos com nota acima da média



Fonte: Elaborada pelo autor

Comparando os resultados obtidos da mineração de dados na disciplina de Programação Orientada a Objetos e a disciplina de Fundamentos de Programação, foi observado

que em Programação Orientada a Objetos, o desempenho dos alunos não evolui da mesma maneira depois do ano de 2015 como foi o caso em Fundamentos de Programação. É importante lembrar que na disciplina de Programação Orientada a Objetos, não foram utilizadas as mesmas metodologias de testes contínuos e aplicação de atividades, trabalhos e provas pelo *Moodle* como foi feito em Fundamentos de Programação. Esse fato pode ser um motivo do desempenho observado pelos alunos na disciplina de Programação Orientada a Objetos. Na Figura 25 pode ser observado uma evolução da quantidade dos alunos com uma nota acima da média de 2013 a 2017.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho contribuiu a demonstrar a possibilidade de encontrar diversos grupos de alunos da Universidade Federal do Ceará Campus de Quixadá usando as técnicas de mineração de dados. Foi igualmente possível acompanhar o desempenho dos alunos do curso de Ciência da Computação nas disciplinas de Fundamentos de programação, Estrutura de Dados e Programação Orientada a Objetos do ano 2013 até o ano de 2017. Com o uso do método de agrupamento de mineração de dados com o algoritmo *K-means*, foi possível agrupar os alunos em função dos atributos especificados.

O objetivo principal deste trabalho foi a análise do impacto da adoção de uma nova metodologia de avaliação dos alunos na disciplina de Fundamentos de Programação. Esse novo método de avaliação é baseado na aplicação de testes contínuos, de provas, atividades e trabalhos em sala de aula com auxílio da ferramenta *Moodle*. O impacto da implementação dessa nova metodologia, como foi observado no Capítulo 5, foi positivo no desempenho dos alunos na disciplina de Fundamentos de Programação a partir do ano de 2015. Os alunos a partir de 2015 para frente tiveram um desempenho melhor e foi observado uma consistência durante os seguintes anos. O resultado realmente mostrou que o uso dessa metodologia como foi seu objetivo principal, realmente ajudou os alunos de Fundamentos de Programação a ter um rendimento melhor. Na disciplina de Estrutura, foi observado que a colaboração dos professores com os monitores da disciplina não teve um efeito imediato no desempenho dos alunos. Foi observado somente em 2016 um real aumento no grupos dos alunos com nota acima da média. O mesmo caso se repetiu com a disciplina de Programação Orientada a Objetos. A melhoria notável na quantidade dos alunos com nota acima da média foi observado somente no ano 2017 e não teve uma padronização nos resultados dos alunos. Portanto, não pode ser afirmado que a colaboração dos professores com os monitores das disciplinas a partir de 2015 teve um impacto negativo ou positivo sobre o desempenho dos alunos nas disciplinas de Estrutura de Dados e Programação Orientada a Objetos. Com os resultados desse trabalho, os gestores serão capazes de analisar e tomar futuras decisões que poderiam ajudar de uma certa maneira os alunos.

No decorrer da realização do trabalho, os aspectos mais desafiadores foram lidar com o tempo para fazer o estudo, a aprendizagem da tecnologia mas também os desafios que surgiram ao longo do trabalho. A investigação na área da mineração de dados proporcionou um interesse particular em conhecer mais as possibilidades de coisas que podem ser feitas usando esta técnica.

Como trabalhos futuros, pode ser feita uma análise que considere um número maior de disciplinas. Também a possibilidade de aplicar os mesmos métodos para os dados dos outros cursos do Campus de Quixadá a fim de ajudar os professores a monitorar o efeito de certas tomadas de decisões sobre o desempenho dos alunos. Seria interessante também no futuro, junto a mineração de dados, fazer um estudo de campo a fim de questionar os alunos sobre suas opiniões sobre as metodologias usadas e disponibilizar em uma plataforma, os resultados dos estudos.

REFERÊNCIAS

- BARADWAJ, B. K.; PAL, S. Mining educational data to analyze students' performance. Rajasthan India, **arXiv preprint arXiv:1201.3417**, 2012.
- CABENA, P.; HADJINIAN, P.; STADLER, R.; VERHEES, J.; ZANASI, A.; JOSE, C. I. B. M. C. S.; JOSE, C. I. T. S. O. S. **Discovering data mining: from concept to implementation**. New Jersey: Prentice Hall PTR, 1997.
- CAMILO, C. O.; SILVA, J. C. d. **Mineração de dados: conceitos, tarefas, métodos e ferramentas**. Goiás, GO: Universidade Federal de Goiás (UFG), p. 1–29, 2009.
- FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. The kdd process for extracting useful knowledge from volumes of data. **Communications of the ACM**, ACM, [S.l.], v. 39, n. 11, p. 27–34, 1996.
- FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P. *et al.* Knowledge discovery and data mining: Towards a unifying framework. **KDD-96: Proceedings**, [S.l.], v. 96, p. 82–88, 1996.
- HAND, D. J.; MANNILA, H.; SMYTH, P. **Principles of data mining (adaptive computation and machine learning)**. Cambridge, MA: MIT Press, 2001.
- KODINARIYA, T. M.; MAKWANA, P. R. Review on determining number of cluster in k-means clustering. **International Journal**, [S.l.], v. 1, n. 6, p. 90–95, 2013.
- MONDARDO, A. H.; PEDON, E. A. Estresse e desempenho acadêmico em estudantes universitários. **Revista de Ciências Humanas**, [S.l.], v. 6, n. 6, p. 159–180, 2005.
- PANG-NING MICHAEL STEINBACH, V. K. **Introdução ao Data Mining Mineração de Dados**. Rio de Janeiro - RJ: Editora Ciência Moderna, 2009. 585–593 p.
- SILVA, M. P. S. Mineração de dados- conceitos, aplicações e experimentos com weka. São José dos Campos-SP, **Instituto Nacional de Pesquisas Espaciais (INEP)**, 2004.
- SILVA, T. J. da. **Extração de conhecimento nos dados da universidade federal do ceará via mineração de dados: descoberta a análise dos perfis dos alunos**. 2014. Disponível em: <http://www.repositoriobib.ufc.br/000017/000017cf.pdf>. Acesso em: 21 ago. 2019.
- VIALARDI, C.; BRAVO, J.; SHAFTI, L.; ORTIGOSA, A. Recommendation in higher education using data mining techniques. **International Working Group on Educational Data Mining**, ERIC, [S.l.], 2009.

APÊNDICE A – FIGURAS RELACIONADAS A FUNDAMENTAÇÃO DE PROGRAMAÇÃO DE 2013 ATÉ 2017

Aqui são listados as diferentes figuras relacionadas a disciplina de Fundamentos de Programação nos anos de 2013 a 2017.

Figura 26 – Tabela descritiva da situação do aluno.

Situação	Descrição
1	TRANCADO
2	CANCELADO
3	APROVT INTERNO
4	MATRICULADO
5	APROVT EXTERNO
6	REPROVADO
7	REP. FALTA
8	APROVADO
9	APROVADO MÉDIA

Fonte: Elaborado pelo autor

Figura 27 – O centroide 1 de FUP 2013.1

Nota	5.690
Frequência	97.290
Situacao	7.600
Cidade	1.100
Trancamento	0.000
Rep falta	0.000
Insucessos	0.300
Qtd matriculas	1.000
Ira individual	4.153
Ira Geral	5.308
Horas Integr	729.600

Fonte: Elaborado pelo autor

Figura 28 – Os centroides 2 e 3 de FUP 2013.1

Nota	0.340	Nota	2.133
Frequência	80.080	Frequência	0.000
Situacao	6.240	Situacao	2.000
Cidade	1.040	Cidade	1.000
Trancamento	0.000	Trancamento	0.333
Rep falta	0.240	Rep falta	0.000
Insucessos	1.000	Insucessos	0.333
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	0.995	Ira individual	0.519
Ira Geral	3.245	Ira Geral	1.612
Horas Integr	133.120	Horas Integr	197.333

Fonte: Elaborado pelo autor

Figura 29 – Grupos de alunos de FUP 2013.1



Fonte: Elaborado pelo autor

Figura 30 – Grupos de alunos de FUP 2014.1



Fonte: Elaborado pelo autor

Figura 31 – Centroide 0 de FUP 2014.1

Nota	6.644
Frequência	87.381
Situacao	8.333
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.556
Qtd matriculas	1.444
Ira individual	3.571
Ira Geral	4.540
Horas Integr	522.667

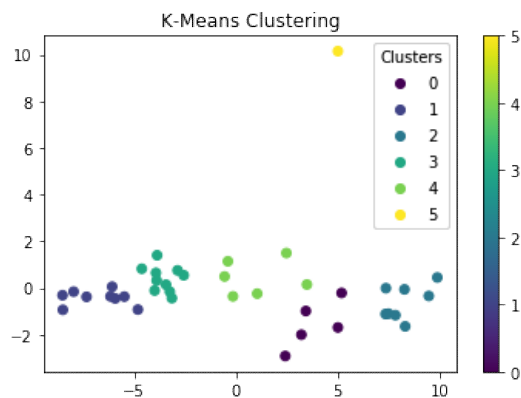
Fonte: Elaborado pelo autor

Figura 32 – Grupos de alunos de FUP 2015.1



Fonte: Elaborado pelo autor

Figura 33 – Grupos de alunos de FUP 2016.1



Fonte: Elaborado pelo autor

Figura 34 – Centroide 2 e 3 de FUP 2016.1

Nota	0.763	Nota	2.460
Frequência	86.716	Frequência	96.248
Situacao	6.250	Situacao	6.000
Cidade	1.625	Cidade	1.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.625	Rep falta	0.000
Insucessos	1.625	Insucessos	1.000
Qtd matriculas	1.625	Qtd matriculas	1.000
Ira individual	1.122	Ira individual	3.877
Ira Geral	1.689	Ira Geral	3.990
Horas Integr	132.000	Horas Integr	460.800

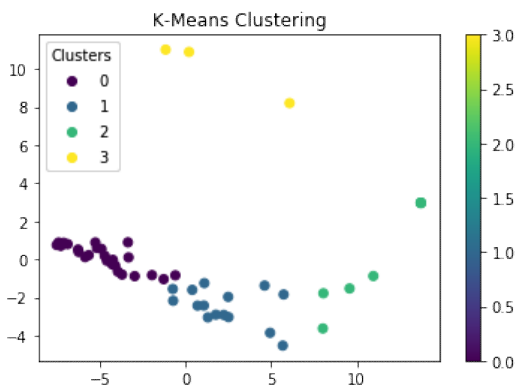
Fonte: Elaborado pelo autor

Figura 35 – Centroide 4 e 5 de FUP 2016.1

Nota	6.317	Nota	5.700
Frequência	97.222	Frequência	0.000
Situacao	8.500	Situacao	3.000
Cidade	1.167	Cidade	2.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	0.000
Insucessos	-0.000	Insucessos	0.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	3.203	Ira individual	0.373
Ira Geral	3.354	Ira Geral	0.780
Horas Integr	608.000	Horas Integr	864.000

Fonte: Elaborado pelo autor

Figura 36 – Grupos de alunos de FUP 2017.1



Fonte: Elaborado pelo autor

Figura 37 – Centroide 2 e 3 de FUP 2017.1

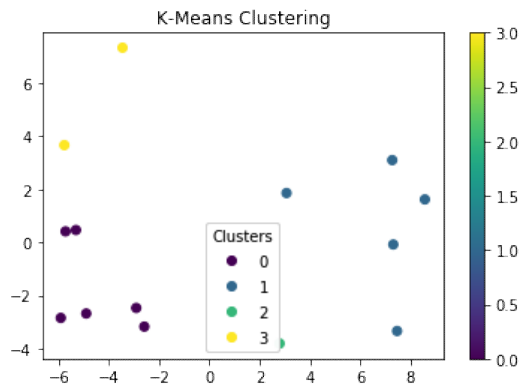
Nota	0.312	Nota	3.433
Frequência	25.780	Frequência	90.969
Situacao	4.500	Situacao	6.667
Cidade	1.125	Cidade	1.333
Trancamento	0.125	Trancamento	0.000
Rep falta	0.500	Rep falta	0.133
Insucessos	0.750	Insucessos	0.733
Qtd matriculas	1.250	Qtd matriculas	1.000
Ira individual	0.330	Ira individual	4.638
Ira Geral	0.482	Ira Geral	4.679
Horas Integr	32.000	Horas Integr	298.667

Fonte: Elaborado pelo autor

APÊNDICE B – FIGURAS RELACIONADAS A ESTRUTURA DE DADOS DE 2013 ATÉ 2017

Aqui são listados as diferentes figuras relacionadas a disciplina de Estrutura de dados nos anos de 2013 até 2017.

Figura 38 – Grupos de alunos de ED 2013.2



Fonte: Elaborado pelo autor

Figura 39 – Centroides 1 e 2 de ED 2013.2

Nota	7.700	Nota	8.050
Frequência	100.000	Frequência	92.185
Situacao	8.000	Situacao	8.500
Cidade	1.000	Cidade	1.500
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	0.000
Insucessos	0.000	Insucessos	0.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	6.739	Ira individual	8.672
Ira Geral	7.757	Ira Geral	8.308
Horas Integr	608.000	Horas Integr	3200.000

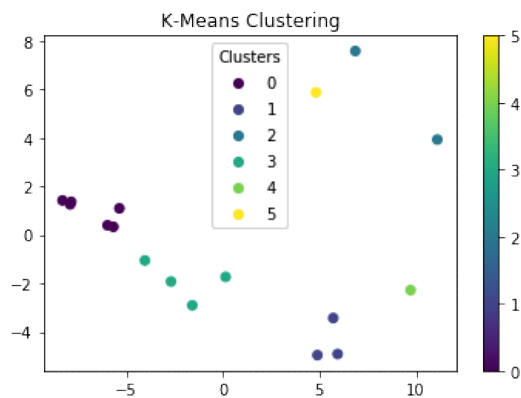
Fonte: Elaborado pelo autor

Figura 40 – Centroide 3 de ED 2013.2

Nota	3.820
Frequência	96.248
Situacao	6.400
Cidade	1.200
Trancamento	0.000
Rep falta	0.000
Insucessos	0.800
Qtd matriculas	1.000
Ira individual	3.762
Ira Geral	4.985
Horas Integr	620.800

Fonte: Elaborado pelo autor

Figura 41 – Grupos de alunos de ED 2014.2



Fonte: Elaborado pelo autor

Figura 42 – Centroide 1 de ED 2014.2

Nota	2.000
Frequência	88.670
Situacao	6.000
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	1.000
Qtd matriculas	1.000
Ira individual	3.471
Ira Geral	4.173
Horas Integr	736.000

Fonte: Elaborado pelo autor

Figura 43 – Centroide 2 e 3 de ED 2014.2

Nota	6.900	Nota	0.000
Frequência	0.000	Frequência	0.000
Situacao	3.000	Situacao	1.000
Cidade	1.000	Cidade	1.000
Trancamento	0.000	Trancamento	1.000
Rep falta	0.000	Rep falta	0.000
Insucessos	0.000	Insucessos	1.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	6.546	Ira individual	4.520
Ira Geral	6.797	Ira Geral	4.683
Horas Integr	864.000	Horas Integr	1440.000

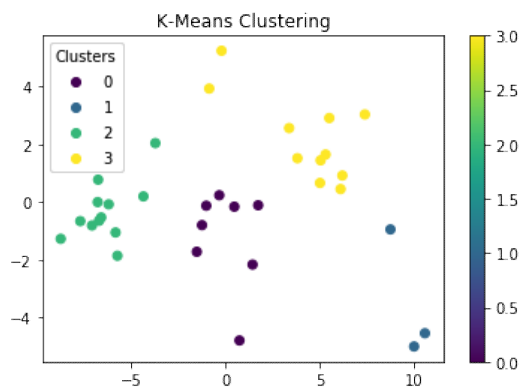
Fonte: Elaborado pelo autor

Figura 44 – Centroide 4 e 5 de ED 2014.2

Nota	7.833	Nota	0.000
Frequência	92.707	Frequência	28.120
Situacao	8.667	Situacao	7.000
Cidade	1.000	Cidade	1.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	1.000
Insucessos	0.667	Insucessos	1.000
Qtd matriculas	1.667	Qtd matriculas	1.000
Ira individual	5.530	Ira individual	0.206
Ira Geral	5.792	Ira Geral	1.059
Horas Integr	1813.333	Horas Integr	560.000

Fonte: Elaborado pelo autor

Figura 45 – Grupos de alunos de ED 2015.2



Fonte: Elaborado pelo autor

Figura 46 – Centroide 1 de ED 2015.2

Nota	1.909
Frequência	81.531
Situacao	6.727
Cidade	1.182
Trancamento	0.091
Rep falta	0.273
Insucessos	1.091
Qtd matriculas	1.273
Ira individual	3.298
Ira Geral	3.468
Horas Integr	695.273

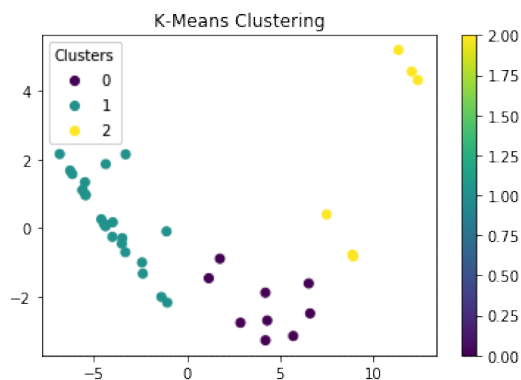
Fonte: Elaborado pelo autor

Figura 47 – Centroide 2 e 3 de ED 2015.2

Nota	2.125	Nota	0.000
Frequência	91.014	Frequência	7.290
Situacao	6.125	Situacao	3.000
Cidade	1.000	Cidade	1.000
Trancamento	0.250	Trancamento	0.667
Rep falta	0.250	Rep falta	0.667
Insucessos	1.625	Insucessos	1.667
Qtd matriculas	1.625	Qtd matriculas	1.667
Ira individual	5.885	Ira individual	3.207
Ira Geral	5.838	Ira Geral	3.633
Horas Integr	1628.000	Horas Integr	394.667

Fonte: Elaborado pelo autor

Figura 48 – Grupos de alunos de ED 2016.2



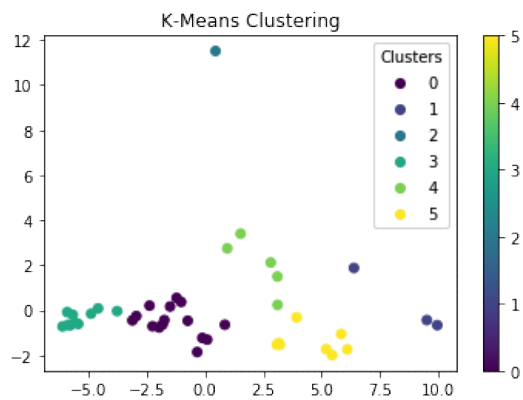
Fonte: Elaborado pelo autor

Figura 49 – Centroide 1 e 2 de ED 2016.2

Nota	1.422	Nota	0.000
Frequência	88.193	Frequência	18.750
Situacao	6.111	Situacao	4.000
Cidade	1.111	Cidade	1.167
Trancamento	0.111	Trancamento	0.500
Rep falta	0.111	Rep falta	0.500
Insucessos	1.222	Insucessos	1.000
Qtd matriculas	1.222	Qtd matriculas	1.000
Ira individual	4.105	Ira individual	2.532
Ira Geral	4.204	Ira Geral	2.745
Horas Integr	821.333	Horas Integr	458.667

Fonte: Elaborado pelo autor

Figura 50 – Grupos de alunos de ED 2017.2



Fonte: Elaborado pelo autor

Figura 51 – Centroide 1 de ED 2017.2

Nota	6.833
Frequência	91.872
Situacao	8.333
Cidade	1.200
Trancamento	0.000
Rep falta	-0.000
Insucessos	0.067
Qtd matriculas	1.000
Ira individual	7.252
Ira Geral	7.092
Horas Integr	629.333

Fonte: Elaborado pelo autor

Figura 52 – Centroide 2 e 3 de ED 2017.2

Nota	3.212	Nota	1.100
Frequência	86.324	Frequência	70.830
Situacao	6.000	Situacao	6.667
Cidade	1.000	Cidade	1.333
Trancamento	0.000	Trancamento	0.333
Rep falta	-0.000	Rep falta	1.000
Insucessos	1.125	Insucessos	2.333
Qtd matriculas	1.125	Qtd matriculas	2.333
Ira individual	4.609	Ira individual	1.929
Ira Geral	4.667	Ira Geral	2.207
Horas Integr	404.000	Horas Integr	480.000

Fonte: Elaborado pelo autor

Figura 53 – Centroide 4 e 5 de ED 2017.2

Nota	8.500	Nota	6.260
Frequência	0.000	Frequência	76.248
Situacao	5.000	Situacao	7.800
Cidade	1.000	Cidade	1.000
Trancamento	0.000	Trancamento	0.400
Rep falta	0.000	Rep falta	0.400
Insucessos	0.000	Insucessos	1.200
Qtd matriculas	1.000	Qtd matriculas	2.000
Ira individual	5.385	Ira individual	4.264
Ira Geral	5.379	Ira Geral	4.351
Horas Integr	1458.000	Horas Integr	742.400

Fonte: Elaborado pelo autor

APÊNDICE C – FIGURAS RELACIONADAS A PROGRAMAÇÃO ORIENTADA A OBJETOS DE 2013 ATÉ 2017

Aqui são listados as diferentes figuras relacionadas a disciplina de Programação Orientada a Objetos nos anos de 2013 até 2017.

Figura 54 – Grupos de alunos de POO 2013.2



Fonte: Elaborado pelo autor

Figura 55 – Centroide 1 de POO 2013.2

Nota	6.067
Frequência	96.873
Situacao	8.000
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.000
Qtd matriculas	1.000
Ira individual	4.900
Ira Geral	5.723
Horas Integr	906.667

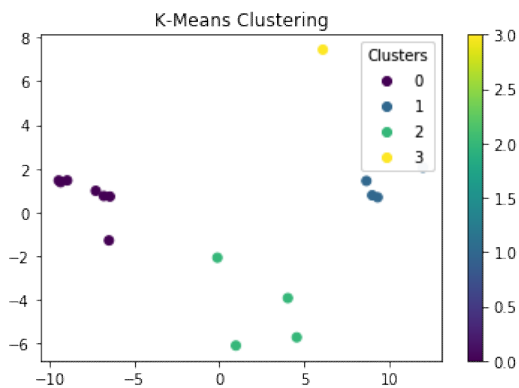
Fonte: Elaborado pelo autor

Figura 56 – Centroide 2 e 3 de POO 2013.2

Nota	3.250	Nota	6.200
Frequência	96.875	Frequência	81.250
Situacao	6.000	Situacao	8.000
Cidade	1.250	Cidade	1.000
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	0.000
Insucessos	1.000	Insucessos	0.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	2.763	Ira individual	6.005
Ira Geral	4.143	Ira Geral	6.963
Horas Integr	388.000	Horas Integr	1056.000

Fonte: Elaborado pelo autor

Figura 57 – Grupos de alunos de POO 2014.2



Fonte: Elaborado pelo autor

Figura 58 – Centroide 1 de POO 2014.2

Nota	0.000
Frequência	3.905
Situacao	5.500
Cidade	1.000
Trancamento	0.250
Rep falta	0.750
Insucessos	1.250
Qtd matriculas	1.250
Ira individual	2.168
Ira Geral	3.217
Horas Integr	468.000

Fonte: Elaborado pelo autor

Figura 59 – Centroide 2 e 3 de POO 2014.2

Nota	2.300	Nota	0.000
Frequência	96.092	Frequência	0.000
Situacao	6.500	Situacao	1.000
Cidade	1.000	Cidade	1.000
Trancamento	0.000	Trancamento	1.000
Rep falta	0.000	Rep falta	0.000
Insucessos	0.750	Insucessos	1.000
Qtd matriculas	1.000	Qtd matriculas	1.000
Ira individual	3.735	Ira individual	5.887
Ira Geral	4.422	Ira Geral	5.840
Horas Integr	776.000	Horas Integr	2176.000

Fonte: Elaborado pelo autor

Figura 60 – Grupos de alunos de POO 2015.2



Fonte: Elaborado pelo autor

Figura 61 – Centroide 1 e 2 de POO 2015.2

Nota	6.638	Nota	0.400
Frequência	96.875	Frequência	23.957
Situacao	8.125	Situacao	7.000
Cidade	1.000	Cidade	1.333
Trancamento	0.000	Trancamento	0.000
Rep falta	0.000	Rep falta	1.333
Insucessos	0.250	Insucessos	1.333
Qtd matriculas	1.125	Qtd matriculas	1.333
Ira individual	5.592	Ira individual	2.769
Ira Geral	5.569	Ira Geral	3.224
Horas Integr	1472.000	Horas Integr	554.667

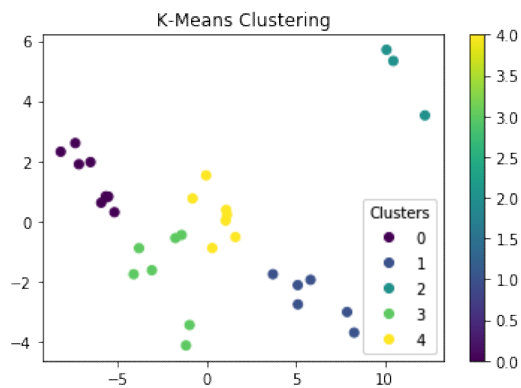
Fonte: Elaborado pelo autor

Figura 62 – Centroide 3 e 4 de POO 2015.2

Nota	7.991	Nota	0.000
Frequência	98.863	Frequência	0.000
Situacao	8.818	Situacao	1.333
Cidade	1.000	Cidade	1.000
Trancamento	0.000	Trancamento	1.333
Rep falta	0.000	Rep falta	0.000
Insucessos	0.000	Insucessos	1.333
Qtd matriculas	1.000	Qtd matriculas	1.667
Ira individual	8.001	Ira individual	2.852
Ira Geral	7.780	Ira Geral	3.439
Horas Integr	2071.273	Horas Integr	421.333

Fonte: Elaborado pelo autor

Figura 63 – Grupos de alunos de POO 2016.2



Fonte: Elaborado pelo autor

Figura 64 – Centroide 1 e 2 de POO 2016.2

Nota	0.983	Nota	3.700
Frequência	73.957	Frequência	87.497
Situacao	6.333	Situacao	6.000
Cidade	1.167	Cidade	1.143
Trancamento	0.000	Trancamento	0.000
Rep falta	0.500	Rep falta	0.000
Insucessos	2.167	Insucessos	1.000
Qtd matriculas	2.167	Qtd matriculas	1.000
Ira individual	2.735	Ira individual	5.307
Ira Geral	2.950	Ira Geral	5.307
Horas Integr	538.667	Horas Integr	964.571

Fonte: Elaborado pelo autor

Figura 65 – Centroide 3 de POO 2016.2

Nota	7.529
Frequência	93.749
Situacao	8.571
Cidade	1.000
Trancamento	0.000
Rep falta	0.000
Insucessos	0.143
Qtd matriculas	1.143
Ira individual	4.974
Ira Geral	5.022
Horas Integr	896.000

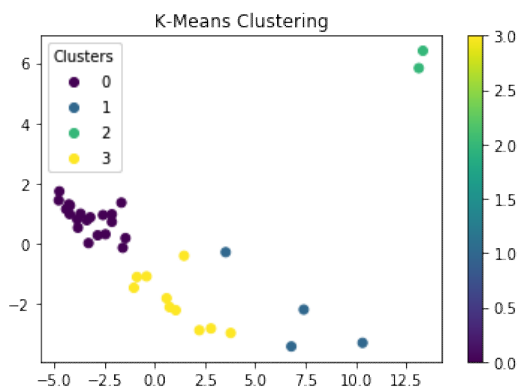
Fonte: Elaborado pelo autor

Figura 66 – Centroide 4 de POO 2016.2

Nota	0.000
Frequência	0.000
Situacao	1.000
Cidade	1.333
Trancamento	1.000
Rep falta	0.000
Insucessos	1.333
Qtd matriculas	1.333
Ira individual	3.048
Ira Geral	3.240
Horas Integr	554.667

Fonte: Elaborado pelo autor

Figura 67 – Grupos de alunos de POO 2017.2



Fonte: Elaborado pelo autor

Figura 68 – Centroide 1 e 2 de POO 2017.2

Nota	0.000	Nota	7.500
Frequência	0.000	Frequência	90.622
Situacao	1.000	Situacao	8.800
Cidade	1.000	Cidade	1.100
Trancamento	1.500	Trancamento	0.100
Rep falta	0.000	Rep falta	-0.000
Insucessos	1.500	Insucessos	0.200
Qtd matriculas	1.500	Qtd matriculas	1.200
Ira individual	3.102	Ira individual	5.489
Ira Geral	3.284	Ira Geral	5.474
Horas Integr	560.000	Horas Integr	499.200

Fonte: Elaborado pelo autor

Figura 69 – Centroide 3 de POO 2017.2

Nota	2.075
Frequência	85.153
Situacao	6.250
Cidade	1.000
Trancamento	0.000
Rep falta	0.250
Insucessos	1.000
Qtd matriculas	1.000
Ira individual	4.021
Ira Geral	4.128
Horas Integr	376.000

Fonte: Elaborado pelo autor