

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICA**

**ESTUDO DAS SIMILARIDADES ENTRE OS CURSOS DE
GRADUAÇÃO EM MATEMÁTICA COM BASE NAS
VARIÁVEIS DO CONCEITO PRELIMINAR DE CURSO: UMA
APLICAÇÃO DO MÉTODO DE AGRUPAMENTO K-MÉDIAS**

Beli Thomaz Marcatti

**Florianópolis
2021**

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICA**

**ESTUDO DAS SIMILARIDADES ENTRE OS CURSOS DE
GRADUAÇÃO EM MATEMÁTICA COM BASE NAS
VARIÁVEIS DO CONCEITO PRELIMINAR DE CURSO: UMA
APLICAÇÃO DO MÉTODO DE AGRUPAMENTO K-MÉDIAS**

Beli Thomaz Marcatti

Trabalho de Conclusão de Curso
apresentado ao Curso de Matemática, do
Departamento de Matemática – Centro de
Ciências Físicas e Matemáticas da
Universidade Federal de Santa Catarina, para
obtenção de grau de Licenciada em
Matemática.

Orientadora: Dra. Andreia Zanella

**Florianópolis
2021**

Beli Thomaz Marcatti

**ESTUDO DAS SIMILARIDADES ENTRE OS CURSOS DE
GRADUAÇÃO EM MATEMÁTICA COM BASE NAS
VARIÁVEIS DO CONCEITO PRELIMINAR DE CURSO: UMA
APLICAÇÃO DO MÉTODO DE AGRUPAMENTO K-MÉDIAS**

Trabalho de Conclusão de Curso apresentado ao Curso de Matemática, do Departamento de Matemática – Centro de Ciências Físicas e Matemáticas da Universidade Federal de Santa Catarina, para obtenção de grau de Licenciada em Matemática.

BANCA EXAMINADORA

Prof^a. Andreia Zanella (Orientadora)

Universidade Federal de Santa Catarina

Prof^a. Andréa Cristina Konrath

Universidade Federal de Santa Catarina

Prof^a. Silvia Martini de Holanda

Universidade Federal de Santa Catarina

**Florianópolis
2021**

AGRADECIMENTOS

Em primeiro lugar, agradeço aos meus pais e à minha irmã, que sempre estiveram do meu lado me apoiando, ajudando e lembrando do que é importante. Agradeço a toda minha família que sempre esteve presente e me trouxe felicidade e amor. E minha prima Giovana Cicala, que me incentivou incansavelmente nesses últimos meses de curso.

A faculdade me proporcionou conhecer algumas pessoas especiais que melhoraram meus dias lá dentro, como a Thais de Carvalho, a Bruna Scheibel e o Daniel Simas. Vocês me ajudaram nos momentos mais difíceis e eu sempre serei grata. Agradeço também àqueles que de alguma forma fizeram parte dessa história cheia de desafios como Matheus Pimenta, Mateus Patrício, Jéssika Silva, Andresa de Liz, Deborah Portela, Carlos de Castro e Pedro Lourenço.

Aos meus amigos do *cheerleading*, agradecimentos sinceros por tornarem a experiência universitária mais rica e divertida. Um agradecimento especial à Luiza Avila e ao Rodrigo Schifini, que já são meus amigos de vida. Vocês me ajudaram e estiveram comigo em incontáveis momentos felizes. Agradeço por tudo, tudo mesmo.

Agradeço aos meus professores por todos os ensinamentos. E à minha orientadora Andreia Zanella, que topou fazer este trabalho todo em 2-3 meses. Obrigada por toda ajuda e orientação.

Sou grata por todos que participaram do meu processo de formação acadêmica e de alguma forma me ajudaram a chegar até aqui. E sou grata a mim mesma, que consegui vencer muitos obstáculos e momentos intensos, para chegar neste ponto, alcançar este objetivo e me tornar quem sou hoje.

RESUMO

A avaliação da Educação Superior é uma importante ferramenta para o desenvolvimento da qualidade do ensino. Atualmente, o Conceito Preliminar de Curso (CPC) é considerado o principal indicador de qualidade da educação superior brasileira. Este trabalho utiliza as notas que compõem o CPC para analisar a existência de agrupamentos entre os cursos de graduação de Licenciatura em Matemática na modalidade presencial em instituições federais. Para o agrupamento dos cursos, utilizou-se o método não hierárquico denominado k-médias, o qual possibilitou dividir os cursos em três grupos. Os resultados mostraram que o primeiro *cluster* foi formado por cursos com notas altas nas componentes referentes ao aprendizado do aluno e às titulações dos professores, o segundo *cluster* uniu os cursos com notas intermediárias na maioria das componentes e o terceiro *cluster* agrupou os cursos com notas mais altas nas componentes referentes à percepção dos alunos sobre a estrutura do curso.

Palavras-chave: Análise de *cluster*, k-médias, Avaliação da educação superior.

ABSTRACT

The evaluation of the Higher Education is an important tool for the improvement of the education's quality. Nowadays, the Preliminary Undergraduate Program Grade (Conceito Preliminar de Curso - CPC) is considered the main quality indicator of the brazilian higher education. This paper uses the scores that compose the CPC to analyze the existence of clusters among the undergraduate programs in Mathematics in face-to-face modality of federal institutions. The clustering technique used was the non-hierarchical method called k-means, which divided the undergraduate courses into three clusters. The results showed that the first cluster was formed by courses with high grades in the components related to the student's learning and to the professors' qualifications, the second cluster brought together the courses with intermediate grades in most of the components, and the third cluster grouped the courses with higher grades in the components related to the students perceptions of the course's structure.

Keywords: Cluster analysis, K-means, Evaluation of undergraduate courses.

LISTA DE FIGURAS

Figura 1 - Dendrograma que ilustra o agrupamento hierárquico	22
Figura 2 - Exemplo de ligação simples.....	23
Figura 3 - Histogramas das variáveis NC, NIDD, NO, NF, NA, NM, ND e NR, respectivamente	34
Figura 4 - Componentes dos centros dos clusters finais	39
Figura 5 - <i>Boxplots</i> das variáveis NC, NIDD, NO, NF, NA, NM e ND respectivamente	41
Figura 6 - Gráfico dos <i>clusters</i> de acordo com o tipo de organização acadêmica	44
Figura 7 - CPC dos cursos por <i>cluster</i>	45
Figura 8 - Gráficos de dispersão relacionando NC com NIDD, NM com ND, NF com NO e NO com NA respectivamente.....	47

LISTA DE TABELAS

Tabela 1 - Composição do CPC e pesos das suas dimensões e componentes	18
Tabela 2 - Principais medidas descritivas para as variáveis	33
Tabela 3 - Distribuição dos cursos nos <i>clusters</i>	37

SUMÁRIO

1 INTRODUÇÃO	9
1.1 Objetivos	10
1.2 Estrutura do Trabalho	10
2 REFERENCIAL TEÓRICO.....	12
2.1 A Avaliação da Educação Superior Brasileira	12
2.1.1 Contextualização histórica: Dos primeiros esforços para avaliar a Educação Superior até os dias atuais	12
2.1.2 O SINAES	15
2.1.3 Conceito Preliminar de Curso (CPC)	17
2.2 Análise de <i>Clusters</i>	20
2.2.1 Hierárquicos	21
2.2.1.1 Ligação Simples	23
2.2.1.2 Ligação Completa	23
2.2.1.3 Ligação Média	24
2.2.1.4 Método Centróide	24
2.2.1.5 Método Ward	25
2.2.2 Não-hierárquicos	25
2.2.2.1 K-médias	26
2.3 Análise de Cluster aplicada na avaliação da educação superior	27
3 METODOLOGIA	31
3.1 Descrição dos dados utilizados na pesquisa	31
3.2 Métodos aplicados para a análise dos dados	31
4 RESULTADOS.....	33
4.1 Análise exploratória e descritiva dos dados	33
4.2 Análise de agrupamento	35
4.3 Análise do curso de Licenciatura em Matemática da UFSC Florianópolis	46
5 CONSIDERAÇÕES FINAIS	49
REFERÊNCIAS.....	51
ANEXO A	54

1 INTRODUÇÃO

A avaliação das Instituições de Ensino Superior (IES) no Brasil começou a ser feita em 1983 e, desde então, diferentes programas de avaliação foram propostos, até que em 2004 foi instituído o Sistema Nacional de Avaliação da Educação Superior (SINAES), que é o programa vigente atualmente (Zainko, 2008). No Brasil, a avaliação é realizada pelo Ministério da Educação (MEC) por meio do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), que busca contribuir com o desenvolvimento da educação no país. Por isso a avaliação da educação superior é tão importante, a partir dela obtém-se as informações necessárias para buscar melhores padrões de qualidade de ensino.

Um dos indicadores de qualidade da educação superior elaborados pelo Inep é o Conceito Preliminar de Curso (CPC). Atualmente é considerado o principal Indicador da Qualidade da Educação Superior voltado aos cursos superiores brasileiros (IKUTA, 2016). O indicador é calculado a partir de uma média ponderada entre oito componentes: Nota dos Concluintes no Enade, Nota do Indicador da Diferença entre os Desempenhos Observado e Esperado, Nota de Proporção de Mestres, Nota de Proporção de Doutores, Nota de Regime de Trabalho, Nota referente à Organização Didático-Pedagógica, Nota referente à Infraestrutura e Instalações Fiscais e Nota referente às Oportunidades de Ampliação da Formação Acadêmica e Profissional. Esta média é transformada em um conceito entre 1 e 5, e todo ano são avaliados cursos diferentes de forma que um mesmo curso seja avaliado a cada três anos (BRASIL, 2020).

O CPC, como indicador de qualidade, é uma ferramenta útil para obter informações a respeito da situação da Educação Superior do país e para analisar e direcionar o desenvolvimento e progresso da qualidade educacional (IKUTA, 2016).

Os cursos de graduação vêm trabalhando para melhorar a qualidade de ensino e, por consequência, obter um conceito alto na avaliação do MEC. Uma forma de identificar oportunidades de melhoria é a observação das boas práticas entre os pares, ou seja, entre cursos com características similares. Como destacado por Piran, Lacerda e Camargo (2018), para que uma análise de *benchmarking* seja eficaz, é importante que as unidades a serem comparadas sejam relativamente similares.

Neste contexto, observa-se a importância de estudar as similaridades entre os cursos de graduação.

Diante do exposto, formulou-se a seguinte questão de pesquisa: tomando como parâmetro as oito componentes utilizadas para construir o Conceito Preliminar de Curso, verificam-se padrões de desempenho similares ou dissimilares suficientemente significativos para permitir afirmar a existência de agrupamentos naturais entre os cursos de graduação em Matemática?

Com base na questão de pesquisa, foi definido o objetivo geral deste trabalho.

1.1 OBJETIVOS

O objetivo geral deste trabalho é analisar a existência de agrupamentos entre os cursos de graduação em Matemática, tendo como referência as variáveis que são utilizadas para a construção do Conceito Preliminar de Curso.

Para atender ao objetivo geral, foram elaborados os seguintes objetivos específicos:

- Realizar uma análise descritiva e exploratória das notas dos cursos de Matemática nas componentes do Conceito Preliminar de Curso;
- Verificar a existência de agrupamentos entre os cursos de graduação em Matemática por meio de um algoritmo de agrupamento;
- Caracterizar os agrupamentos obtidos por meio da identificação das similaridades (características comuns) dentro dos *clusters* e dissimilaridade entre os *clusters*;
- Analisar comparativamente as notas do curso de Matemática da UFSC nas componentes do CPC em relação aos demais cursos do grupo ao qual o curso foi classificado.

1.2 ESTRUTURA DO TRABALHO

Este trabalho está organizado em 5 capítulos. Este primeiro capítulo teve como propósito contextualizar o tema a ser estudado, apresentar a motivação para realização do trabalho e os objetivos do mesmo. No segundo capítulo serão

desenvolvidos os conceitos mais relevantes para o desenvolvimento do trabalho. No terceiro capítulo será apresentada a metodologia a ser seguida para este estudo. No quarto capítulo serão apresentados os resultados do agrupamento aplicado e as análises decorrentes dele. Por fim, o capítulo 5 traz as principais conclusões do estudo e sugestões para trabalhos futuros.

2 REFERENCIAL TEÓRICO

2.1 A AVALIAÇÃO DA EDUCAÇÃO SUPERIOR BRASILEIRA

2.1.1 **Contextualização histórica: dos primeiros esforços para avaliar a Educação Superior até os dias atuais**

Durante o governo militar no país, de 1964 a 1985, a educação foi submetida a um modelo “produtivista” de eficiência. Neste período, em 1968, foi proposta uma Reforma Universitária que resultou em um aumento no número de instituições e matrículas no ensino superior, sem democratização do acesso e garantia de permanência e de qualidade. Isso gerou uma “crise” nas universidades (Zainko, 2008).

No fim do governo militar, em junho de 1983, foi criado o Programa de Avaliação da Reforma Universitária (PARU). O PARU foi desenvolvido por incentivo do Conselho Federal de Educação (CFE) como consequência de discussões internas sobre as greves nas universidades federais (Barreyro e Rothen, 2008). Este programa tinha o propósito de avaliar o sistema de educação superior a fim de conhecer as verdadeiras condições deste local de produção e disseminação de conhecimento. A primeira fase desta pesquisa consistia num “estudo base” para coletar informações essenciais sobre as Instituições de Educação Superior (IES), o que criava um instrumento de comparação entre as instituições, e na segunda fase ocorriam “estudos de casos” com análises específicas.

Apesar de ter sido desativado em 1984 sem apresentar os resultados, Barreyro e Rothen (2008) destacam ainda que o PARU deixou de herança a ideia de uma avaliação formativa e emancipatória e de uma atenção à Gestão das Instituições de Educação Superior.

Segundo Zainko (2008), após o fim do PARU surgiu a necessidade de ter novamente um tipo de avaliação para controle da educação superior, então foi criada a Comissão Nacional de Reformulação da Educação Superior em 1985 pelo governo de José Sarney, atendendo demandas da comunidade universitária.

A comissão criou um relatório que era constituído de cinco partes: 1) Os princípios norteadores da proposta de avaliação; 2) A função do Conselho Federal de

Educação, de ser responsável pela realização das avaliações; 3) Autonomia e avaliação, que tratava sobre autonomia das universidades e formas de avaliar; 4) Acreditação, tratando sobre o controle de funcionamento das faculdades; 5) A lógica do documento (Barreyro e Rothen, 2008).

As partes deste relatório que tratavam diretamente sobre as avaliações no Ensino Superior eram a 2) e 3), que destacavam que o Conselho Federal de Educação deveria ser o órgão responsável por avaliar as instituições de Ensino Superior e depois identificar as condições de cada universidade, em que o alto padrão de qualidade seria a permissão da sociedade para uma maior autonomia da instituição. Autonomia essa na gestão financeira, escolha de currículo dos cursos e pesquisas desenvolvidas.

No ano de 1986 foi criado o Grupo Executivo da Reforma da Educação Superior (GERES), constituído por pessoas do Ministério da Educação. O GERES tinha a responsabilidade de elaborar uma proposta de reforma universitária e para isso usou o relatório da Comissão Nacional de Reformulação da Educação Superior (Barreyro e Rothen, 2008).

Segundo Zainko (2008), essa nova proposta na verdade foi uma reformulação do relatório da Comissão Nacional de Reformulação da Educação Superior, pois o relatório inicial produzido pela Comissão não ia ao encontro das expectativas do Estado quanto à forma de controle da educação superior.

Dentro deste novo relatório, Barreyro e Rothen (2008) trazem que a avaliação da educação superior deveria ser responsabilidade da Secretaria da Educação Superior, e não mais do Conselho Federal de Educação (CFE), mas que os resultados continuariam a ser usufruídos pelo CFE. Trazem também que a avaliação e o financiamento das instituições deveriam ter relação direta: Um recurso mínimo seria assegurado para as instituições federais, porém parte do financiamento dependeria dos resultados das avaliações de desempenho. Além disso, são essas avaliações que controlariam a qualidade da Educação Superior, tanto para as universidades, como para estruturas com outros modelos.

Ao ser divulgado, esse projeto novo foi muito criticado por professores, estudantes e funcionários, por temerem que o governo não fosse mais obrigado a financiar a educação superior.

Em 1993 foi criada a Comissão Nacional de Avaliação das Universidades Brasileiras (chamado de PAIUB), pela Secretaria de Educação Superior, para definir diretrizes e propiciar a implementação da avaliação institucional nas universidades do Brasil. O documento resultante da Comissão definia que a avaliação tinha o objetivo de mostrar um processo de aprimoramento do desempenho acadêmico, que era um dever das universidades, de ser usada como recurso para planejamento e gestão da universidade, e que seria também como uma prestação de contas frente à sociedade (Barreyro e Rothen, 2008).

Esses objetivos são uma mescla dos relatórios apresentados pelas grandes organizações que até então existiram: o PARU, a Comissão Nacional de Reformulação da Educação Superior e o GERES.

Em 1995, o modelo de avaliação do PAIUB passou a não ser considerado suficiente para o Ministério da Educação, e então foi implementado o Exame Nacional de Cursos, conhecido popularmente como Provão, constituindo de provas nacionais sobre os conteúdos mínimos de cada curso, a ser realizada anualmente pelos estudantes em fase de conclusão (Zainko, 2008).

Dez anos depois da criação do PAIUB, durante o governo de Lula em 2003, foi criada a Comissão Especial de Avaliação (CEA) a fim de desenvolver uma proposta de avaliação da educação superior (Barreyro e Rothen, 2014).

Depois de discussões de diferentes propostas, em dezembro de 2003 foi instituído o Sistema Nacional de Avaliação e Progresso do Ensino Superior (SINAPES), em que previa a criação de duas comissões nacionais de avaliação com funções de execução e consulta e revogava a existência do Provão (Barreyro e Rothen, 2014). Previa também a criação das Comissões Próprias de Avaliação (CPA) dentro de cada instituição, e a regulamentação da avaliação ficava a cargo do MEC.

Finalmente em 2004 foi promulgada a lei que instituiu o Sistema Nacional de Avaliação da Educação Superior (SINAES) que, como Barreyro e Rothen (2014) explicam, é estruturado por três grandes partes: a avaliação institucional, composta pela autoavaliação e a avaliação externa; a avaliação de cursos, por meio de visitas *in loco* de avaliadores externos; a avaliação dos estudantes, pelo Exame Nacional de Desempenho dos Estudantes (Enade). Manteve-se as CPAs de cada instituição com

autonomia para gerir o processo e criou-se a Comissão Nacional de Avaliação da Educação Superior (CONAES), para cuidar desse sistema.

O SINAES foi instituído pela Lei nº 10.861, de 2004 (BRASIL, 2004). Esta lei é a que rege a avaliação da educação superior até os dias atuais. Ao longo do tempo e vendo os resultados na prática, ela passou por algumas mudanças. A principal é a criação do Conceito Preliminar de Cursos, em 2008, que leva em conta dados advindos do Enade, o Censo da Educação Superior, entre outros e produz um conceito de 1 a 5.

2.1.2 O SINAES

O Sistema Nacional de Avaliação da Educação Superior (SINAES) é, como o próprio nome diz, um sistema que envolve toda a nação, e por isso ele deve envolver todas as instituições de ensino superior, tanto no âmbito público como no privado (Ristoff e Giolo, 2006). Neste sentido, a avaliação toda deve ter caráter formativo e construtivo, voltada à efetividade científica e social, e não deve ter o foco apenas em medir, classificar e gerar *rankings* sobre a produtividade e nível de conhecimento das instituições. Mais do que isso, em cada fase avaliativa deve-se reconhecer o meio e as condições reais em que aquela instituição e comunidade envolvida estão inseridas (BRASIL, 2003). Dessa forma, o SINAES deve cumprir com seus objetivos apresentados pela lei n.10.861/2004 de melhorar a qualidade da educação superior.

Como mencionado anteriormente, o SINAES possui três grandes pilares na sua estrutura. Um dos pilares é a avaliação institucional, composta pela autoavaliação e a avaliação externa, o outro é a avaliação de cursos, e o terceiro é a avaliação dos estudantes, por meio do Exame Nacional de Desempenho dos Estudantes (Enade).

O processo de avaliação começa na *avaliação institucional*. Na primeira etapa, cada IES realiza uma autoavaliação organizada pela Comissão Própria de Avaliação (CPA), que é uma comissão interna que tem a atribuição de conduzir os processos de autoavaliação das instituições. Numa segunda etapa, acontece a avaliação externa, por meio da visita de avaliadores do Inep.

O resultado dessa fase da avaliação deve mostrar, além dos dados de aprendizado, a possibilidade de identificação de problemas que existem e também as potencialidades que a instituição oferece. Deve ser avaliado também a responsabilidade social dentro da concepção de formação pelo Projeto Pedagógico Institucional e a responsabilidade pública com relação a gestão e infraestrutura (BRASIL, 2003).

Pela lei n. 10.861/2004, a avaliação do *desempenho dos estudantes* deve ser realizada por meio do Enade. Este exame é efetuado a cada três anos e deve analisar o desempenho dos estudantes com relação aos conteúdos programáticos de cada curso, além de comparar os conhecimentos dos alunos ingressantes e dos concluintes. Lacerda e Ferri (2017) explicam que este exame é composto por uma prova de formação geral, com peso de 75%, e uma de componente específico, com peso de 25% da nota.

De acordo com a portaria n. 2051/2004 (BRASIL, 2004), a *avaliação dos Cursos de Graduação* deve ser feita por uma Comissão Externa de Avaliação de Cursos, nomeada pelo Inep, que tem foco em três dimensões: 1) Organização didático-pedagógica, 2) Corpo docente e tutorial e 3) Infraestrutura. Dentro dessas dimensões, deve-se observar os seguintes aspectos:

I - O perfil do corpo docente;

II - As condições das instalações físicas;

III - A organização didático-pedagógica;

IV - O desempenho dos estudantes da IES no Enade;

V - Os dados do questionário socioeconômico preenchido pelos estudantes, disponíveis no momento da avaliação;

VI - Os dados atualizados do Censo da Educação Superior e do Cadastro Geral das Instituições e Cursos; e

VII - Outros considerados pertinentes pela CONAES.

Esta etapa da avaliação, bem como a etapa de avaliação do desempenho dos estudantes, está diretamente ligada ao Conceito Preliminar de Curso (CPC), pois a

partir delas que são determinadas as notas usadas para o cálculo do CPC. Como o CPC é objeto de estudo do presente trabalho, será apresentado em detalhe na seção seguinte.

2.1.3 Conceito preliminar de curso (CPC)

Criado em 2008, o Conceito Preliminar de Curso é considerado oficialmente o indicador de qualidade da educação superior no país. São consideradas as seguintes dimensões para fazer seu cálculo: o desempenho dos estudantes, o valor agregado pelo processo formativo do curso, indicadores do corpo docente e as condições ofertadas para o desenvolvimento do processo formativo (Ikuta, 2016).

O cálculo do CPC resulta em conceitos de 1 a 5 em que maior ou igual a 3 é considerado satisfatório e menos que 3 é insatisfatório. Essa classificação pelo conceito passou a permitir que instituições com nota satisfatória não necessitem da visita *in loco*, ou seja, ela só é obrigatória para instituições com CPC menor que três (Ikuta, 2016).

Para a fórmula do CPC, as notas usadas são padronizadas pelo recurso estatístico de afastamento padronizado e posteriormente são reescaladas para que variem em uma escala entre 0 e 5, e assim é feita então uma soma ponderada com os componentes e pesos descritos na Tabela 1 (BRASIL, 2020):

Tabela 1 – Composição do CPC e pesos das suas dimensões e componentes

Dimensão	Componentes	Pesos	
Desempenho dos Estudantes	Nota dos Concluintes no Enade (NC)	20%	
Valor Agregado pelo Processo Formativo Oferecido pelo Curso	Nota do Indicador da Diferença entre os Desempenhos Observado e Esperado (NIDD)	35%	
Corpo Docente	Nota de Proporção de Mestres (NM)	7,5%	30%
	Nota de Proporção de Doutores (ND)	15%	
	Nota de Regime de Trabalho (NR)	7,5%	
Percepção Discente sobre as Condições do Processo Formativo	Nota referente à Organização Didático-Pedagógica (NO)	7,5%	15%
	Nota referente à Infraestrutura e Instalações Fiscais (NF)	5,0%	
	Nota referente às Oportunidades de Ampliação da Formação Acadêmica e Profissional (NA)	2,5%	

Fonte: BRASIL (2020, p.10)

A primeira dimensão tem a ver com o desempenho dos estudantes sendo 20% do peso atribuído a ela no indicador, e é dada a partir da Nota dos Concluintes no Enade (NC).

A segunda dimensão tem como única componente a Nota do Indicador de Diferença entre o Observado e o Esperado (NIDD), isto é, ela compara o desempenho dos concluintes do curso pelo Enade com as características dos estudantes ao ingressar no curso pelo ENEM com o objetivo de verificar o valor agregado pelo processo formativo dos alunos oferecido pelo curso, e esta dimensão tem peso de 35% no indicador.

A terceira dimensão é sobre o Corpo Docente em que 7,5% do peso da nota é atribuído à Nota de Proporção de Mestres (NM) em relação a todos os professores,

15% à Nota de Proporção de Doutores e 7,5% à Nota de Regime de Trabalho (NR) totalizando 30% no indicador.

A quarta dimensão se trata da Percepção Discente sobre o Processo Formativo e observa-se que 7,5% do peso é atribuído à Nota Referente à Organização Didático-Pedagógica (NO), 5% à Nota Referente à Infraestrutura e Instalações Físicas (NF), e 2,5% à Nota referente às Oportunidades de Ampliação da Formação Acadêmica e Profissional (NA), totalizando os últimos 15% de peso na nota do CPC.

Dessa forma, a fórmula do cálculo da Nota Contínua do CPC é a seguinte:

$$NCPC = 0,2NC + 0,35NIDD + 0,075NM + 0,15ND + 0,075NR \quad (i) \\ + 0,075NO + 0,05NF + 0,025NA$$

A Nota dos Concluintes do Enade (NC) é determinada a partir dos resultados deste exame realizado pelos alunos concluintes dos cursos e é padronizada e transformada para uma escala de 0 a 5, resultando na nota do curso.

A nota deste exame também faz parte da construção da componente Nota do Indicador de Diferença entre Desempenho Observado e o Esperado (NIDD), que leva em conta o número de estudantes concluintes participantes no Enade com resultados válidos, o desempenho geral dos estudantes participantes do exame, o desempenho dos estudantes no ENEM e o número de participantes no Enade com nota do ENEM recuperada.

As notas referentes ao Corpo Docente são obtidas do Censo da Educação Superior: as notas de Proporção de Mestres (NM) e Doutores (ND) são dadas pela proporção entre docentes com titulação igual ou superior a Mestres e a Doutores, em relação ao número total de docentes do curso; a nota de Regime de Trabalho (NR) vem da proporção de docentes com regime de trabalho parcial ou integral em relação ao número total de docentes do curso.

As demais notas referentes a Organização Didático-Pedagógica (NO), Infraestrutura (NI) e de Oportunidades de Ampliação da Formação (NA) são retiradas de um questionário socioeconômico respondido pelos estudantes que realizam o Enade.

Apesar do CPC ser considerado o principal indicador da qualidade da educação superior, ainda existem discussões acerca do modo de avaliação desse conceito. Não há dúvidas de que avaliações são necessárias, mas avaliar diferentes instituições com os pesos atribuídos a cada nota pode prejudicar as que têm diferentes focos, especializações e organizações (Zanella e Oliveira, 2020).

A partir dessa indagação, existem pesquisas que utilizam ferramentas na intenção de tornar o cálculo do CPC mais “pessoal” para cada instituição de acordo com suas especializações. Por exemplo, Zanella e Oliveira (2020) que calcularam o Indicador Compósito a partir da técnica de Análise Envoltória de Dados. Neste caso, o indicador de desempenho valoriza o que a instituição tem de melhor. Esta técnica é interessante, pois identifica as potencialidades e as fraquezas de cada IES, o que pode ser usado a favor de um dos objetivos do SINAES de ser uma avaliação formativa.

Outro exemplo da relação do CPC com diferentes realidades é no caso da Educação à Distância. Como Oliveira e Piconez (2017) analisaram em uma pesquisa feita com dados de 2015, a maioria dos cursos EaD tiveram conceito 3, o que pode significar que estão fazendo o mínimo para não precisar das avaliações externas *in loco*; além disso, ao comparar os CPCs desta modalidade com os dos cursos presenciais, a distribuição dos conceitos se apresenta de maneira bastante diferente.

2.2 ANÁLISE DE *CLUSTERS*

A análise de agrupamento, também chamada de *clustering*, se trata de uma coleção de técnicas em que o principal objetivo é separar objetos em grupos, de acordo com suas características, unindo objetos similares ou próximos e separando os diferentes ou distantes, seguindo apenas um critério de similaridade pré-determinado (Linden, 2009).

Essas técnicas são aplicadas em várias áreas, como: medicina, engenharia, marketing, biologia, bioinformática, etc. A ideia é agrupar os dados sem a necessidade de conhecimentos prévios sobre o assunto e que mantenham características comuns dentro dos grupos e distintas em grupos diferentes. Na bioinformática, por exemplo,

pode ser usada para classificar genes de acordo com seus processos de atuação (Faceli, 2007). A clusterização também pode ser aplicada no esporte, gerando uma classificação quanto ao potencial do atleta de acordo com dados físicos e de desempenho, como na pesquisa em que Paes, Uezu, Marcelo e Bohme (2008) realizaram com jovens atletas do projeto Esporte Talento (USP) classificando-os em melhores, médios e piores na qual teve uma concordância de 94,74% com a classificação subjetiva dos técnicos dos esportistas pesquisados.

Os algoritmos de agrupamento podem ser classificados em dois grupos: os hierárquicos e não-hierárquicos.

2.2.1 Hierárquicos

Os procedimentos hierárquicos de agrupamentos possuem uma estrutura do tipo árvore, ou dendrograma, e são divididos em aglomerativos ou divisivos. Nos métodos aglomerativos, inicia-se pelos indivíduos de forma separada que a cada passo se unem a outros pela similaridade ou proximidade, até que se forme um único grupo com todos os elementos. Nos métodos divisivos, começa-se por um grupo com todos os indivíduos que a cada passo se divide em subgrupos considerando as diferenças ou distâncias entre os elementos (Hair, Black, Babin, Anderson e Tatham, 2009).

De acordo com Hair *et al.* (2009), pode-se dizer que os métodos divisivos funcionam quase como os métodos aglomerativos ao contrário. Dessa forma, pode ser explorado o procedimento aglomerativo para compreensão de um procedimento hierárquico, que tem um passo a passo simples:

a. Cada objeto/indivíduo faz parte de seu próprio grupo unitário, ou seja, o número de agrupamentos é igual ao número de objetos.

b. A primeira ação é unir os dois agrupamentos mais similares, de acordo com a medida de similaridade determinada; resultando num número de agrupamentos uma unidade menor que anteriormente.

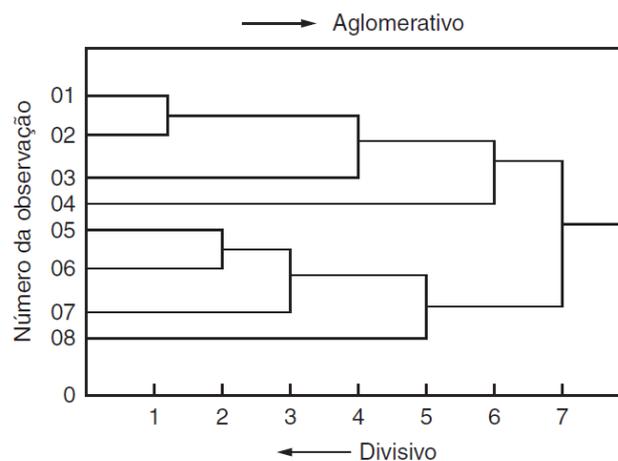
c. Em seguida o processo se repete, novamente de acordo com a medida de similaridade, unindo os dois grupos mais próximos e diminuindo o número total de agrupamentos.

d. Esse processo segue até que todos os objetos façam parte do mesmo grupo, ou seja, ele será aplicado $n - 1$ vezes, sendo n o número total de objetos.

Exemplo:

Tem-se 8 objetos. Isso significa que inicialmente são 8 agrupamentos com 1 indivíduo em cada. No primeiro passo, são unidos os dois grupos mais similares, restando 7 grupos. No segundo passo, também são unidos os dois grupos mais parecidos, resultando em 6 grupos no total. Assim segue sucessivamente por 7 passos até restar 1 grupo com 8 objetos, conforme Figura 1.

Figura 1 – Dendrograma que ilustra o agrupamento hierárquico



Fonte: Hair, Black, Babin, Anderson e Tatham, 2009

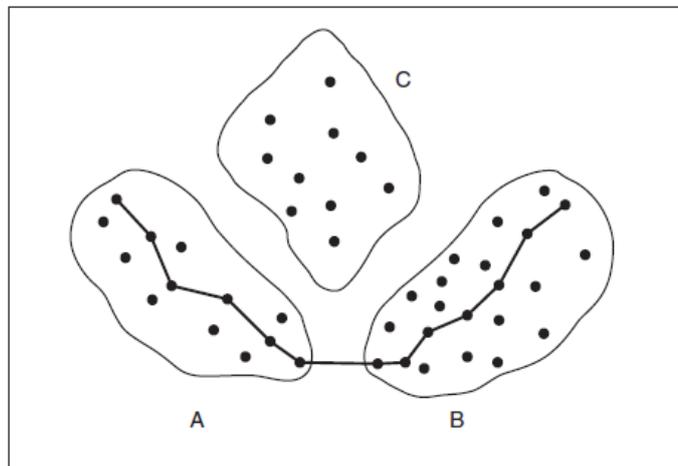
Para medir a similaridade entre agrupamentos com mais de um membro, existem diversas possibilidades. Hair *et al.* (2009) lista e explica os cinco algoritmos aglomerativos mais populares, que serão apresentados a seguir.

2.2.1.1 Ligação simples

Este método determina a similaridade entre agrupamentos como a menor distância entre qualquer indivíduo de um grupo e qualquer indivíduo do outro. Por isso também é chamado de **método do vizinho mais próximo**. O procedimento consiste em determinar todas as distâncias entre os objetos de agrupamentos diferentes e escolher a menor como medida de similaridade.

A vantagem é que esta técnica é bastante versátil, pois pode definir diversos padrões de aglomeração. Um exemplo citado no livro *Análise Multivariada de Dados* é o de agrupamentos em círculos concêntricos, como os anéis de um alvo. No entanto, a desvantagem aparece quando os agrupamentos são mal delineados, pois podem criar longas cadeias em que os objetos dos extremos são muito distantes, resultando em grupos que a homogeneidade interna fica “estranha”. A Figura 2 mostra que a ligação simples fará com que os agrupamentos A e B se unam, mas pensando na homogeneidade interna, faria mais sentido unir C com A ou com B.

Figura 2 – Exemplo de ligação simples



Fonte: Hair, Black, Babin, Anderson e Tatham, 2009

2.2.1.2 Ligação completa

Seguindo a mesma ideia da ligação simples, este método também compara todas as distâncias entre indivíduos de aglomerados distintos, porém é escolhida a maior distância como medida de similaridade. Por esse motivo, este método também

é chamado de **método do vizinho mais distante**, ou ainda, **método do diâmetro**. O segundo nome é pela ideia de que a similaridade entre agrupamentos pode ser visualizada como a menor esfera que inclui todos os objetos dos dois agrupamentos, ou seja, os indivíduos de um mesmo agrupamento estão ligados uns aos outros por no máximo um diâmetro de distância.

Dessa forma, o problema da ligação simples está resolvido, pois essa técnica não gera aquele tipo de encadeamento citado, por isso muitos pesquisadores consideram a mais apropriada para diversas aplicações.

2.2.1.3 Ligação média

O processo desse método consiste em tomar todas as distâncias entre indivíduos de agrupamentos distintos e utilizar a média dessas distâncias como medida de similaridade.

Esta técnica permite que as medidas não sejam tão afetadas por valores extremos e também não dependam de um único parâmetro, sendo quase como um meio-termo entre a ligação simples e a completa. Além disso, a tendência é que os agregados gerados tenham aproximadamente a mesma variação interna.

2.2.1.4 Método Centróide

Neste método, a medida de similaridade entre dois agrupamentos é a distância de seus centróides. O centróide de um agrupamento é o ponto resultante da média das distâncias dos indivíduos deste agrupamento.

Hair *et al.* (2009) destaca que a vantagem desse método é como a da ligação média, em que as distâncias não são tão afetadas por valores extremos. No entanto, este método pode criar soluções confusas, já que a cada união de dois agrupamentos o centróide do mesmo muda e, portanto, a distância de dois agrupamentos pode ser menor do que a distância de outros dois que foram unidos na fase anterior.

2.2.1.5 Método Ward

Para este método, a medida de similaridade é a soma dos quadrados das distâncias dentro dos agrupamentos feita sobre todas as variáveis. Em outras palavras, a cada passo serão combinados os dois agrupamentos que têm a menor soma total de quadrados entre todas as variáveis em todos os agrupamentos.

Esta técnica tende a associar agrupamentos com um número pequeno de indivíduos, já que a soma dos quadrados é totalmente influenciada pelo número de observações envolvidas; e tende também a gerar agrupamentos com aproximadamente o mesmo número de indivíduos. Esta segunda característica pode ser decisiva na escolha de método de clusterização, caso seja algo que o pesquisador deseja ter em seus resultados.

2.2.2 Não-hierárquicos

Diferente dos métodos hierárquicos, os não-hierárquicos não têm o processo de árvore e o agrupamento não é feito de dois em dois. Neste processo de clusterização, o agrupamento depende do número k de *clusters* total que se deseja formar ao final. Em Análise Multivariada de Dados (Hair, Black, Babin, Anderson, Tatham, Ronald; 2009) é apresentado o processo em dois passos:

- a. Primeiro deve-se determinar, em geral de forma aleatória, os pontos de partida, também chamados de **sementes de agrupamento**, para cada aglomerado.
- b. Em seguida, cada objeto é associado a uma semente de agrupamento de acordo com a similaridade. A medida de similaridade depende da abordagem escolhida.

As sementes de agrupamento podem ser escolhidas pelo pesquisador, que neste caso já deve ter algum conhecimento sobre os dados avaliados e suas similaridades, ou podem ser geradas pela amostra por critérios ou ferramentas de programas, e os pontos de partida podem ser mudados se o pesquisador julgar necessário ao observar o resultado dos agrupamentos e das médias dos mesmos.

2.2.2.1 k-médias

O método k-médias foi proposto de forma pioneira por S. Lloyd em 1982 (Porto, 2017) e é um dos não-hierárquicos mais conhecidos. É um método iterativo que tem como objetivo encontrar a melhor divisão de n dados em k agrupamentos, de forma que a distância total entre o centro de um grupo e seus respectivos dados, somada para todos os grupos, seja minimizada (Machado, 2011).

Segundo Machado (2011) e Porto (2017), considerando o conjunto de objetos $X = \{p_1, p_2, \dots, p_n\}$, sendo n o número de objetos a ser agrupado em k agrupamentos, o algoritmo inicia com a determinação dos k centros iniciais e então forma os k grupos seguindo uma medida de similaridade e associando cada objeto ao centro mais próximo, neste caso a medida de similaridade costuma ser a distância Euclidiana, formando os *clusters* X_i . Após o primeiro agrupamento, os centros são recalculados para que as distâncias entre ele e os objetos do grupo sejam as menores possíveis, em seguida a divisão é refeita com base nos novos centros e os objetos são rearranjados. Este processo deve ser feito até que os objetos se estabilizem e não mudem mais de agrupamentos, ou até que seja atingido um número pré-definido de iterações.

O centroide de cada *cluster* que Lloyd escolheu é o ponto que minimiza a soma do quadrado da distância Euclidiana entre ele e cada objeto do conjunto,

$$\arg \min_c \sum_{p_j \in X_i} |c - p_j|^2. \quad (\text{ii})$$

Esse ponto é justamente o centro de massa do agrupamento X_i e é dado por

$$c = \frac{p_1 + \dots + p_n}{|X_i|}, \quad (\text{iii})$$

em que $|X_i|$ corresponde a cardinalidade de X_i (Porto, 2017).

Agora que vimos um pouco sobre os diferentes métodos e seus tipos, podemos compreender quando Hair *et al.* (2009) diz que não existe uma resposta definitiva para a questão de quais métodos devem ser utilizados - hierárquicos ou não-hierárquicos-. A resposta é: depende.

Por um lado, os métodos hierárquicos têm a vantagem de ser mais simples por sua estrutura de árvore, podem ser aplicados a uma grande variedade de tipos de pesquisa, trazem resultados rápidos e um grande número de soluções de agrupamento, que vão desde os indivíduos separados até um único agrupamento que abrange todos eles. No entanto, eles podem sofrer grandes impactos com observações atípicas e podem se tornar inviáveis para tamanhos amostrais grandes. Por outro lado, os métodos não-hierárquicos são menos suscetíveis às observações anômalas e podem ser usados na clusterização de um grande número de dados. Em contrapartida, dependem muito da escolha inicial das sementes de agrupamentos e não são vantajosos quando se procura uma maior quantidade de soluções, pois geram apenas uma solução de exatamente k agrupamentos a cada análise feita.

2.3 ANÁLISE DE CLUSTER APLICADO NA EDUCAÇÃO SUPERIOR

A Análise de *Cluster* já foi aplicada no contexto da avaliação da educação superior. Fazendo uma busca na literatura com as palavras-chave deste trabalho - Análise de *cluster*, *k*-médias e Avaliação da educação superior -, podem ser encontrados estudos que envolvem esses assuntos. Seguem alguns exemplos.

O primeiro artigo é o “Classificação da Similaridade dos Cursos de Graduação da UFSM em Relação ao Índice do Aluno Equivalente – Uma Ferramenta de Gestão” em que Huppés, Souza e Ansuji (2016) tinham como objetivo classificar os cursos que apresentam maior e menor Nfte(G) dentro da Universidade Federal de Santa Maria e identificar alguma variável que possa sofrer alterações para melhorar esse índice. Este Nfte(G) é o indicador de aluno equivalente da graduação, utilizado na análise de custos de manutenção das Instituições Federais da Educação Superior e as variáveis incluem a duração média do curso, número de diplomados, número de matriculados, entre outros. Para isso, foi analisada uma amostra de 60 cursos de graduação, dentro dos oito Centros de Ensino, usando os dados entre os anos de 2010 e 2013.

Neste artigo foram usadas as ferramentas estatísticas de análise descritiva e a análise de *cluster*. Para a Análise de *Cluster*, foi utilizado o método Hierárquico de Ligação Simples, aquele que também é chamado de “método do vizinho mais próximo”. E assim foi feita a divisão em agrupamentos.

Foram feitas três análises de *clusters* diferentes e sempre comparando com a análise descritiva. Desta forma, foi possível identificar que a variável número de diplomados (Ndi) é a que impacta diretamente no índice Nfte(G) e, sabendo disso, a gestão pode investir em medidas didáticas e de incentivo para diminuir o número de desistências nos cursos.

Outro artigo encontrado na literatura é o “Análise dos Programas de Pós-graduação em Administração Reconhecidos pela CAPES: uma Abordagem Multivariada por Análise de *Cluster*”. Nele, Leite, Viana, Pereira e Mantovani (2007) buscaram agrupar os programas de pós-graduação em Administração de 2001 a 2003 a partir de algumas variáveis quantitativas medidas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e identificar quais variáveis são mais relevantes na composição do conceito dos programas. As variáveis foram selecionadas para que não houvessem dados faltantes e para que o número fosse coerente com a quantidade de programas pesquisados; então utilizou-se as seguintes: dimensão discente, titulados, tempo médio de titulação do mestrado acadêmico e publicação média do triênio.

Foi feito então o agrupamento utilizando o método não-hierárquico para gerar quatro *clusters*. Analisando os resultados, observou-se que a variável que mais diferencia os *clusters* foi a Publicação Média que inclusive contribuiu para a criação de um *cluster* com um único programa. Já a Dimensão Discente e a Tempo Médio de Titulação não apresentam grandes diferenças para os agrupamentos finais. A variável Titulados apresenta alguma variação entre os grupos formados.

Depois foi feito o agrupamento através do método hierárquico e foi considerado o resultado com 4 *clusters* para fazer uma comparação com os resultados anteriores. A partir dos resultados, observou-se que a Dimensão Discente não apresentou diferenças para os *clusters*, assim como no método não-hierárquico. A variável Titulados por sua vez apresentou uma diferença maior que antes. A Tempo Médio de Titulação apresentou uma diferença pequena, porém possível de notar entre os

agrupamentos. E a variável Publicação Média do Triênio, assim como no método não-hierárquico, é a que mostrou mais diferenciação entre os *clusters* obtidos. Leite, Viana, Pereira e Mantovani (2007) consideraram que o método hierárquico apresentou uma melhor distribuição dos programas do que o não-hierárquico.

Observando de forma geral, notou-se que a diferenciação dos grupos vem principalmente do nível de Publicação apresentado pelos programas. Assim, concluiu-se que esta pode ser a variável que possibilita uma melhora no conceito da avaliação da CAPES.

O próximo artigo é o “Proposição de Indicadores para o Corpo Discente e Análise de Agrupamentos Aplicada aos Cursos de Graduação da UFES”, em que Junior, Penholato, Erler e Carneiro (2013) propuseram indicadores de desempenho para o corpo discente para os cursos de graduação da Universidade Federal do Espírito Santo e, a partir da análise de agrupamentos, testaram o potencial desses indicadores como classificadores dos cursos de graduação desta instituição. Os indicadores propostos foram: demanda de candidato por vaga no vestibular, número de alunos ingressantes, matriculados e diplomados, índice de evasão, envolvimento do discente com pesquisa, índice de reprovação, conceito no Enade e o CPC.

Os indicadores foram padronizados para terem todos o mesmo peso no processo e então foi feito o agrupamento utilizando o método hierárquico de Ward. Analisando o dendrograma resultante, observaram a formação de quatro grandes grupos e assim foram agrupados os 63 cursos considerados.

Feito o agrupamento, observou-se que os cursos do Grupo 1 se destacaram por ter alta demanda no vestibular, alto número de matrículas, baixo índice de reprovação, pouco envolvimento dos discentes com pesquisa e baixo índice de evasão. O Grupo 2 se destacou por ter grande envolvimento dos discentes com pesquisa e índices moderados de demanda no vestibular e evasão. O Grupo 3 se destacou pelo baixo envolvimento com pesquisa. O Grupo 4 agrupou os cursos com maiores índices de reprovação e evasão, e concluíram ser o que mais demanda atenção da universidade. A heterogeneidade entre os clusters mostrou que os índices sugeridos são significativos na classificação dos cursos.

Juntamente com uma análise de correlação linear feita, observou-se que os indicadores estão relacionados entre si, com destaque para a relação entre a demanda no vestibular e o número de alunos ingressantes, matriculados e diplomados do curso, e também a relação entre os índices de reprovação e evasão. Junior *et al.* (2013) sugerem a busca de estratégias para melhorar o índice da demanda dos cursos no vestibular, por ser um indicador que impacta nos resultados dos cursos.

3 METODOLOGIA

De acordo com Gil (2008), os estudos descritivos são os que têm como objetivo descrever fatos e características de determinada população ou fenômeno, ou estabelecer relações entre variáveis. Levando em consideração os objetivos determinados, este trabalho é de caráter quantitativo e descritivo.

3.1 DESCRIÇÃO DOS DADOS UTILIZADOS NA PESQUISA

Para o desenvolvimento deste trabalho, foram utilizados os dados das oito componentes que formam o CPC, apresentadas na seção 2.1.3 Conceito Preliminar de Curso (CPC) deste trabalho. Foram escolhidos os dados do CPC do ano de 2017 divulgados pelo Ministério da Educação¹, pois foi a última vez que os cursos de Matemática passaram pela avaliação. Foram selecionados, então, os dados correspondentes aos cursos de Licenciatura em Matemática de instituições Públicas Federais na modalidade Presencial, para que tivesse relação com o contexto deste trabalho, justamente por ser um Trabalho de Conclusão de Curso em um curso presencial de Licenciatura em Matemática em uma Universidade Federal. Outra restrição feita, foi utilizar dados de instituições em que o número de concluintes inscritos fosse maior que 5, com o objetivo de aumentar a confiabilidade nos resultados da avaliação.

Com todas essas determinações, a base de dados é composta por 163 cursos, sendo 94 de Universidades e 69 de Institutos Federais. Todos esses cursos pertencem a um total de 78 Instituições de Educação Superior diferentes.

3.2 MÉTODOS APLICADOS PARA A ANÁLISE DOS DADOS

Num primeiro momento, foram utilizadas técnicas de análise exploratória e descritiva de dados para descrever e resumir os dados de uma maneira geral e, num

¹ Dados disponíveis em <http://portal.inep.gov.br/educacao-superior/indicadores-de-qualidade/resultados>

segundo momento após a clusterização, foram exploradas as características dos cursos dentro dos agrupamentos. Essas características também foram observadas na intenção de entender melhor quais variáveis foram mais significativas para formar os grupos.

Para o agrupamento dos dados, escolheu-se o método não-hierárquico k-médias. Ele foi escolhido por ser menos sensível a possíveis valores discrepantes e anomalias, como visto na subseção 2.2 do Referencial Teórico anteriormente, além de ser mais prático para a interpretação dos resultados do que um esquema de árvore hierárquica com o número de informações deste problema. O agrupamento dos cursos foi feito usando o programa SPSS versão 20 com o método k-médias. Este programa é um software que oferece um conjunto integrado de recursos de análise estatística e por isso ele também foi utilizado para a construção dos gráficos e tabelas apresentados ao longo dos resultados.

4 RESULTADOS

4.1 ANÁLISE EXPLORATÓRIA E DESCRITIVA DOS DADOS

Os dados utilizados nesse trabalho foram extraídos das variáveis que geraram o CPC, descritas anteriormente na seção 2.1.3, referem-se ao ano de 2017 e são divulgados pelo Inep. Inicialmente tomou-se as oito variáveis já transformadas em uma escala de 0 a 5. Na Tabela 2 são apresentadas as principais medidas descritivas dessas variáveis consideradas.

Tabela 2 – Principais medidas descritivas para as variáveis

	Mínimo	Máximo	Mediana	Média	Desvio Padrão
NC	0,39	5,00	2,40	2,50	0,86
NIDD	0,00	4,61	2,48	2,56	0,66
NO	0,77	4,43	2,89	2,84	0,81
NF	1,04	4,68	2,85	2,89	0,70
NA	1,21	4,52	3,06	3,03	0,67
NM	0,00	5,00	4,40	4,11	0,90
ND	0,00	5,00	2,57	2,46	1,45
NR	3,73	5,00	5,00	4,96	0,17

Fonte: Elaborada pela autora

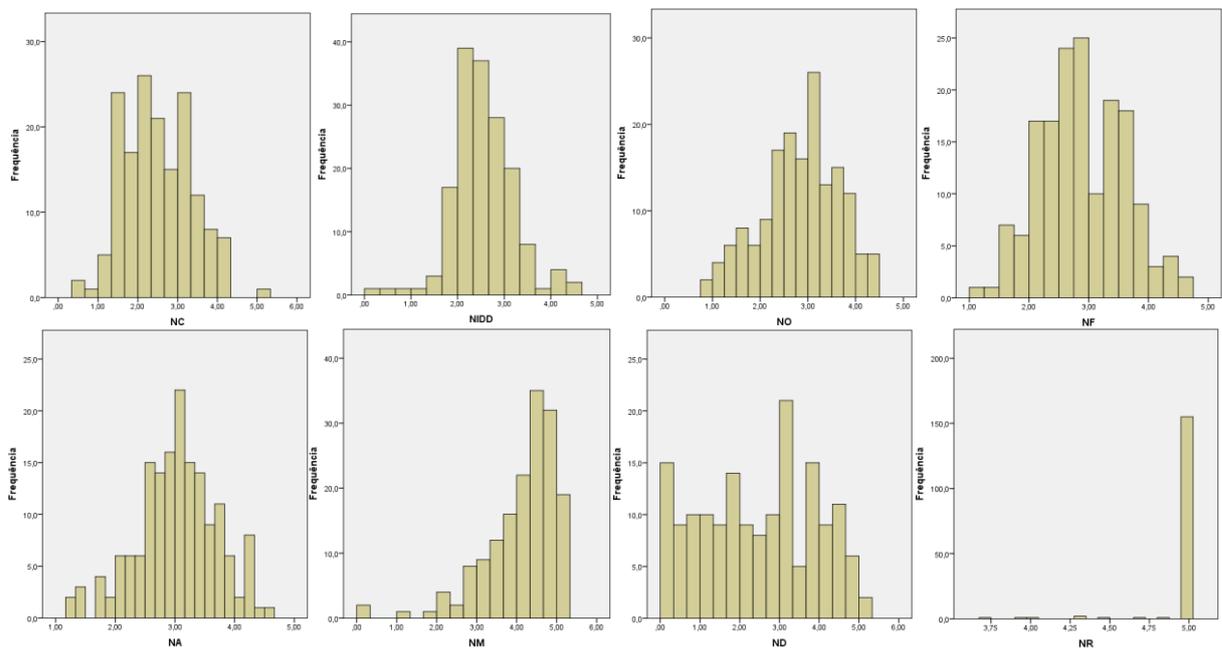
Observando os valores das medidas descritivas apresentadas na Tabela 2, é possível compreender um pouco os comportamentos dos dados. Pelo valor do desvio padrão, que é uma medida descritiva que capta o grau de afastamento dos dados em relação à média, é possível notar que as maiores discrepâncias de valores estão nas componentes relacionadas à Proporção de Doutores (ND) seguida pela Proporção de Mestres (NM). Já a nota referente ao Regime de Trabalho (NR) é a que possui a menor variação interna. Sobre essa componente, também é possível observar que os valores

da média e da mediana são bastante elevados, então compreende-se que grande parte dos cursos apresentaram nota do regime de trabalho próxima ou igual ao valor máximo. Todas as outras componentes possuem desvios-padrões com valores intermediários a altos, o que indica ainda uma heterogeneidade nas notas apresentadas pelos cursos de graduação considerados neste trabalho.

Olhando para os valores das médias e das medianas, percebe-se que a maioria das componentes apresentam valores nestas medidas em torno de 2,5 e 3,0, indicando uma menor assimetria na distribuição das notas dos cursos, com exceção da proporção de Mestres (NM) e o Regime de Trabalho (NR) já comentados anteriormente.

Além da tabela com as medidas descritivas, é possível gerar histogramas que refletem o comportamento de cada uma das variáveis. A Figura 3 apresenta os histogramas das oito componentes do CPC para os 163 cursos considerados.

Figura 3 – Histogramas das variáveis NC, NIDD, NO, NF, NA, NM, ND e NR, respectivamente



Fonte: Elaborada pela autora

Analisando os histogramas das variáveis, pode-se observar que as variáveis das notas dos Concluintes no Enade (NC), da Diferença de Desempenhos Observado e Esperado (NIDD), da Organização Didático Pedagógica (NO), da Infraestrutura (NF)

e das Oportunidades de Ampliação (NA), ou seja, a maioria das variáveis, tendem a ter grande parte dos valores centralizados e não têm assimetrias muito bruscas. As informações se complementam ao comparar com as médias e medianas apresentadas pela Tabela 2. A nota referente à Proporção de Mestres (NM) possui uma assimetria à esquerda, isso significa que a maioria dos cursos possui um grande número de docentes com titulação mínima de mestre, o que faz sentido já que estão sendo consideradas as instituições públicas federais.

Pela imagem, nota-se que a componente Nota de Regime de Trabalho (NR) teve apenas oito cursos com nota diferente de 5. Como foram considerados apenas os cursos de universidades ou institutos federais, era esperado que esta variável apresentasse pouca ou nenhuma variabilidade. Por isso, essa variável será excluída da etapa de análise de agrupamento, pois não agrega informações para a análise e pode ser um fator de confusão.

4.2 ANÁLISE DE AGRUPAMENTO E CARACTERIZAÇÃO DOS GRUPOS

Na base de dados considerada, todas as variáveis (componentes do CPC) são disponibilizadas pelo Inep já transformadas em uma escala de 0 a 5, portanto, já estão na mesma unidade de medida. No entanto, as componentes têm variabilidades muito distintas, como pode-se observar pela medida do desvio-padrão mostrada na Tabela 2.

Ao conduzir a análise de agrupamento, em geral, variáveis com maior dispersão (ou seja, maiores desvios-padrão) têm maior impacto sobre o valor de similaridade final (Hair *et al.*, 2009). Portanto, antes de conduzir a análise de agrupamento, as componentes foram padronizadas de maneira que cada componente tenha média igual a zero e desvio-padrão igual a 1. Ou seja, para cada componente i , os valores que cada um dos cursos apresentou (x_{ij}) foram subtraídos da média da componente ($Média_i$) e, posteriormente, essa diferença foi dividida pelo desvio padrão da componente (DP_i):

$$z_{ij} = \frac{x_{ij} - Média_i}{DP_i} \quad (iv)$$

A partir das sete componentes padronizadas, decidiu-se fazer o agrupamento dos cursos pelo método não-hierárquico k-médias. Para utilização desse método, deve-se decidir o número k de *clusters* que se quer formar no final. Então, foram gerados alguns cenários utilizando métodos de agrupamentos hierárquicos, que possibilitam ver os agrupamentos em forma de árvore, a fim de ter um *insight* acerca do número k de *clusters* a serem formados pelo outro método. Os resultados destes testes mostraram que faria sentido dividir os cursos em 3 agrupamentos distintos e, portanto, o algoritmo k-médias seria aplicado com três *clusters* ($k = 3$). Assim, definido $k = 3$ clusters finais, $n = 163$ cursos a serem agrupados, aplicou-se o método k-médias.

Feito o agrupamento, tem-se os seguintes clusters apresentados na Tabela 3, em que cada curso está nomeado pela sigla da instituição de Educação Superior, o nome do município que se encontra e entre parêntesis está o código do curso no MEC. O *Cluster 1* resultou em 49 cursos, o *Cluster 2* em 66 cursos e o *Cluster 3* em 48 cursos.

Tabela 3 – Distribuição dos cursos nos clusters

cluster 1	cluster 2	cluster 3
FURG Rio Grande (1035)	IF Catarinense Camboriú (1103025)	IF Catarinense Sombrio (1103032)
IF SUL DE MINAS Inconfidentes (1117203)	IF Catarinense Concórdia (1103031)	IFBA Camaçari (1161404)
IFBA Salvador (1122743)	IF Catarinense Rio do Sul (1102795)	IFCE Cedro (85320)
IFES Vitória (108730)	IFAC Cruzeiro do Sul (5000487)	IFCE Juazeiro do Norte (84374)
IFMG Formiga (116820)	IFAL Maceió (1103556)	IFES Cachoeiro de Itapemirim (1103657)
IFRJ Nilópolis (102430)	IFAM Manaus (5000493)	IFFarroupilha Alegrete (1128936)
IFSP Bragança Paulista (1128375)	IFBA Barreiras (116730)	IFFarroupilha Júlio de Castilhos (1103185)
IFSP Guarulhos (122122)	IFBA Valença (123519)	IFFarroupilha Santa Rosa (1128938)
IFSP São Paulo (113490)	IFCE Canindé (1160979)	IFFarroupilha São Borja (1168573)
UFABC Santo André (1102402)	IFCE Crateús (5000138)	IFFluminense Campos dos Goytacazes (52442)
UFBA Salvador (13276)	IFCE Fortaleza (63383)	IFMA Buriticupu (1264336)
UFES Vitória (12806)	IFG Goiânia (1103461)	IFMA Caxias (1327336)
UFF Niterói (12702)	IFMA São Luís (21443)	IFMA Codó (1103914)
UFG Goiânia (1131697)	IFNMG Januária (99503)	IFMA Zé Doca (1160371)
UFG Goiânia (1304557)	IFNMG Salinas (1102641)	IFMG São João Evangelista (1142320)
UFJF Juiz de Fora (1166038)	IFPA Belém (48289)	IFMT Campo Novo do Parecis (114388)
UFLA Lavras (101556)	IFPI Angical do Piauí (1103910)	IFMT Juína (1160424)
UFMG Belo Horizonte (12969)	IFPI Teresina (63062)	IFPB Cajazeiras (1128096)
UFMS Campo Grande (15833)	IFRJ Volta Redonda (121013)	IFPB Campina Grande (1128030)
UFMS Três Lagoas (15865)	IFRN Mossoró (123787)	IFPE Pesqueira (100690)
UFMT Rondonópolis (34)	IFRN Natal (1123270)	IFPI Corrente (1103911)
UFPE Recife (13619)	IFRS Ibirubá (1156565)	IFPI Floriano (83929)
UFPEL Capão do Leão (15000)	IFS Aracaju (96913)	IFPI Piripiri (1103813)
UFPEL Pelotas (122746)	IFTO Paraíso do Tocantins (1106721)	IFPI Uruçuí (1103912)
UFPR Curitiba (12553)	UFAL Arapiraca (102152)	IFRJ Paracambi (1153801)
UFPR Curitiba (49478)	UFAL Maceió (107520)	IFRN Santa Cruz (1184450)
UFRGS Porto Alegre (13741)	UFAM Manaus (379)	IFRO Cacoal (1292902)
UFRJ Rio de Janeiro (14328)	UFBA Salvador (117052)	IFRO Vilhena (1182765)
UFRN Natal (312333)	UFC Fortaleza (38246)	IFRR Boa Vista (1152312)
UFRPE Recife (14505)	UFCG Cajazeiras (1152880)	IFRS Bento Gonçalves (116144)
UFRRJ Nova Iguaçu (96158)	UFCG Campina Grande (13448)	IFRS Caxias do Sul (1126144)
UFRRJ Seropédica (12918)	UFCG Cuité (99806)	IFSEMG Rio Pomba (113589)
UFSC Florianópolis (14233)	UFF Santo Antônio de Pádua (12713)	IFSP Araraquara (1129073)
UFSCAR São Carlos (624)	UFG Catalão (14121)	IFSP Birigui (1128360)
UFSCAR São Carlos (637)	UFGD Dourados (15868)	IFSP Campos do Jordão (1181042)

(continua)

Tabela 3 – Distribuição dos cursos nos clusters

cluster 1	cluster 2	cluster 3
UFSCAR Sorocaba (115086)	UFMA São Luís (11439)	IFSP Caraguatatuba (1128378)
UFSJ São João del Rei (48946)	UFMS Aquidauana (18382)	IFTO Palmas (123494)
UFSM Santa Maria (13873)	UFMT Cuiabá (26)	UFAC Rio Branco (11540)
UFSM Santa Maria (41069)	UFMT Pontal do Araguaia (18354)	UFES São Mateus (116882)
UFTM Uberaba (114556)	UFOPA Santarém (12039)	UFG Jataí (18954)
UFU Uberlândia (1428)	UFPA Abaetetuba (114853)	UFMS Corumbá (15862)
UFV Florestal (122358)	UFPA Belém (11999)	UFMS Paranaíba (52139)
UFV Viçosa (696)	UFPA Bragança (12033)	UFMS Ponta Porã (121792)
UNB Brasília (44372)	UFPA Breves (12044)	UFMT Sinop (100775)
UNIFAL-MG Alfenas (96955)	UFPA Cametá (12074)	UFU Ituiutaba (102944)
UNIFEI Itajubá (120389)	UFPA Castanhal (12035)	UNIPAMPA Itaqui (5000912)
UNIRIO Rio de Janeiro (1101770)	UFPB João Pessoa (13402)	UTFPR Cornélio Procópio (1127676)
UTFPR Curitiba (1128169)	UFPB Rio Tinto (99045)	UTFPR Toledo (1152637)
UTFPR Pato Branco (14542)	UFPE Caruaru (118100)	
	UFPI Parnaíba (102576)	
	UFPI Picos (102588)	
	UFPI Teresina (300520)	
	UFRB Amargosa (100412)	
	UFRN Caicó (12354)	
	UFS Itabaiana (95055)	
	UFS São Cristóvão (297)	
	UFT Araguaína (1105218)	
	UFT Arraias (17154)	
	UFT Palmas (1121464)	
	UFVJM Teófilo Otoni (100916)	
	UNB Brasília (161)	
	UNIFAP Macapá (17196)	
	UNIFESSPA Marabá (12037)	
	UNIPAMPA Bagé (104282)	
	UNIR Ji-Paraná (16010)	
	UNIR Porto Velho (16009)	

Nota: Cursos apresentados pela sigla da instituição, nome do município e código do curso no MEC

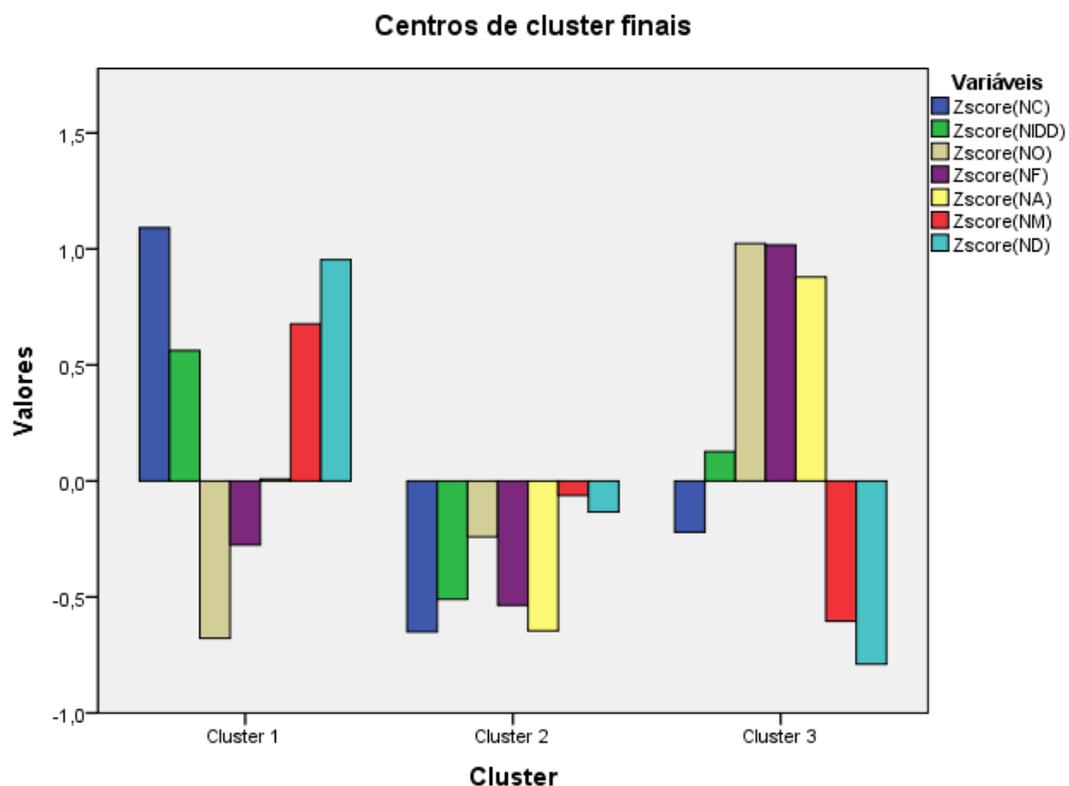
Fonte: Elaborada pela autora

(conclusão)

Um gráfico interessante de ser analisado é o que mostra os valores do centro final de cada uma das sete componentes avaliadas em cada um dos clusters. A Figura

4 apresenta este gráfico, em que cada cor representa uma componente diferente. Deve-se lembrar que os valores das componentes estão padronizadas, então a notação “Zscore(NC)” corresponde ao valor da Nota dos Concluintes no Enade (NC) na padronização descrita anteriormente, da mesma forma acontece para as outras componentes. Também por causa dessa padronização, a média é zero e, portanto, os valores próximos de zero estão perto da média, os valores positivos estão acima da média e os negativos abaixo.

Figura 4 – Componentes dos centros dos clusters finais



Fonte: Elaborada pela autora

Algumas observações interessantes podem ser feitas ao analisar esse gráfico. Como os valores dos centros dos *clusters* correspondem às médias dos valores dos objetos de cada agrupamento, eles dão uma noção muito boa das características dos cursos de cada *cluster*.

É possível notar que o *Cluster 2* é formado por cursos que possuem todas as notas observadas abaixo da média. As proporções de Mestres (NM) e de Doutores (ND) estão próximas à média geral, mas ainda se encontram abaixo dela. As notas dos Concluintes no Enade (NC) e a Diferença dos Desempenhos Observado e

Esperado (NIDD) estão muito abaixo da média geral, assim como as notas advindas das opiniões dos alunos sobre as Oportunidade de Ampliação da Formação Acadêmica (NA) e Infraestrutura e Instalações Físicas (NF).

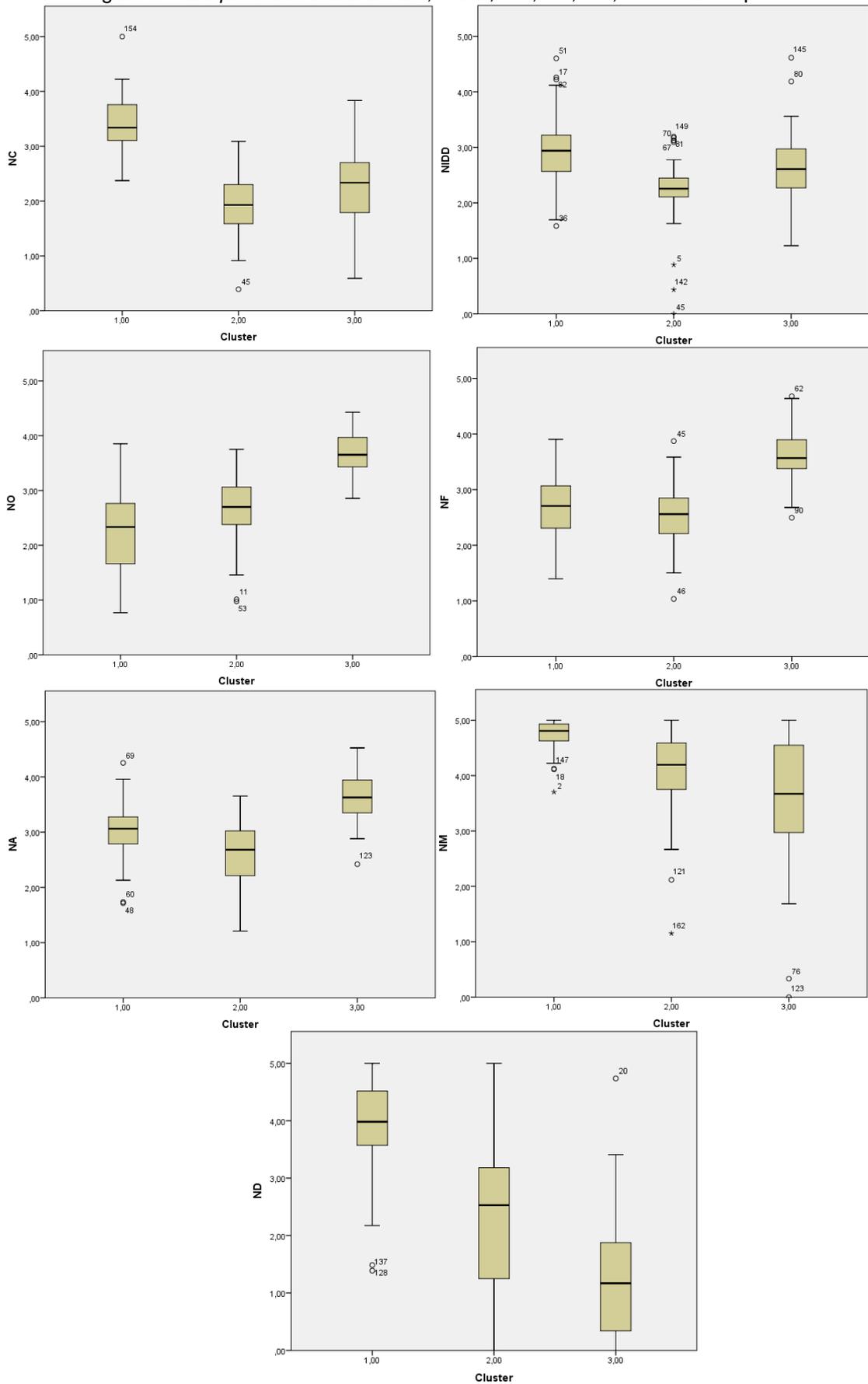
O *Cluster 1* é o que possui os cursos com maiores notas NC e NIDD. Ambas as componentes estão relacionadas aos alunos dos cursos de Matemática, seja pelo conhecimento dos alunos concluintes dos cursos, mensurado pelo exame Enade, seja pelo valor agregado pelo processo formativo oferecido pelo curso, mensurado pelo NIDD. Este *cluster* também é o que possui as maiores proporções de professores com titulação de Mestre e Doutor (NM e ND), na verdade, para essas duas componentes, é o único *cluster* que possui as médias das componentes maiores que as médias gerais. Em contrapartida, as médias referentes às componentes que têm a ver com a percepção do aluno sobre a organização didático-pedagógica, as oportunidades de ampliação da formação e a infraestrutura e instalações físicas (NO, NA e NF) estão abaixo da média geral, ou no máximo muito próximas à média.

O *Cluster 3* por sua vez, quase que ao contrário do *Cluster 1*, possui médias altas nas componentes que se referem à percepção do estudante (NO, NA e NF) sobre o curso, possui médias NC e NIDD, que se referem ao conhecimento adquirido abaixo ou bem próximas das médias gerais e apresentam as piores médias de proporções de Mestres e Doutores (NM e ND) em relação aos outros dois agrupamentos.

Agora serão usados diagramas de *boxplot* para fazer observações a respeito das características dos cursos pertencentes a cada cluster. Os diagramas de *boxplot* representam os valores típicos da distribuição, ou seja, o valor mínimo, o quartil 1, a mediana, o quartil 3 e o valor máximo observado.

A Figura 5 mostra os *boxplots* referentes aos valores de cada uma das sete componentes consideradas – no eixo vertical - em relação a cada um dos três *clusters* – no eixo horizontal -. Quando colocados lado a lado, os diagramas facilitam bastante a comparação dos valores típicos das componentes por *cluster*, complementando assim a informação comparativa das médias, apresentada na Figura 4.

Figura 5 – Boxplots das variáveis NC, NIDD, NO, NF, NA, NM e ND respectivamente



Fonte: Elaborada pela autora

Na variável Nota dos Concluintes no Enade (NC), pode-se observar que certamente o *cluster 1* agrupou cursos que tinham notas bem melhores do que nos demais cursos, inclusive pertence a ele o curso da UFABC de Santo André (denotado pelo número 154 na Figura 5) que aparece como um *outlier* (valor atípico na distribuição) por ser o único com a nota máxima 5,0; é possível observar que a mediana é aproximadamente 3,4. Por outro lado, os cursos dos *clusters 2* e *3* tiveram desempenhos razoáveis ou baixos no Enade em comparação com os cursos do *cluster 1*.

Na variável Nota do Indicador de Diferença entre os Desempenhos Observado e Esperado (NIDD), o comportamento é similar à componente anterior, com o *cluster 1* sendo o que obtém os cursos com as notas mais altas, mas nesse caso a mediana está em torno de 3,0. Os *clusters 2* e *3* também possuem cursos com desempenhos razoáveis ou baixos, em especial o *cluster 2* que contém os cursos da UNB de Brasília, UFT de Palmas e UFBA de Salvador com as piores notas. Ainda, vale observar que a amplitude entre o valor mínimo e máximo do *boxplot* no *cluster 2* é consideravelmente menor que a amplitude observada nos demais clusters, sugerindo que, excluindo os valores atípicos, os cursos deste *cluster* apresentaram valores de NIDD mais homogêneos do que os cursos dos demais agrupamentos.

Quando se trata da componente Nota de Organização Didático-Pedagógica (NO) o cenário muda. O *cluster 1* contém os cursos que receberam as piores avaliações dos alunos, o *cluster 2* teve avaliações medianas e o *cluster 3* concentrou os que tiveram as melhores notas nesse quesito. Nesta componente, a nota mínima observada no *cluster 3* é superior ao valor do quartil 3 do *cluster 1*, portanto, ao menos 75% dos cursos do *cluster 1* tiveram nota inferior à nota mínima observada nos cursos do *cluster 3*.

O comportamento dos *clusters* em relação à Nota Referente à Infraestrutura e Instalações Físicas (NF) é muito similar à componente NO, no entanto no *cluster 2* aparece um *outlier* com a pior nota entre todos os cursos sendo da UFPB de João Pessoa (curso denotado pelo número 46 na Figura 5).

Em relação à Nota Referente à Oportunidade de Ampliação da Formação Acadêmica (NA), os cursos com maiores notas também se concentraram no *cluster 3*,

porém, diferentemente das componentes anteriores, o *cluster 1* teve avaliações razoáveis e o *cluster 2* foi o que concentrou os cursos com as piores avaliações.

Observando o *boxplot* da Nota de Proporção de Mestres (NM), nota-se que o *cluster 1* agrupou quase que apenas cursos com nota maior que 4,0, com exceção do curso da UFTM de Rondonópolis (denotado pelo número 2 na Figura 5) que tem nota nesta componente em torno de 3,6. Os *clusters 2* e 3 também contém cursos com nota máxima, porém estão inclusos cursos superiores com notas intermediárias e baixas, inclusive o *cluster 3* contém o curso do IFPI de Uruçuí (denotado pelo número 123 na Figura 5) que teve nota igual a zero. Esta é a componente que apresentou as maiores discrepâncias na comparação entre os *clusters*. Pode-se observar que o valor do quartil 1 do *cluster 1* (aproximadamente 4,5) é superior aos valores do quartil 3 dos demais *clusters*, indicando que 75% dos cursos do primeiro *cluster* tiveram notas superiores a 75% dos cursos dos *cluster 2* e 3.

Sobre a Nota de Proporção de Doutores (ND), o *cluster 2* chama atenção, pois ele engloba toda a extensão de notas, o que significa que neste agrupamento é possível encontrar cursos em toda a faixa de valores de 0 a 5, então neste quesito é um *cluster* muito heterogêneo. Enquanto isso, no *cluster 3* observa-se que os 75% inferiores tiveram nota inferior a 2,0 e no *cluster 1* os 50% superiores tiveram nota superior a 4,0.

A partir dessas constatações, observa-se que uma característica muito forte do *cluster 1* é ter um grande número de professores com formação de mestre ou doutor, já o *cluster 3* é quase que o oposto com relação à formação de doutor.

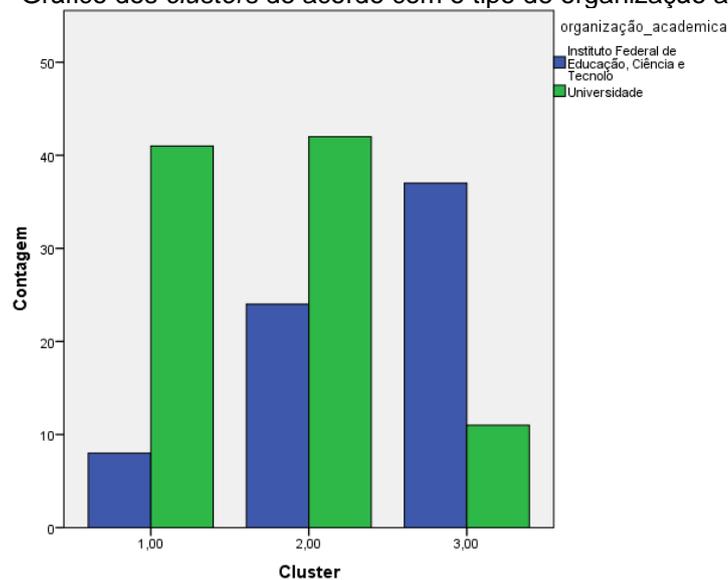
Outro fator que é possível notar, é que os cursos do *cluster 1* possuem notas altas nas componentes relacionadas ao desempenho do aluno (NC e NIDD), enquanto que as notas que dependem da percepção dos alunos sobre o curso e a instituição (NO, NF e NA) costumam ser as mais baixas comparadas aos outros cursos.

Quase que de uma forma antagônica, os cursos do *cluster 3* possuem as notas mais baixas quando se trata do desempenho dos alunos e da proporção de mestres e doutores, e receberam as melhores avaliações por parte dos alunos em relação à estrutura geral do curso.

No meio disso, o *cluster 2* reuniu os cursos com notas intermediárias em praticamente todas as componentes avaliadas, sem apresentar relação direta com as dimensões observadas.

Tirando o foco das variáveis, uma comparação que vale a pena fazer é sobre os tipos de organização acadêmica – Universidade ou Instituto Federal – dos cursos de cada *cluster*. A Figura 6 mostra o número de instituições de cada tipo de organização acadêmica dentro dos grupos.

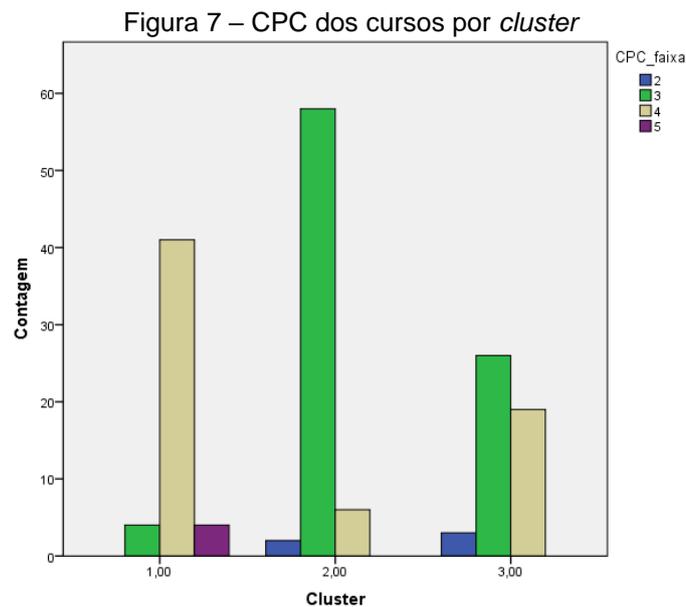
Figura 6 – Gráfico dos *clusters* de acordo com o tipo de organização acadêmica



Fonte: Elaborada pela autora

Nota-se que esse fator é marcante neste agrupamento, em especial para o *cluster 1* que é majoritariamente formado por cursos de universidades e inclui poucos institutos federais, e o *cluster 3* que é formado predominantemente por institutos federais. Já o *cluster 2* é um pouco mais heterogêneo neste aspecto.

Com todas essas observações, não poderia deixar de ser feita a comparação dos clusters com o Conceito Preliminar de Curso que cada curso recebeu. A Figura 7 mostra o gráfico da relação dos conceitos dos cursos de cada *cluster*.



Fonte: Elaborada pela autora

Com este gráfico, constata-se que o *cluster 1* é o que possui os cursos com os maiores CPCs, nele estão todos os cursos com conceito 5 e a maioria com conceito 4. Este resultado poderia ser esperado com as análises anteriores, pois este agrupamento apresenta notas boas em muitas variáveis. O *cluster 3*, por sua vez, possui um pouco mais de equilíbrio com cursos conceito 3 e 4, e alguns conceito 2. O *cluster 2* é o que possui o maior número de cursos que tiveram conceito 3, e possui poucos cursos com conceito 2 e 4. As informações detalhadas sobre os cursos de cada cluster, incluindo as notas das componentes do CPC e nota geral do CPC, podem ser encontradas no Anexo A.

É interessante observar que nenhum curso de Matemática considerado teve conceito 0 ou 1, apenas cinco tiveram conceito 2, a maioria – oitenta e oito – tiveram conceito 3, sessenta e seis tiveram conceito 4 e quatro tiveram conceito 5.

Apesar dessas comparações, é importante lembrar que a fórmula do cálculo do CPC inclui uma variável que foi descartada neste trabalho por apresentar variabilidade muito baixa. Além disso, cada componente deste cálculo possui um peso diferente, enquanto que a clusterização deste trabalho utilizou inclusive ferramentas de padronização na tentativa de que as variáveis tivessem o mesmo peso.

Essas similaridades reforçam a ideia das técnicas de agrupamentos e análise de *clusters*, pois mostram como dentro de um mesmo grupo existem as semelhanças e em grupos diferentes existem as diferenças. Além disso, essas técnicas aplicadas neste contexto contribuem para o objetivo do MEC de avaliar a educação, pois como os *clusters* agrupam cursos similares (relativamente homogêneos), facilita a identificação e observação dos exemplos de boas práticas entre cursos, com o propósito de promover melhorias no desempenho dos cursos.

4.3 ANÁLISE DO CURSO DE LICENCIATURA EM MATEMÁTICA DA UFSC FLORIANÓPOLIS

Como os dados foram escolhidos de acordo com o contexto de realização deste trabalho, é válido analisar as variáveis também para o curso de Licenciatura em Matemática da Universidade Federal de Santa Catarina do campus de Florianópolis. Para isso, a Figura 8 apresenta diagramas de dispersão que cruzam as variáveis dentro de cada dimensão, isto é, mostra as relações entre as variáveis NC e NIDD, entre NM e ND, entre NF e NO e entre NO e NA, com diferenciação dos *clusters* que estão destacados por cor: o azul corresponde ao *cluster* 1, o laranja ao 2 e o verde ao 3. Em cada um dos gráficos está destacado o curso da UFSC de Florianópolis com um círculo vermelho.

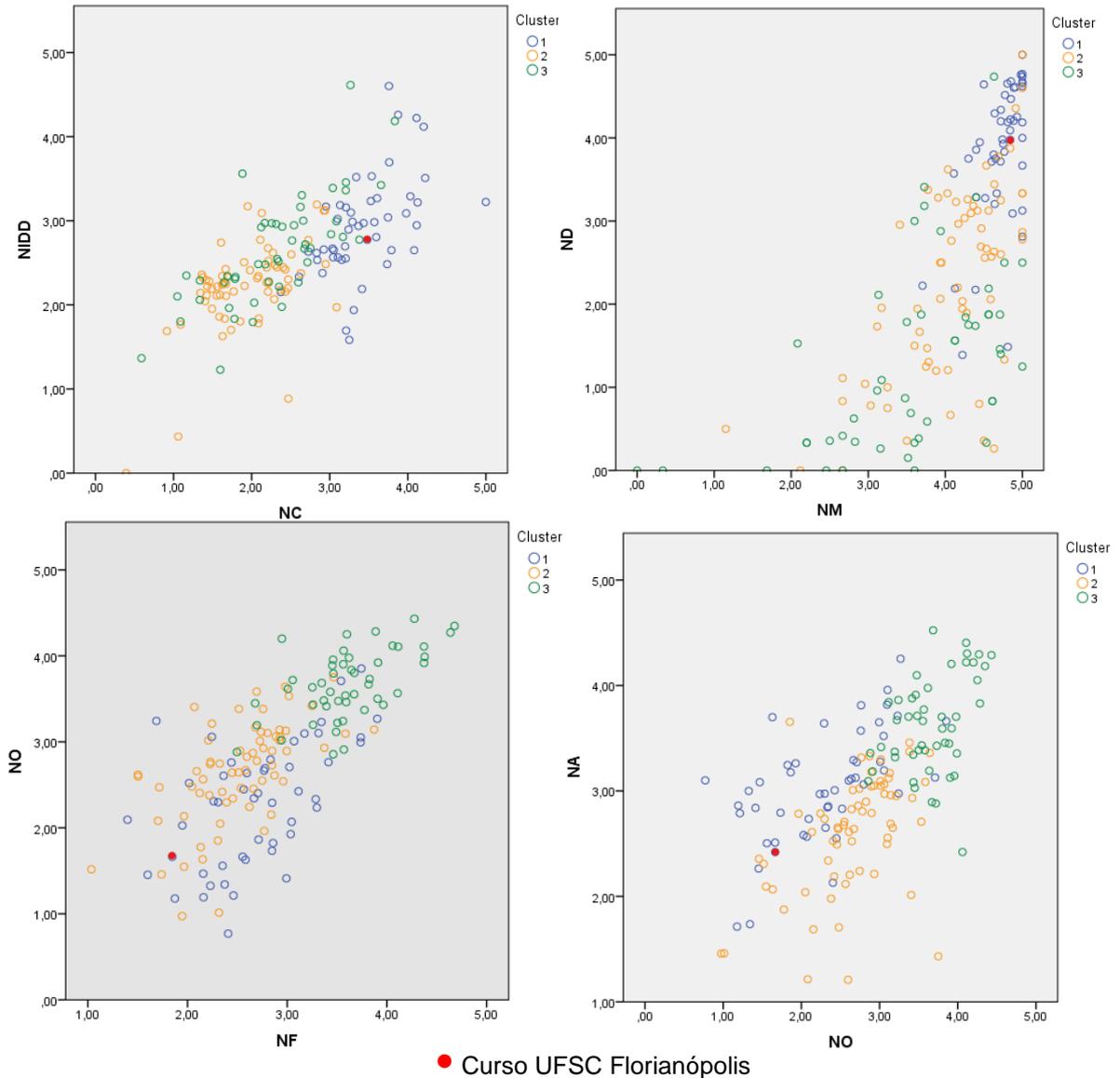
Esses gráficos mostram de forma visual o agrupamento dos cursos em cada *cluster* e tem-se uma boa visualização das similaridades internas de cada *cluster* e das relações entre as componentes de cada dimensão.

Individualmente em cada diagrama, observa-se relações diretas entre as variáveis. Os cursos com notas altas em uma variável, em geral, têm uma tendência de ter notas altas na outra variável; o oposto também acontece.

Além disso, analisando conjuntamente os quatro diagramas, observa-se que os cursos que tiveram notas altas nos componentes do primeiro e segundo diagramas (relacionadas ao desempenho dos alunos e titulação docente), tendem a ter notas baixas nos outros dois diagramas (relacionados à percepção dos alunos) e vice-versa.

Essas informações reforçam e complementam as constatações a respeito das características dos *clusters* discutidas na seção anterior.

Figura 8 – Gráficos de dispersão relacionando NC com NIDD, NM com ND, NF com NO e NO com NA respectivamente



Fonte: Elaborada pela autora

Observando os gráficos, é possível identificar o desempenho do curso da UFSC de Florianópolis em todas as componentes, e o desempenho por dimensões. A Nota dos Concluintes no Enade (NC) e a da Diferença entre o Desempenho Observado e Esperado (NIDD), que dependem do aprendizado do aluno, estão próximas de 3,0, são boas comparadas aos outros cursos e aparentemente estão centralizadas no *cluster* 1.

O segundo diagrama apresenta as notas de proporção de Mestres (NM) e Doutores (ND), em que as notas do curso da UFSC são visivelmente altas, o que significa que a grande maioria dos professores deste curso possui título de mestre ou doutor.

Os outros dois diagramas mostram as notas de Organização Didático-Pedagógica (NO), de Infraestrutura e Instalações Físicas (NF) e de Oportunidade de Ampliação da Formação Acadêmica (NA), referentes à dimensão que depende da percepção dos alunos sobre estrutura e oportunidades do curso e, nesta dimensão, o curso de Florianópolis foi mal avaliado pelos estudantes em comparação com os outros cursos. Portanto, esta é a dimensão que apresenta maior potencial de melhoria comparativamente aos demais cursos.

É relevante destacar que o curso de Licenciatura em Matemática da UFSC de Florianópolis pertence ao *cluster* 1 e essas características analisadas estão de acordo com o que foi observado desse *cluster*. Este curso recebeu conceito 4 no CPC avaliado no ano de 2017.

5 CONSIDERAÇÕES FINAIS

Neste trabalho, foi apresentada a avaliação dos cursos de Matemática de acordo com as componentes do Conceito Preliminar de Curso; e as notas dos cursos em cada componente foram analisadas de forma descritiva e exploratória. A partir dessas componentes, foi aplicado o método de clusterização k-médias com a finalidade de formar agrupamentos de cursos com elevada homogeneidade interna (dentro dos grupos) e heterogeneidade externa (entre os grupos). Como resultado obteve-se 3 grupos de cursos, a partir disso foi feita uma análise exploratória a fim de caracterizar cada *cluster* de acordo com o perfil dos cursos pertencentes a ele; verificando-se, portanto, padrões de desempenho similares suficientemente significativos para permitir afirmar a existência de agrupamentos naturais entre os cursos de graduação em Matemática.

O *cluster 1* ficou com 49 cursos, o *cluster 2* com 66 cursos e o *cluster 3* com 48 cursos. A análise dos *clusters* mostrou que o *cluster 1* agrupou os cursos com maiores notas nas componentes das dimensões Desempenho dos Estudantes, Valor Agregado pelo Processo Formativo Oferecido pelo Curso e Corpo Docente. O *cluster 3* agrupou os cursos com maiores notas na dimensão Percepção Discente sobre as Condições do Processo Formativo. O *cluster 2*, por sua vez, agrupou os cursos com notas intermediárias na maioria das componentes. Observando o conceito que cada curso recebeu, nota-se que o *cluster 1* agrupou os melhores cursos de acordo com o Conceito Preliminar de Curso.

O curso de Licenciatura em Matemática da UFSC de Florianópolis, considerado no agrupamento, apresentou notas baixas nas componentes referentes a Organização Didático-Pedagógica, Infraestrutura e Instalações Físicas e Oportunidades de Ampliação na Formação Acadêmica e Profissional. Assim, essas são as componentes mais críticas para se pensar em melhorias. É interessante notar que as três componentes fazem parte da dimensão Percepção Discente sobre as Condições do Processo Formativo. Isso significa que uma forma de aumentar as notas é melhorar a percepção dos alunos do curso sobre as condições oferecidas. Uma outra forma é melhorar de fato as condições destacadas por cada componente, se assim for necessário, e evidenciar as melhorias para os estudantes.

Os algoritmos de agrupamento proporcionam a comparação entre cursos que possuem desempenho similar ou distinto, com isso podem ajudar na identificação de características predominantes positivas ou que podem melhorar em cada curso e, portanto, podem contribuir para o desenvolvimento da educação no país.

Como sugestão, recomenda-se para trabalhos futuros inserir na análise mais variáveis explicativas que reflitam as características dos cursos, por exemplo, número de docentes, número de alunos matriculados e diplomados, índices de evasão e reprovação e turno em que o curso é oferecido, de maneira a aprimorar a identificação de cursos com características similares. Assim, quanto mais similares forem os cursos dentro de cada *cluster*, mais efetiva poderá ser a observação entre pares em busca de exemplos de boas práticas.

REFERÊNCIAS

BARREYRO, B. G.; ROTHEN, J. C. Para uma história da avaliação da educação superior brasileira: análise dos documentos PARU, CNRES, GERES e PAIUB. **Avaliação**, Sorocaba, v. 13, n. 1, p. 131-152, mar. 2008.

BARREYRO, B. G.; ROTHEN, J. C. Percurso da avaliação da educação superior nos governos Lula. **Educação e Pesquisa**, São Paulo, v. 40, n.1, p. 61-76, mar. 2014.

BRASIL. **Lei 10.861, de 14 de abril de 2004**. Institui o Sistema Nacional de Avaliação da Educação Superior – SINAES e dá outras providências. Diário Oficial da União, Brasília, DF, 15 abr. 2004.

BRASIL. Ministério da Educação. **Sistema Nacional de Avaliação da Educação Superior (Sinaes)**: bases para uma nova proposta de avaliação da educação superior. Brasília, DF: MEC, 2003. Disponível em: <<http://portal.mec.gov.br/arquivos/pdf/sinaes.pdf>>. Acesso em: 23 mar. 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Nota técnica n. 58, de 04 de dezembro de 2020**. Inep, 2020. Disponível em: <https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2019/NOTA_TECNICA_N_58-2020_CGCQES-DAES_Metodologia_de_calculo_do_CPC_2019.pdf>. Acesso em 24 abr. 2021.

BRASÍLIA. **Portaria n. 2.051, de 9 de julho de 2004**. Regulamenta os procedimentos de avaliação do Sistema Nacional de Avaliação da Educação Superior (SINAES), instituído na Lei no 10.861, de 14 de abril de 2004. Disponível em: <http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=533-portaria-2051&category_slug=documentos-pdf&Itemid=30192>. Acesso em: 23 mar. 2021.

FACELLI, Katti. **Um Framework para análise de agrupamento baseado na combinação multi-objetivo de algoritmos de agrupamento**. 2007. Tese (Doutorado em Ciências da Computação e Matemática Computacional) — Universidade de São Paulo, São Carlos, 2007. Disponível em: <<https://teses.usp.br/teses/disponiveis/55/55134/tde-12012007-082216/publico/teseKatti.pdf>>. Acesso em 23 fev. 2021.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

HAIR, Joseph F.; BLACK, William C.; BABIN, Barry, J.; ANDERSON, Rolph, E.; TATHAM, Ronald L. **Análise Multivariada de Dados**. 6 ed. Porto Alegre: Bookman, 2009.

HUPPES, J. C.; SOUZA, A. M.; ANSUJ, A. P. Classificação da Similaridade dos Cursos de Graduação da UFSM em Relação ao Índice do Aluno Equivalente – Uma Ferramenta de Gestão. **Revista Espacios**, v. 37, n.13, p. 22-30, mar. 2016. Disponível em: <<https://www.revistaespacios.com/a16v37n13/16371322.html>>. Acesso em: 12 abr. 2021.

IKUTA, C. Y. S. Sobre o Conceito Preliminar de Curso: concepção, aplicação e mudanças metodológicas. **Estudos em Avaliação Educacional**, São Paulo, v. 27, n. 66, p. 938–969, dez. 2016. Disponível em: <<http://dx.doi.org/10.18222/ea.e.v27i66.4039>>. Acesso em: 29 abr. 2021.

JUNIOR, J. S. S.; PENHOLATO, J. P.; ERLER, I. da S.; CARNEIRO, T. C. J. Proposição de Indicadores para o Corpo Discente e Análise de Agrupamentos Aplicada aos Cursos de Graduação da UFES. **Revista Gestão Universitária na América Latina**, Santa Catarina, v. 6, n. 2, p. 106-125, mai. 2013. Disponível em: <<https://www.redalyc.org/articulo.oa?id=319327519007>>. Acesso em: 12 abr. 2021.

LEITE, M. F. B.; VIANA, A. B. N.; PEREIRA, G. G.; MANTOVANI, D. M. N. Análise dos Programas de Pós-graduação em Administração Reconhecidos pela CAPES: uma Abordagem Multivariada por Análise de Cluster. **FACEF Pesquisa**, v. 10, n. 1, p. 20-34, fev. 2007. Disponível em: <<http://periodicos.unifacef.com.br/index.php/facefpesquisa/article/view/87>>. Acesso em: 12 abr. 2021.

LINDEN, R. Técnicas de Agrupamento. **Revista de Sistemas de Informação da FSMA**, Rio de Janeiro, n. 4, p. 18-36, 2009. Disponível em: <https://www.researchgate.net/publication/267710538_Tecnicas_de_Agrupamento>. Acesso em: 24 fev. 2021.

MACHADO, Raquel de Lima. **Desenvolvimento de um Algoritmo Imunológico para Agrupamento de Dados**. 2011. Trabalho de Conclusão de Curso (Bacharel em Ciência da Computação) - Universidade de Caxias do Sul, Caxias do Sul, 2011. Disponível em: <<https://repositorio.ucs.br/xmlui/bitstream/handle/11338/1486/TCC%20Raquel%20de%20Lima%20Machado.pdf?sequence=1&isAllowed=y>>. Acesso em: 10 mar. 2021.

OLIVEIRA, É. T. de; PICONEZ, S. C. B. Avaliação da educação superior nas modalidades presencial e a distância: análises com base no Conceito Preliminar de Cursos (CPC). **Avaliação**, Sorocaba, v. 22, n. 03, p. 833-851, nov. 2017.

PIRAN, Fabio Sartori; LACERDA, Daniel Pacheco; CAMARGO, Luis Felipe Riehs. **Análise e Gestão Da Eficiência**. 1 ed. Rio de Janeiro: Elsevier, 2018.

PORTO, Julianna Pinele Santos. **Geometria do Modelo Estatístico das Distribuições Normais Multivariadas**. 2017. Tese (Doutorado em Matemática Aplicada) - Universidade Estadual de Campinas, Campinas, 2017. Disponível em: <http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/325689/1/Porto_JuliannaPineleSantos_D.pdf>. Acesso em: 11 mar. 2021.

RISTOFF, D.; GIOLO, J. O Sinaes Como Sistema. **Revista Brasileira de Pós-Graduação**, Brasília, v.3, n. 6, p. 193-213, dez. 2006. Disponível em: <<http://ojs.rbpg.capes.gov.br/index.php/rbpg/article/view/106/100>>. Acesso em: 23 mar. 2021.

ZAINKO, M. A. S. Avaliação da Educação Superior no Brasil: Processo de Construção Histórica. **Revista da Avaliação da Educação Superior**, Sorocaba, v. 13, n. 3, p. 827-831, nov. 2010. Disponível em: <<https://www.scielo.br/pdf/aval/v13n3/12.pdf>>. Acesso em: 15 mar. 2021.

ZANELLA, A.; OLIVEIRA, R. M. e S. de. Avaliação de desempenho na educação superior: guiando melhorias e valorizando as especialidades de cada curso. **Artigo em avaliação na Revista Avaliação da Educação Superior**, dez. 2020.

ANEXO A – Base de dados utilizada no agrupamento

Código do Curso (MEC)	Sigla da IES	Município do Curso	Organização Acadêmica	NID									CPC	Cluster
				NC	D	NM	ND	NR	NO	NF	NA			
1102402	UFABC	Santo André	Universidade	5,00	3,22	5,00	5,00	5,00	3,05	3,74	3,52	5	1	
13873	UFMS	Santa Maria	Universidade	3,76	4,60	4,77	4,52	5,00	1,63	2,58	3,70	5	1	
122358	UFV	Florestal	Universidade	4,11	4,22	5,00	3,67	5,00	2,76	2,44	3,57	5	1	
96955	UNIFAL-MG	Alfenas	Universidade	4,20	4,12	4,30	3,75	5,00	3,85	3,74	3,66	5	1	
1035	FURG	Rio Grande	Universidade	2,91	2,38	4,11	3,57	5,00	2,07	3,04	2,57	4	1	
1117203	IF SUL DE MINAS	Inconfidentes	Instituto Federal	3,10	3,02	4,12	2,19	4,44	2,69	2,78	3,13	4	1	
1122743	IFBA	Salvador	Instituto Federal	3,76	3,70	4,64	3,21	5,00	3,06	2,24	3,19	4	1	
108730	IFES	Vitória	Instituto Federal	3,31	1,94	4,67	3,33	5,00	3,10	3,17	3,82	4	1	
116820	IFMG	Formiga	Instituto Federal	2,95	3,17	4,22	1,39	5,00	2,43	3,11	3,01	4	1	
113490	IFSP	São Paulo Bragança Paulista	Instituto Federal	3,53	3,23	4,87	3,09	5,00	3,24	1,69	2,98	4	1	
1128375	IFSP	São Paulo Bragança Paulista	Instituto Federal	2,83	2,62	5,00	2,81	5,00	3,01	3,07	3,28	4	1	
13276	UFBA	Salvador	Universidade	2,92	2,66	4,74	3,98	5,00	1,33	2,23	3,00	4	1	
12702	UFF	Niterói	Universidade	3,79	2,65	4,91	4,61	5,00	2,76	3,41	3,81	4	1	
1131697	UFG	Goiânia	Universidade	3,05	2,65	4,84	4,09	5,00	1,93	3,04	3,26	4	1	
1304557	UFG	Goiânia	Universidade	3,24	2,90	4,72	4,20	5,00	2,33	3,29	2,84	4	1	
1166038	UFJF	Juiz de Fora	Universidade	4,03	3,29	4,64	4,29	5,00	1,73	2,85	2,79	4	1	
101556	UFLA	Lavras	Universidade	3,04	2,67	5,00	4,19	5,00	2,99	3,74	3,65	4	1	
12969	UFMG	Belo Horizonte	Universidade	4,08	2,65	4,98	4,76	5,00	0,77	2,41	3,10	4	1	
15833	UFMS	Campo Grande	Universidade	3,27	3,10	4,77	3,83	5,00	3,10	3,31	3,96	4	1	
15865	UFMS	Três Lagoas	Universidade	3,54	3,53	5,00	3,13	5,00	2,40	2,71	2,13	4	1	
34	UFMT	Rondonópolis	Universidade	3,21	3,16	3,70	2,22	5,00	2,09	1,40	2,73	4	1	
13619	UFPE	Recife	Universidade	3,60	2,81	4,72	3,72	5,00	1,18	1,87	1,71	4	1	
15000	UFPEL	Capão do Leão	Universidade	3,88	4,26	4,81	4,19	5,00	1,47	2,16	3,08	4	1	
122746	UFPEL	Pelotas	Universidade	2,84	2,53	4,45	3,95	5,00	2,61	2,36	3,10	4	1	
12553	UFPR	Curitiba	Universidade	3,29	2,99	4,85	4,68	5,00	1,45	1,60	2,26	4	1	
49478	UFPR	Curitiba	Universidade	3,61	3,27	4,50	4,64	5,00	1,86	2,71	3,18	4	1	
13741	UFRGS	Porto Alegre	Universidade	3,98	3,09	5,00	4,62	5,00	2,03	1,95	2,58	4	1	
14328	UFRJ	Rio de Janeiro	Universidade	3,20	2,70	4,81	4,65	5,00	2,29	2,85	3,64	4	1	
312333	UFRN	Natal	Universidade	2,73	2,63	5,00	4,00	5,00	2,24	3,30	2,97	4	1	
14505	UFRPE	Recife	Universidade	3,37	2,94	4,63	3,80	5,00	1,34	2,37	1,74	4	1	
12918	UFRRJ	Seropédica	Universidade	3,20	2,55	4,72	4,34	5,00	2,31	2,26	2,65	4	1	
96158	UFRRJ	Nova Iguaçu	Universidade	2,37	2,15	5,00	4,77	5,00	2,45	2,56	2,55	4	1	
14233	UFSC	Florianópolis	Universidade	3,48	2,77	4,84	3,98	5,00	1,66	1,85	2,42	4	1	
624	UFSCAR	São Carlos	Universidade	3,10	2,56	4,89	4,60	5,00	1,41	2,99	2,84	4	1	
637	UFSCAR	São Carlos	Universidade	3,43	2,97	4,85	4,47	5,00	1,21	2,46	2,79	4	1	
115086	UFSCAR	Sorocaba São João del Rei	Universidade	3,74	2,48	5,00	4,74	5,00	3,23	3,34	3,71	4	1	
48946	UFSJ	São João del Rei	Universidade	3,57	2,98	4,65	3,75	5,00	3,71	3,54	3,13	4	1	
41069	UFMS	Santa Maria	Universidade	3,13	3,19	4,60	3,71	5,00	1,66	2,55	2,51	4	1	
114556	UFTM	Uberaba	Universidade	3,34	3,52	4,40	3,29	5,00	2,67	2,77	3,29	4	1	
1428	UFU	Uberlândia	Universidade	2,61	2,33	5,00	4,66	5,00	2,34	2,67	2,85	4	1	
696	UFV	Viçosa	Universidade	3,41	2,19	5,00	4,67	5,00	1,82	2,86	3,25	4	1	

120389	UNIFEI	Itajubá	Universidade	4,22	3,51	4,40	3,86	5,00	3,27	3,90	4,25	4	1
1101770	UNIRIO	Rio de Janeiro	Universidade	4,11	2,95	4,75	3,93	5,00	2,52	2,02	2,83	4	1
14542	UTFPR	Pato Branco	Universidade	3,75	3,04	4,52	3,28	5,00	2,71	3,02	3,27	4	1
1128169	UTFPR	Curitiba	Universidade	4,13	3,22	4,89	4,21	5,00	2,64	2,60	3,12	4	1
102430	IFRJ	Nilópolis	Instituto Federal	3,04	2,57	4,81	1,49	5,00	2,79	2,83	3,06	3	1
122122	IFSP	Guarulhos	Instituto Federal	3,15	2,53	4,39	2,17	5,00	2,30	2,31	2,97	3	1
12806	UFES	Vitória	Universidade	3,25	1,58	4,84	4,22	5,00	1,19	2,16	2,86	3	1
44372	UNB	Brasília	Universidade	3,21	1,69	4,93	4,25	5,00	1,56	2,35	2,50	3	1
1103031	IF Catarinense	Concórdia Campina Grande	Instituto Federal	2,73	2,77	5,00	2,78	5,00	2,75	2,32	2,24	4	2
13448	UFCG		Universidade	2,95	2,48	4,47	2,91	5,00	2,41	2,12	2,53	4	2
1152880	UFCG	Cajazeiras	Universidade	2,33	2,45	4,42	3,13	5,00	3,42	3,25	2,93	4	2
15868	UFGD	Dourados Pontal do	Universidade	2,13	3,09	4,46	2,69	5,00	2,55	2,75	2,71	4	2
18354	UFMT	Araguaia	Universidade	2,22	2,26	4,53	3,67	5,00	3,38	2,76	3,46	4	2
118100	UFPE IF	Caruaru	Universidade	2,40	2,45	4,29	3,26	5,00	2,90	2,85	3,18	4	2
1102795	Catarinense IF	Rio do Sul	Instituto Federal	1,49	1,95	4,22	1,94	5,00	2,47	1,72	2,64	3	2
1103025	Catarinense	Camboriú Cruzeiro do Sul	Instituto Federal	1,85	1,80	4,59	2,06	5,00	2,24	2,62	2,79	3	2
5000487	IFAC		Instituto Federal	1,63	2,25	1,15	0,50	5,00	2,78	2,69	2,81	3	2
1103556	IFAL	Maceió	Instituto Federal	1,40	2,04	3,78	1,30	5,00	3,38	2,51	3,37	3	2
5000493	IFAM	Manaus	Instituto Federal	2,21	2,67	3,75	1,25	5,00	3,12	2,74	3,12	3	2
116730	IFBA	Barreiras	Instituto Federal	2,94	3,12	3,25	0,75	5,00	2,62	1,50	2,20	3	2
123519	IFBA	Valença	Instituto Federal	1,91	2,23	3,03	0,78	5,00	2,56	2,16	2,12	3	2
63383	IFCE	Fortaleza	Instituto Federal	2,29	2,07	4,28	1,90	5,00	2,74	2,23	3,02	3	2
1160979	IFCE	Canindé	Instituto Federal	2,43	2,42	4,44	0,80	5,00	3,06	2,93	3,06	3	2
5000138	IFCE	Crateús	Instituto Federal	2,37	2,18	4,07	0,67	5,00	2,93	3,37	2,21	3	2
1103461	IFG	Goiânia	Instituto Federal	1,98	2,41	3,93	2,07	5,00	2,64	2,45	2,83	3	2
21443	IFMA	São Luís	Instituto Federal	2,19	2,27	3,12	1,73	5,00	2,66	2,09	3,01	3	2
99503	IFNMG	Januária	Instituto Federal	1,90	2,51	2,96	1,04	5,00	2,46	2,54	2,65	3	2
48289	IFPA	Belém	Instituto Federal	2,47	2,20	4,22	2,04	4,67	3,21	2,25	2,95	3	2
63062	IFPI	Teresina Angical do Piauí	Instituto Federal	2,63	2,38	4,03	1,21	5,00	2,67	2,58	2,64	3	2
1103910	IFPI		Instituto Federal	1,62	2,24	2,12	0,00	5,00	2,83	2,60	2,90	3	2
121013	IFRJ	Volta Redonda	Instituto Federal	3,09	1,97	4,72	2,60	5,00	2,54	2,96	2,68	3	2
123787	IFRN	Mossoró	Instituto Federal	1,67	2,25	4,63	0,26	5,00	3,09	3,58	2,50	3	2
1123270	IFRN	Natal	Instituto Federal	2,51	2,60	4,16	2,20	5,00	2,34	2,45	2,34	3	2
1156565	IFRS	Ibirubá	Instituto Federal	2,84	3,19	4,50	0,36	5,00	2,77	2,23	3,32	3	2
96913	IFS	Aracaju Paraíso do Tocantins	Instituto Federal	1,59	1,86	4,77	1,33	5,00	2,61	2,87	2,61	3	2
1106721	IFTO		Instituto Federal	1,95	3,17	3,50	0,36	5,00	1,85	2,31	3,65	3	2
102152	UFAL	Arapiraca	Universidade	2,30	2,48	4,51	2,56	5,00	3,01	2,73	3,09	3	2
107520	UFAL	Maceió	Universidade	2,21	2,52	4,69	3,78	5,00	1,63	2,15	2,07	3	2
379	UFAM	Manaus	Universidade	2,09	2,34	4,17	2,97	5,00	3,10	2,80	2,56	3	2
38246	UFC	Fortaleza	Universidade	2,19	2,44	5,00	4,60	5,00	0,97	1,95	1,46	3	2
99806	UFCG	Cuité Santo Antônio de Pádua	Universidade	2,11	2,77	4,60	2,57	5,00	1,96	2,77	2,78	3	2
12713	UFF		Universidade	1,66	1,83	4,08	2,76	5,00	3,13	2,99	3,17	3	2
14121	UFG	Catalão	Universidade	2,09	1,84	4,03	3,33	5,00	2,15	2,84	1,69	3	2
11439	UFMA	São Luís	Universidade	2,09	1,78	5,00	2,87	5,00	2,13	1,96	2,61	3	2

18382	UFMS	Aquidauana	Universidade	1,42	2,11	3,25	1,00	5,00	2,46	2,66	2,49	3	2
26	UFMT	Cuiabá	Universidade	1,74	1,70	3,77	3,38	5,00	2,05	2,33	2,04	3	2
12039	UFOPA	Santarém	Universidade	1,63	1,63	5,00	5,00	5,00	3,64	2,98	3,36	3	2
11999	UFPA	Belém	Universidade	1,37	2,33	4,64	3,44	5,00	3,17	2,68	2,65	3	2
12033	UFPA	Bragança	Universidade	1,59	2,12	4,63	2,63	5,00	2,08	1,70	1,21	3	2
12035	UFPA	Castanhal	Universidade	1,55	2,11	4,36	3,18	5,00	2,48	2,04	1,71	3	2
12044	UFPA	Breves	Universidade	1,42	2,29	2,67	1,11	5,00	3,40	2,07	2,01	3	2
12074	UFPA	Cametá	Universidade	1,56	2,22	5,00	3,33	5,00	3,53	3,02	2,71	3	2
114853	UFPA	Abaetetuba	Universidade	1,34	2,14	3,95	2,50	5,00	2,60	1,50	1,21	3	2
13402	UFPB	João Pessoa	Universidade	2,47	2,30	4,84	3,88	5,00	1,52	1,04	2,31	3	2
99045	UFPB	Rio Tinto	Universidade	1,49	2,19	4,25	3,04	5,00	2,93	2,76	3,05	3	2
102576	UFPI	Parnaíba	Universidade	1,77	2,16	3,17	1,96	5,00	1,78	2,13	1,88	3	2
102588	UFPI	Picos	Universidade	1,46	2,28	2,67	0,83	5,00	1,01	2,32	1,46	3	2
300520	UFPI	Teresina	Universidade	2,43	2,16	4,72	3,25	5,00	2,30	2,39	2,73	3	2
100412	UFRB	Amargosa	Universidade	2,27	2,62	3,93	2,50	5,00	3,02	2,21	3,05	3	2
12354	UFRN	Caicó	Universidade	1,67	2,11	4,33	3,10	5,00	2,64	2,51	2,52	3	2
297	UFS	São Cristóvão	Universidade	2,07	2,32	4,03	3,62	5,00	2,73	2,85	2,86	3	2
95055	UFS	Itabaiana	Universidade	1,61	2,74	4,56	3,13	5,00	1,55	1,96	2,09	3	2
17154	UFT	Arraias	Universidade	1,36	2,36	3,60	1,50	5,00	2,89	2,99	3,04	3	2
1105218	UFT	Araguaína	Universidade	1,67	2,43	3,67	1,67	5,00	3,06	2,89	2,97	3	2
1121464	UFT	Palmas	Universidade	1,06	0,43	5,00	3,33	5,00	3,75	3,47	1,43	3	2
100916	UFVJM	Teófilo Otoni	Universidade	2,93	3,12	3,88	1,20	5,00	2,38	2,22	1,98	3	2
161	UNB	Brasília	Universidade	2,47	0,88	4,91	4,36	5,00	1,46	1,74	2,36	3	2
17196	UNIFAP	Macapá	Universidade	1,62	2,34	3,64	1,94	5,00	3,14	2,92	2,69	3	2
12037	UNIFESSP A	Marabá	Universidade	0,91	1,69	3,76	1,47	5,00	3,58	2,69	3,08	3	2
104282	UNIPAMPA	Bagé	Universidade	2,21	2,16	3,91	3,28	5,00	3,02	2,92	3,30	3	2
16009	UNIR	Porto Velho	Universidade	1,43	2,22	3,41	2,95	5,00	2,42	2,35	2,19	3	2
16010	UNIR	Ji-Paraná	Universidade	1,69	2,33	4,53	2,67	5,00	2,90	2,53	2,79	3	2
1102641	IFNMG	Salinas	Instituto Federal	1,09	1,77	2,67	0,00	5,00	2,87	2,66	3,33	2	2
117052	UFBA	Salvador	Universidade	0,39	0,00	4,14	3,23	5,00	3,14	3,87	2,96	2	2
IF													
1103032	Catarinense IF	Sombrio	Instituto Federal	2,17	2,97	4,56	1,88	5,00	3,89	3,45	3,45	4	3
52442	Fluminense	Goytacazes	Instituto Federal	3,66	3,42	3,65	0,38	4,32	3,92	3,91	3,59	4	3
1103657	IFES	Itapemirim	Instituto Federal	3,38	2,78	3,50	1,79	5,00	3,84	3,64	3,46	4	3
1103185	IFFarroupilh a	Júlio de Castilhos	Instituto Federal	1,88	3,56	5,00	2,50	5,00	4,11	4,11	4,40	4	3
1128936	IFFarroupilh a	Alegrete	Instituto Federal	2,26	2,97	4,12	1,56	5,00	4,11	4,37	4,22	4	3
1128938	IFFarroupilh a	Santa Rosa	Instituto Federal	3,27	4,61	4,61	0,83	5,00	4,27	4,64	4,30	4	3
1168573	IFFarroupilh a	São Borja	Instituto Federal	2,63	3,16	4,61	0,83	5,00	3,50	3,91	3,39	4	3
1184450	IFRN	Santa Cruz	Instituto Federal	3,21	3,46	4,30	1,75	5,00	3,92	4,37	4,20	4	3
1292902	IFRO	Cacoal	Instituto Federal	2,64	3,30	4,53	0,33	5,00	3,68	3,35	4,52	4	3
116144	IFRS	Bento Gonçalves	Instituto Federal	3,04	3,39	4,39	1,74	5,00	3,24	3,56	3,86	4	3
1126144	IFRS	Caxias do Sul	Instituto Federal	3,83	4,19	4,56	1,88	5,00	3,46	3,59	3,91	4	3
113589	IFSEMG	Rio Pomba	Instituto Federal	2,31	2,96	4,56	2,19	5,00	3,48	3,37	3,71	4	3
1128378	IFSP	Caraguatatuba	Instituto Federal	3,17	2,81	4,12	1,56	5,00	3,98	3,62	3,70	4	3
116882	UFES	São Mateus	Universidade	2,69	2,72	3,94	2,88	5,00	3,54	3,58	3,43	4	3

100775	UFMT	Sinop	Universidade	2,75	2,67	3,60	3,00	5,00	4,20	2,95	4,22	4	3
102944	UFU	Ituiutaba	Universidade	2,39	2,93	4,63	4,74	5,00	3,57	4,11	3,77	4	3
5000912	UNIPAMPA	Itaqui	Universidade	2,54	2,95	3,69	1,88	5,00	3,72	3,06	2,88	4	3
1127676	UTFPR	Cornélio Procópio	Universidade	3,21	3,36	4,40	3,29	5,00	4,35	4,68	4,18	4	3
1152637	UTFPR	Toledo	Universidade	2,71	2,51	4,77	2,50	5,00	3,12	3,49	3,84	4	3
1161404	IFBA	Camaçari	Instituto Federal	2,39	1,98	4,71	1,46	5,00	4,25	3,60	4,05	3	3
84374	IFCE	Juazeiro do Norte	Instituto Federal	2,68	2,67	3,55	0,69	5,00	3,43	3,96	3,34	3	3
85320	IFCE	Cedro	Instituto Federal	2,12	2,92	2,83	0,34	5,00	2,86	3,46	3,09	3	3
1103914	IFMA	Codó	Instituto Federal	1,05	2,10	2,81	0,63	5,00	3,63	3,25	3,19	3	3
1264336	IFMA	Buriticupu	Instituto Federal	1,34	2,29	2,50	0,36	3,73	4,28	3,89	3,83	3	3
1327336	IFMA	Caxias	Instituto Federal	1,65	2,27	2,45	0,00	5,00	3,90	3,56	3,13	3	3
1142320	IFMG	São João Evangelista	Instituto Federal	2,34	2,52	3,60	0,33	5,00	3,80	3,67	3,59	3	3
114388	IFMT	Campo Novo do Parecis	Instituto Federal	1,79	2,31	3,60	0,00	5,00	3,19	2,70	3,37	3	3
1160424	IFMT	Juína	Instituto Federal	2,66	3,00	2,20	0,33	5,00	3,62	3,01	3,98	3	3
1128030	IFPB	Campina Grande	Instituto Federal	2,35	2,22	3,48	0,87	5,00	3,67	3,82	2,89	3	3
1128096	IFPB	Cajazeiras	Instituto Federal	3,09	2,99	3,52	0,15	3,92	3,22	3,50	3,67	3	3
100690	IFPE	Pesqueira	Instituto Federal	1,69	2,34	3,17	1,09	5,00	3,20	3,26	3,32	3	3
83929	IFPI	Florianópolis	Instituto Federal	2,09	2,48	1,68	0,00	5,00	2,91	3,57	3,19	3	3
1103813	IFPI	Piripiri	Instituto Federal	1,78	1,83	2,67	0,00	5,00	3,55	3,67	3,66	3	3
1103911	IFPI	Corrente	Instituto Federal	1,16	2,35	2,20	0,33	5,00	3,43	3,26	3,08	3	3
1153801	IFRJ	Paracambi	Instituto Federal	2,03	2,03	5,00	1,25	5,00	4,43	4,28	4,29	3	3
1152312	IFRR	Boa Vista	Instituto Federal	1,34	2,06	3,12	0,96	5,00	3,02	2,94	3,42	3	3
1128360	IFSP	Birigui	Instituto Federal	2,33	2,55	4,71	1,88	5,00	3,79	3,46	3,70	3	3
1129073	IFSP	Araraquara	Instituto Federal	2,60	2,27	4,26	1,84	5,00	3,73	3,83	3,43	3	3
1181042	IFSP	Campos do Jordão	Instituto Federal	3,02	2,84	3,16	0,26	5,00	3,45	2,68	3,03	3	3
123494	IFTO	Palmas	Instituto Federal	1,79	2,33	3,76	0,59	5,00	3,58	3,43	3,39	3	3
11540	UFAC	Rio Branco	Universidade	2,52	2,77	3,13	2,11	4,80	3,37	3,77	3,70	3	3
18954	UFG	Jataí	Universidade	2,02	1,80	3,73	3,41	5,00	4,12	4,06	4,30	3	3
15862	UFMS	Corumbá	Universidade	2,22	2,28	4,72	1,40	4,29	3,42	3,39	3,88	3	3
52139	UFMS	Paranaíba	Universidade	1,71	1,96	3,73	3,18	5,00	3,99	4,38	3,36	3	3
121792	UFMS	Ponta Porã	Universidade	2,17	2,48	2,08	1,53	4,01	2,88	2,50	3,36	3	3
1160371	IFMA	Zé Doca	Instituto Federal	0,59	1,37	0,33	0,00	5,00	3,96	3,46	3,14	2	3
1103912	IFPI	Uruçuí	Instituto Federal	1,09	1,80	0,00	0,00	5,00	4,06	3,57	2,42	2	3
1182765	IFRO	Vilhena	Instituto Federal	1,60	1,23	2,67	0,42	5,00	3,48	3,51	4,10	2	3