

## APOIO À DECISÃO NA GESTÃO DE PROGRAMAS DE PÓS-GRADUAÇÃO: MODELOS E INDICADORES

### Resumo:

São vários os desafios ligados à gestão acadêmica e produção científica. Os esforços empregados para apoiar as tarefas ligadas à tomada de decisão são as mais diversas. Uma das maneiras para minimizar essas dificuldades é disponibilizar, a esses gestores, informações com qualidade e de forma segura; e a tecnologia da informação pode contribuir para a segurança e confiabilidade das informações. Este trabalho busca entender de que forma os itens de avaliação da CAPES podem contribuir para apoio à decisão em programas de pós-graduação. Para isso, foram utilizadas técnicas estatísticas, em especial, relacionadas à regressão logística. Além disso, propõe métodos para a extração e transformação dos dados. Modelos dimensionais foram construídos como solução para apoio no processo decisório. Os resultados evidenciam que três itens de avaliação da ficha de avaliação da CAPES influenciam fortemente o resultado de uma avaliação. Esses modelos foram testados com dados reais de um programa de pós-graduação e parecem se mostrar eficientes para apoio à gestão.

### Palavras-chave:

sistemas de apoio à decisão, *business intelligence*, gestão acadêmica.

## DECISION SUPPORT IN GRADUATE PROGRAMS MANAGEMENT: MODELS AND INDICATORS

### Abstract:

There are several challenges related to academic management and scientific production. The efforts employed to support the tasks related to decision making are the most diverse. One of the ways to minimize these difficulties is to provide these managers with information with quality and in a safe way; and information technology can contribute to the security and reliability of the information. This paper seeks to understand how CAPES assessment items can contribute to support decision making in graduate programs. For this, statistical techniques were used, especially related to logistic regression. In addition, it proposes methods for the extraction and transformation of data. Dimensional models were built as a solution to support the decision-making process. The results show that three evaluation items on the CAPES evaluation form strongly influence the result of an evaluation. These models have been tested with real data from a graduate program and appear to be efficient for supporting management.

### Keywords:

*decision support systems, business intelligence, academic management.*

## INTRODUÇÃO

Os ambientes de produção científica, em especial, universidades, podem ser entendidos como ambientes complexos (Burigo et al., 2016) e produzem uma grande quantidade de informação, de forma cada vez mais acelerada, o que contribui para tornar o processo de descoberta de padrões e recomendação personalizada ainda mais difícil. Outra questão diz respeito à gestão nessas instituições: docentes assumem cargo de gestão e, muitas vezes, não recebem formação para a tarefa, exigindo das instituições outras alternativas para apoio à gestão (Rodrigues & Villardi, 2017). Além da necessidade de prestação de contas à sociedade, como é o caso das universidades financiadas com recurso público (Cecilia & Lopes, 2015; Comim et al., 2017).

Uma das maneiras de minimizar essas dificuldades é disponibilizar, a esses gestores, informações com qualidade e de forma segura; e a tecnologia da informação pode contribuir para a segurança e confiabilidade das informações (Beal, 2012).

Há uma variedade de ferramentas baseadas em tecnologia da informação que apoiam a organização na gestão da informação e na tomada de decisão. Dentre essas ferramentas, estão os Sistemas de Processamento de Transações (SPT), os Sistemas de Informações Gerenciais (SIG), os Sistemas de Apoio à Decisão (SAD) e o Sistemas de Apoio ao Executivo (SAE) (Turban et al., 2005, 2009).

Cada tipo de sistema surge para solução de demandas computacionais de gestão da informação em um nível organizacional. Os SPTs trabalham a informação em nível operacional; os SIGs trabalham as informações geradas pelos SPT em um nível mais tático, com informações mais consolidada. Já os SADs trabalham com informações oriundas dos SPT e dos SIG e com modelos que apoiam a tomada de decisão. Os SAEs são sistemas que suportam a informação no nível estratégico, de uma forma mais “amigável” ao usuário, inclusive com informações externas à organização (Beal, 2012; Turban et al., 2005, 2009).

Ainda no contexto das IFES, são realizadas, pela Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES)<sup>1</sup>, avaliações quadrienais com o objetivo de melhoria da qualidade dos cursos ofertados. Ou seja, a coordenação do curso, para além das decisões demandadas no cotidiano, deve se preocupar, também, com esses processos de avaliação.

Tem-se no processo de avaliação o objetivo de trazer maior qualidade aos Programas de Pós-Graduação (PPG), alinhando também os objetivos dos programas às necessidades do contexto social em que se encontra e, em especial, com os mestrados profissionais, formação profissional de qualidade para o contexto “não-acadêmico”. É nesse contexto que estão inseridos os Programas de Pós-Graduação (PPGs) com cursos de Mestrado Profissional.

Até 2018, existiam no Brasil, 703 Programas de Pós-Graduação que ofertam cursos de mestrado profissional. Desses, 75 se enquadram na Área de Avaliação “Administração Pública e de Empresas, Ciências Contábeis e Turismo”, uma das 49 áreas de avaliação da CAPES. As informações são do Sistema de Informações Georreferenciadas (GEOCAPES)<sup>2</sup>.

O objetivo desse trabalho, portanto, é propor modelos e indicadores para apoio à decisão para PPG com cursos de mestrado profissional da Área de “Administração Pública e de Empresas, Ciências Contábeis e Turismo”.

Neste contexto, as seguintes questões de pesquisa são colocadas:

- QP1: Quais itens da ficha de avaliação utilizada pela CAPES mais contribuem para o sucesso de um PPG com mestrado profissional?

---

<sup>1</sup><https://www.capes.gov.br/>

<sup>2</sup> <https://geocapes.capes.gov.br/geocapes>

- QP2: De que forma as técnicas relacionadas a SAD podem contribuir na gestão de um PPG?

Este trabalho está organizado em, além desta seção, outras quatro seções. Na Seção “Trabalhos Correlatos” são apresentados trabalhos que apresentam relação com este estudo. Na Seção “Metodologia” são apresentados os passos seguidos para alcance do objetivo proposto e resposta às questões colocadas. Os resultados encontrados são apresentados na Seção “Resultados”. As contribuições dos estudos e expectativas para trabalhos futuros são apresentadas na Seção “Conclusão”.

## TRABALHOS CORRELATOS

Há uma variedade de conceitos que relacionados aos SADs. Os conceitos *Business Intelligence e Analytics (BI&A)*, *Data Warehouse (DW)*, *Data Mart (DM)* são exemplos dos mais utilizados para designar ferramentas e metodologias relacionadas ao tema (Laudon & Laudon, 2011; Turban et al., 2005, 2009). Uma das áreas em que estes conceitos têm sido amplamente aplicados é a de Educação. Nesse contexto, os SADs passam a ser entendidos como sistemas de *Learning Analytics (LA)*, quando o foco é a análise de desempenho em nível pessoal, de um curso ou departamento; e *Academic Analytics (AA)*, quando o foco é em nível institucional, regional, nacional ou internacional (Dawson et al., 2014; Siemens et al., 2011).

Como o foco deste trabalho é o apoio no processo de tomada de decisão em um Programa de Pós-graduação de uma Instituição Federal de Ensino Superior (IFES), o termo *Learning Analytics (LA)* parece estar mais próximo da solução a ser proposta e será adotado como proposto por Siemens et al. (2011).

Borges et al. (2016) utilizam técnicas relacionadas à modelagem dimensional de dados, em especial o *star schema (Kimball, 1997; Kimball & Ross, 2013)*, para um sistema de apoio à decisão direcionado à educação. Os autores propõem o processo de Extração, Transformação e Carga (ETL - do inglês *Extraction, Transformation and Load*). Ou seja, os dados são extraídos das diversas fontes de dados, transformados de acordo com as necessidades de análise e carregados em *Data Warehouse (DW)*, banco de dados específico para sistemas de apoio à decisão.

Uma vez no DW, os dados ficam disponíveis para acesso por meio de cubos OLAP, que possibilitam as análises dos dados considerando as várias dimensões de um modelo. O usuário pode acessar esses dados por meio de painéis de bordo que possibilitam a filtragem e análise dos dados combinando uma ou mais dimensões do modelo, considerando a forma que os dados foram disponibilizados no cubo OLAP. A relação com este trabalho está na concepção dos autores de tratamento das informações, com técnicas de ETL, OLAP e visualização das informações de diferentes perspectivas e granularidades, facilitando a gestão de programas de pós-graduação.

Villegas-Ch et al. (2020) propõem uma arquitetura para adoção de BI e mineração de dados na área acadêmica. O estudo compreende o entendimento do processo de gestão de uma ferramenta capaz de dar apoio à decisão na área acadêmica, utilizando os conceitos fundamentais em sistemas de apoio à decisão, dentre eles o uso de ETL e mineração de dados. Em uma análise mais ampla, a proposta trabalha com os dados acadêmicos e com dados socioeconômicos dos estudantes. O estudo demonstra que a adoção de técnicas e métodos relacionados ao sistema de BI, em especial utilizando ferramentas de ETL e mineração de dados, podem apoiar a gestão educacional.

Em seus estudos, (Gutiérrez et al., 2020) trabalham com a construção e avaliação de aceitação de *dashboards* em duas universidades. A pesquisa envolveu usuários especialistas e não especialistas. Em comparação com outra ferramenta, a proposta dos autores apresenta-se como uma solução para o problema, mas há uma demanda de maior transparência nos algoritmos utilizados para predição. Um ponto importante é que os autores identificaram que a solução é capaz de apoiar na decisão com dados acadêmicos, uma vez que o usuário faz várias análises antes da tomada de decisão.

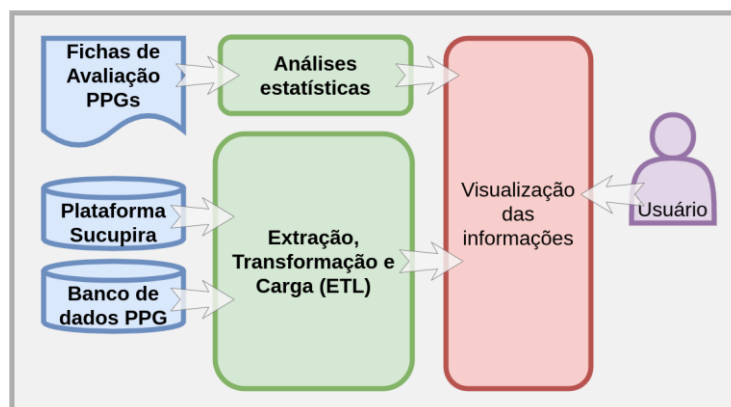
Outros estudos também propõem o uso de ferramentas de LA, utilizando técnicas de ETL, modelagem dimensional, cubos OLAP e visualização de dados por meio de *dashboards*. Em geral, esses estudos possuem foco mais específico em informações de estudantes (Jayakody & Perera, 2016; Nieto et al., 2016).

## METODOLOGIA

Na busca por responder às questões de pesquisa colocadas (Seção Introdução), este estudo foi dividido em duas fases. Na primeira fase, buscou-se analisar quais os itens de avaliação propostos pela CAPES mais contribuem para o sucesso de um PPG com mestrado profissional. E na segunda fase, foram propostos modelos multidimensionais para a construção de um sistema de apoio à decisão.

Ambas as etapas demandaram de busca de dados em fontes da CAPES (fichas de avaliação e dados do PPG). Foram utilizadas técnicas tratamento dos dados, análise estatística, além de técnicas relacionadas à construção de modelos dimensionais. O modelo proposto é apresentado, de forma geral, na Figura 1.

Figura 1 - modelo geral da metodologia proposta



Fonte: elaborado pelo autor

### Obtenção e tratamento dos dados

Na primeira fase, os dados são obtidos das fichas de avaliação da CAPES, referente aos processos de avaliação dos 75 cursos com mestrado profissional da área de “Administração Pública e de Empresas, Ciências Contábeis e Turismo”.

As informações utilizadas das fichas de avaliação foram as notas obtidas pelos PPGs em cada item da ficha. Essas notas correspondem a um conceito dado pelo avaliadores que pode ser “Muito Bom”, “Bom”, “Regular”, “Fraco” e “Inexistente”. Neste estudo, cada uma das variáveis independentes foi codificada em valores numéricos entre conforme Tabela 1.

Para a variável dependente, os valores 0 ou 1 foram atribuídos. A escolha de valores dicotômicos, neste contexto, se deu em razão da busca por itens de avaliação da CAPES que mais contribuem para o sucesso de um PPG com mestrado profissional. Entende-se que o curso com nota maior ou igual a 4 obteve sucesso no processo de avaliação. Do mesmo modo, o curso com nota menor ou igual a 3 fracassou no processo de avaliação. Portanto, o valor 0 foi atribuído aos cursos cuja avaliação final, dada na ficha de avaliação, estivesse entre 0 e 3, inclusive. Caso o valor fosse maior ou igual a 4, o valor 1 foi atribuído.

Na segunda fase, para a construção dos modelos dimensionais, foram utilizados dados da Plataforma Sucupira e de um PPG, caso de estudo para a validação dos modelos, referente ao período de 2013 a 2018. Os dados foram processados utilizando técnicas de Extração, Transformação e Carga (ETL, sigla para *Extraction, Transformatin and Load*), conforme descrito na Seção “Metodologia”.

**Tabela 1: Codificação das variáveis independentes**

Conceito Item (CAPES)	Valor
Muito bom	5
Bom	4
Regular	3
Fraco	2
Insuficiente	1

Fonte: Mesquita e Nogueira (2015)

### **Análise estatística**

Em uma análise inicial, foram realizados testes de normalidade para as variáveis independentes. O teste Kolmogorov-Smirnov foi utilizado para verificar se os dados apresentam distribuição normal. O teste de Kolmogorov-Smirnov tem como hipótese nula ( $H_0$ ) que a variação dos dados segue uma distribuição normal na população e, como hipótese alternativa ( $H_1$ ), que a variação dos dados não segue uma distribuição normal na população, sendo que  $H_0$  deve ser rejeitada quando o nível de significância (p-valor) for menor que 0,05 (Field, 2009).

Após a avaliação da distribuição das variáveis independentes técnicas de regressão logística foram utilizadas (Mesquita & Nogueira, 2015). A utilização de um modelo logístico se justifica pela possibilidade de análise de como cada variável independente, neste caso, os itens de avaliação que compõem a ficha, podem contribuir para o bom desempenho de um PPG, ou seja, nota final para a avaliação maior ou igual a 4.

Diferente da regressão linear, a regressão logística não tem como pressuposto a distribuição normal para as variáveis independentes, sendo necessário que as variáveis não apresentem multicolinearidade dos dados, razão pelo uso dos testes de Tolerância e VIF (Field, 2009).

Um modelo nulo foi considerado para a avaliação dos modelos obtidos com a regressão logística. O modelo nulo considera que todos os indivíduos estão na categoria de maior frequência.

### **Modelos dimensionais e indicadores**

De forma a possibilitar a análise dos indicadores e modelos de simulação propostos, dois modelos dimensionais foram criados: “Atividade Discente” e “Atividade Docente”.

A metodologia adotada para criação dos modelos dimensionais foi a proposta por Kimball e Ross (2013), *star schema* e cubos OLAP. Além de apresentar uma melhor performance para uma quantidade maior de dados, os modelos *star schema* contribuem para um melhor entendimento por parte do usuário. Os modelos foram construídos com a possibilidade de armazenamento ao menor nível de granularidade, o que permite um detalhamento maior de análise durante a combinação dos parâmetros na navegação nos Cubos OLAP.

O modelo “Atividade Docente” considerou as análises feitas por meio das dimensões “Tempo”, “Situação atividade”, “Estrato”, “Categoria Docente”, “Docente” e “Convênio”. Os fatos, ou as métricas, a serem analisados foram “Total de publicações”, “Pontuação em publicações”, “Total de orientações”, “Total de disciplinas”, “Total de entrada e saídas das categorias” e “Proporção do corpo docente”.

Já o modelo para “Atividade Discente” considerou as análises realizadas utilizando as dimensões “Tempo”, “Convênio”, “Situação Discente” e “Orientador”. Os fatos que podem ser analisados utilizando o modelo são “Total discentes”, “Proporção entre matriculados e titulados”, “Média de meses para titulação” e “Mediana de meses para titulação”.

Para que a proposta fosse validade por um PPG, foram construídos *dashboards* que possibilitassem aos usuários os filtros propostos nos modelos, para avaliação por parte da coordenação do curso, dos dados trabalhados.

Os resultados obtidos nesse processo, são apresentados na próxima seção.

## RESULTADOS

Nesta Seção, são apresentados os resultados obtidos com a aplicação da metodologia proposta, para as questões de pesquisa colocadas (Seção Introdução).

### Itens de avaliação da CAPES

Como descrito na Seção 3, a regressão logística não tem como pressuposto a distribuição normal das variáveis, mas é necessário que as variáveis não apresentem multicolinearidade. Após a realização do teste Kolmogorov-Smirnov ( $p < 0,01$ ), verificou-se que as variáveis independentes não atenderam aos parâmetros de normalidade, razão pela qual o método de correlação de pontos de Spearman foi escolhido. Os resultados são apresentados na Tabela 2.

Já para o teste de multicolinearidade, nenhuma das variáveis apresentou valores que demonstrem que há relação entre elas, ficando todos os valores para Tolerância maior que 0,1 e os para VIF menores que 10.

**Tabela 2: Estatísticas descritivas e coeficiente de correlação de Spearman dos itens de avaliação com a Nota CAPES (Y)**

Variável	MD (DP)/Med	Coeficiente	Variável	MD (DP)/Med	Coeficiente
I12	3,74(1,118) /4,00	0,701*	I43	4,49(1,029) /5,00	0,491*
I51	3,81(0,952) /4,00	0,635*	I11	3,85(0,861) /4,00	0,487*
I53	3,63(1,184) /5,00	0,635*	I33	4,05(0,780) /4,00	0,482*

<b>I41</b>	3,99(0,773) /4,00	0,602*	I23	3,74(0,986) /4,00	0,443*
<b>I14</b>	3,42(1,117) /3,00	0,593*	I13	4,44(0,707) /5,00	0,432*
<b>I52</b>	3,44(1,130) /3,00	0,591*	I54	3,74(0,986) /4,00	0,401*
<b>I44</b>	3,82(0,962) /4,00	0,585*	I31	4,30(1,009) /5,00	0,385*
<b>I32</b>	3,88(0,798) /4,00	0,570*	I42	4,00(1,041) /4,00	0,366*
<b>I21</b>	3,92(1,038) /4,00	0,538*	I22	4,51(0,729) /5,00	0,200

MD: Média; DP: Desvio Padrão; Med: Mediana; \*: p-valor  $\leq 0,01$ . Elaborado pelo autor (2019)

Fonte: elaborado pelo autor (2020)

Com o objetivo de tentar uma otimização do modelo de regressão, utilizando um menor número de preditores, foram escolhidas, inicialmente, as variáveis independentes com coeficiente de correlação maior ou igual a 0,500 com a variável dependente.

Para a variável Nota CAPES, foram atribuídos os valores 0 ou 1. Os cursos que obtiveram nota final na avaliação da CAPES menor ou igual a 3 receberam o valor 0. Já os cursos com nota maior ou igual a 4 receberam valor 1.

Para o modelo nulo, criado para avaliação dos modelos obtidos a partir da regressão, a categoria 0 apresentou maior frequência e foi a atribuída para a variável dependente em todos os registros. Considerando um valor de corte de 0,500, neste caso, o modelo teria um percentual de acerto de 60,3%.

A Tabela 3 traz o quantitativo de valores observados e preditos para a variável dependente e o percentual de acertos de cada modelo. Os valores de acerto para os Modelos 2 e 3 são idênticos e que, a princípio, poderia indicar que adotar o Modelo 2 seria a decisão mais acertada, por adotar o modelo mais simples para predição dos valores.

No entanto, para o processo decisório, partindo da análise da correlação, verificou-se que a inclusão da variável I32 melhora a qualidade do modelo, se considerarmos os aumentos dos valores de verossimilhança, dos coeficientes de regressão, tanto de Cox & Snell quanto de Nagelkerke, e dos níveis de significância do teste de Hosmer & Lemeshow (Field, 2009).

**Tabela 3: Valores observados e previstos por modelo**

	Observados	Preditos		% acertos
		0	1	
<b>Modelo 1</b>	0	40	4	90,9
	1	7	22	75,9
	% geral de acertos			84,9
<b>Modelo 2</b>	0	41	3	93,2
	1	2	27	93,1
	% geral de acertos			93,2
<b>Modelo 3</b>	0	42	2	95,5
	1	3	26	89,7
	% geral de acertos			93,2

Fonte: elaborado pelo autor

Para cada modelo são listadas as variáveis independentes que compõem o modelo, os valores de  $\beta$  (beta) para cada variável e para a constante. São apresentados, também, os valores para erro padrão (E.P), para o teste Wald, graus de liberdade (G.L.), o p-valor para o teste Wald, do  $\text{Exp}(\beta)$ , além dos valores mínimos e máximos para o intervalo de confiança de  $\text{Exp}(\beta)$ .

**Tabela 4: Valores gerais dos modelos propostos**

		$\beta$	E.P.	Wald	G.L.	p-valor	Exp( $\beta$ )	95% I.C. Exp( $\beta$ )	
								Mín.	Máx
<b>Modelo 1</b>	I12	2,108	0,442	22,716	1	0,000	8,231	3,459	19,585
	Constante	-8,847	1,886	21,995	1	0,000			
<b>Modelo 2</b>	I12	2,264	0,581	15,206	1	0,000	9,619	3,083	30,012
	I41	2,579	0,819	9,929	1	0,002	13,190	2,651	65,616
	Constante	-20,288	4,947	16,822	1	0,000			
<b>Modelo 3</b>	I12	2,389	0,667	12,828	1	0,000	10,905	2,950	40,314
	I32	2,389	1,030	5,378	1	0,020	10,899	1,448	82,064
	I41	2,819	1,067	6,982	1	0,008	16,762	2,071	135,654
	Constante	-31,338	8,912	12,366	1	0,000			

$\beta$  – coeficiente; E.P. – Erro padrão; Wald – Coeficiente para o teste Wald; G.L. – Graus de Liberdade; Exp( $\beta$ ) – Razão de chances; I.C – Intervalos de confiança. Fonte: elaborado pelo autor

O valor de  $\beta$ , correspondentes aos coeficientes do modelo, foram submetidos ao teste de Wald. Pode-se dizer que o teste de Wald equivale ao teste T na regressão linear e tem como hipótese nula ( $H_0$ ) que o coeficiente  $\beta$  é igual a 0 (zero) e, como hipótese alternativa ( $H_1$ ), que o coeficiente é diferente de 0 (zero). É possível observar, na coluna p-valor, que todos os valores de  $\beta$  são menores que 0,05, levando a rejeição de  $H_0$ , permitindo afirmar que os coeficientes são diferentes de 0 (zero).

Exp ( $\beta$ ) indica a razão de chances de um determinado indivíduo pertencer a categoria de sucesso em relação aos indivíduos que estão na categoria de fracasso. Ou seja, a cada variação na variável independente, provoca um aumento na variável dependente em um percentual de “Exp( $\beta$ )” pontos. Nesse estudo, a título de exemplo, a cada ponto aumentado no valor da variável I41, a chance de um programa pertencer à categoria de sucesso aumenta em 16,76%.

A Tabela 5 apresenta valores que permitem a avaliação dos modelos gerados pela inserção das variáveis, passo a passo. A cada passo, o método “Forward” verifica se o modelo obtido é relevante e tem significância estatística, mantendo a variável inserida, caso seja. Essa verificação se dá pelo teste Qui-quadrado ( $X^2$ ), que tem como hipótese nula ( $H_0$ ) que o ajuste no modelo atual é igual ao ajuste do modelo anterior e, conseqüentemente, como hipótese alternativa ( $H_1$ ), que o ajuste no modelo atual é diferente do ajuste do modelo anterior. É possível verificar que todos os modelos apresentam diferenças significativas (p-valor < 0,01).

Além das diferenças nos ajustes entre os modelos de regressão, evidenciadas pelo teste  $X^2$ . A Tabela 5 traz, também, os valores para a verossimilhança de cada modelo. Os valores se referem à quantidade de informação que não é explicada pelo modelo. Ou seja, quanto menor o valor na coluna “Verossimilhança” maior é a quantidade de informação explicada por cada modelo. Portanto, é possível verificar que a verossimilhança entre os modelos sofre melhoria à medida que as variáveis são inseridas.

**Tabela 5: Valores para Qui-quadrado, Verossimilhança e coeficientes de regressão**

Modelo	$X^2$	p-valor	GL	Verossimilhança	$R^2$ Cox & Snell	$R^2$ Nagelkerke
--------	-------	---------	----	-----------------	----------------------	---------------------



1	44,465	0,000	1	53,631	0,456	0,617
2	60,767	0,000	2	37,328	0,565	0,764
3	68,999	0,000	3	29,096	0,611	0,827

Fonte: elaborado pelo autor

Os coeficientes de regressão ( $R^2$ ) de Cox & Snell e de Nagelkerke também apresentam aumento. É importante colocar que o coeficiente de regressão de Cox & Snell apresenta valor máximo de 0,750, ficando sempre abaixo do coeficiente de Nagelkerke, em que o valor máximo possível é 1 (FIELD, 2009).

Na Tabela 6 são apresentados os valores para o teste de Hosmer & Lemeshow. A hipótese nula ( $H_0$ ) é de que as categorias previstas correspondem às categorias observadas e a hipótese alternativa ( $H_1$ ) é de que as categorias observadas são diferentes das categorias previstas pelo modelo. Todos os modelos apresentam valores de significância (p-valor) maiores que 0,05 para o teste, o que permite admitir  $H_0$ , ou seja, as categorias previstas são iguais as categorias observadas, demonstrando a confiabilidade dos modelos.

**Tabela 6: Valores para o teste de Hosmer & Lemeshow**

	$X^2$	G.L.	p-valor
<b>Modelo 1</b>	0,127	2	0,938
<b>Modelo 2</b>	7,320	6	0,292
<b>Modelo 3</b>	3,917	8	0,865

Fonte: elaborado pelo autor

É possível concluir que o sucesso de um programa com curso de mestrado profissional, na avaliação quadrienal da CAPES, é melhor explicado pelo Modelo 3, composto pelas variáveis I12, I32 e I41, sendo que I41 apresenta um maior peso no modelo ( $\beta = 2,819$ ), seguida de I12 e I32, que apresentam o mesmo peso para o modelo ( $\beta = 2,389$ ). Abaixo, a função de probabilidade de um programa com mestrado profissional obter sucesso (nota maior ou igual a 4), caso sejam mantidos os itens de avaliação:

$$P(\text{Sucesso}) = \frac{1}{1 + e^{31,338 - 2,819 * I41 - 2,389 * I12 - 2,389 * I32}}$$

Considerando a Ficha de Avaliação da CAPES, utilizada no quadriênio 2013-2016, o item que apresenta maior peso na avaliação seria o “4.1 - Publicações qualificadas do Programa por docente permanente”, seguido dos itens “1.2 - Coerência, consistência e abrangência dos mecanismos de interação efetiva com outras instituições, atendendo a demandas sociais, organizacionais” e “3.2 - Qualidade dos trabalhos de conclusão produzidos por discentes e egressos”.

Com a proposta de uma nova ficha de avaliação, após a avaliação quadrienal, 2017-2020, um novo modelo poderá ser construído, para verificar se os mesmos itens são capazes de estimar a probabilidade de uma PPG ter sucesso (Capes, 2019). O Quadro 2 apresenta os novos itens de avaliação que englobará os itens atuais que foram considerados na construção do modelo proposto neste trabalho.

**Quadro 1: Comparação das variáveis I12, I32 e I41 na ficha anterior e na nova ficha de avaliação**

Ficha anterior (2013-2016)	Nova ficha (2017-2020)
----------------------------	------------------------

<b>1. Proposta do Programa</b>	1.2. Coerência, consistência e abrangência dos mecanismos de interação efetiva com outras instituições, atendendo a demandas sociais, organizacionais”	1. Programa	1.3. Planejamento estratégico do programa, considerando também articulações com o planejamento estratégico da instituição, com vistas à gestão do seu desenvolvimento futuro, adequação e melhorias da infraestrutura e melhor formação de seus alunos, vinculada à produção do conhecimento
<b>3. Corpo Docente</b>	3.2. Qualidade dos trabalhos de conclusão produzidos por discentes e egressos	2. Formação	2.1. Qualidade e adequação das teses, dissertações ou equivalente em relação às áreas de concentração e linhas de pesquisa do programa.
<b>4. Corpo Docente</b>	4.1. Publicações qualificadas do Programa por docente permanente	2. Formação	2.4. Qualidade das atividades de pesquisa e da produção intelectual do corpo docente no programa.

**Fonte: Elaborado pelo autor**

Considerando a nova ficha de avaliação da CAPES, uma análise comparativa entre os itens da ficha de avaliação utilizada no último processo de avaliação e a nova ficha proposta foi realizada e apresentada no Quadro 1. Assim, é possível verificar que os itens de avaliação identificados como preditores para o modelo de regressão, proposto neste trabalho, permanecem na nova ficha de avaliação, sendo incorporados por outros itens, em especial, pelo item “Formação” (CAPES, 2019).

### **Modelos e indicadores para pós-graduação**

Os modelos dimensionais propostos para apoio à tomada de decisão são descritos nesta seção. Os modelos são representações em alto nível de abstração para os cubos OLAP. Os indicadores são obtidos pela combinação dos parâmetros passados para as dimensões de análise.

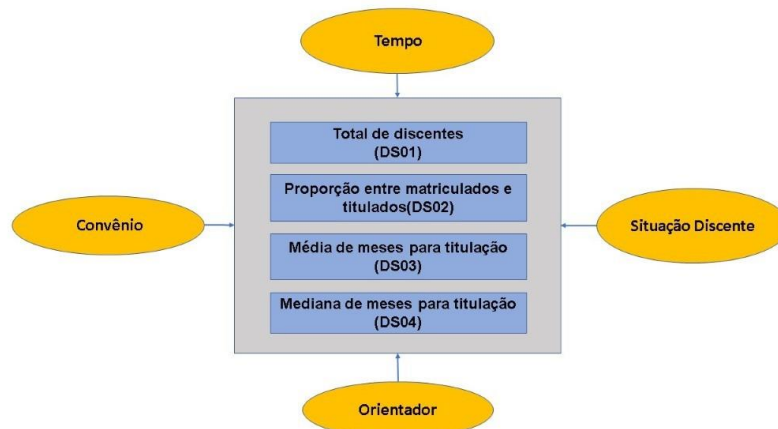
#### *Modelos para Fluxo Discente*

A Figura 2 apresenta o modelo dimensional proposto para gestão do fluxo discente. O indicador (ou métrica) DS01 permite a análise do total de discentes. Assim como o DS02 permite a análise da proporção de discentes matriculados e titulados. Já os indicadores DS03 e DS04 permitem a análise da média e mediana dos meses até a titulação. As análises podem ser realizadas pelas dimensões “situação discente”, “tempo”, “convênio” e “orientador”.

A dimensão “situação discente” permite a análise com atribuição dos parâmetros “matriculado”, “titulado”, “desligado”, “abandono”. A dimensão “tempo” permite a análise informando o ano e semestre desejado. A dimensão “convênio” permite a análise escolhendo o convênio a ser analisado. Já a dimensão “orientador” permite a análise por docente orientador.

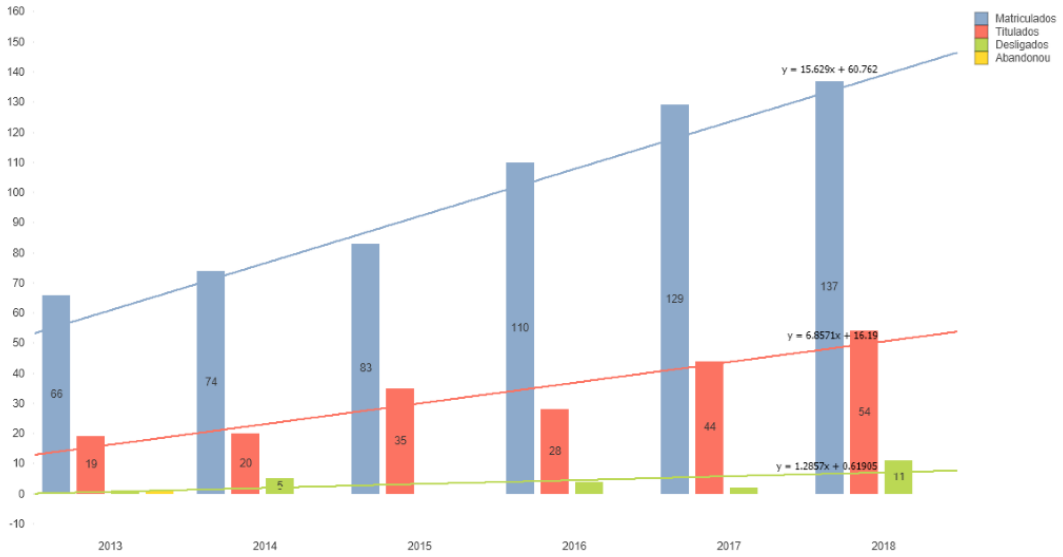
Assim, de maneira geral, o modelo permite a análise de indicadores informando parâmetros para uma dimensão ou pela combinação entre as dimensões. A Figura 3 apresenta total de discentes (DS01) matriculados, titulados, desligados e em abandono, por ano. Não foram atribuídos valores para as dimensões “tempo” e “convênio”, ou seja, o indicador, dado como exemplo, é o número de discentes por situação e ano. Sendo assim, a cada parâmetro informado, o gestor pode navegar pelo modelo, gerando novas análises dos indicadores e melhorando o processo de tomada de decisão.

**Figura 2: modelo para fluxo discente**



Fonte: elaborado pelo autor

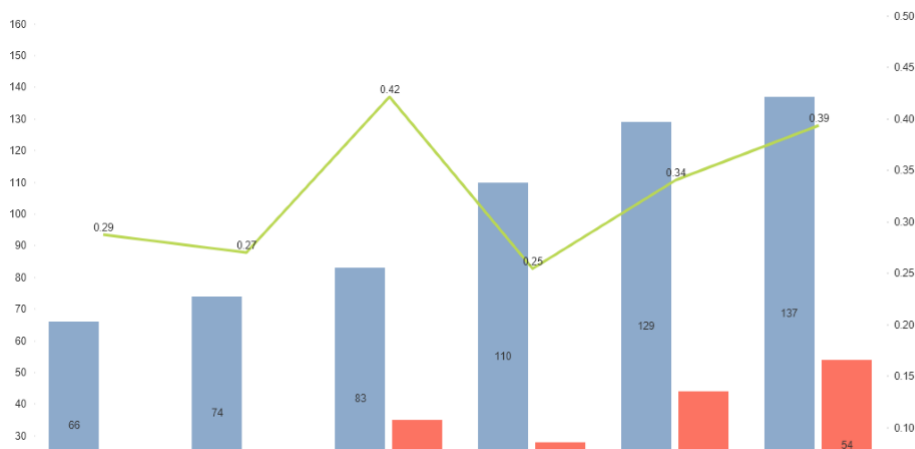
**Figura 3: exemplo de indicador obtido com o modelo “Total de discentes”**



Fonte: elaborado pelo autor

A Figura 4 apresenta a análise de discentes titulados versus matriculados por ano. A linha apresentada traz a proporção entre discentes titulados e matriculados. Não foram passados como parâmetros valores para as dimensões “convênio” e “orientador”.

**Figura 4: Exemplo de indicador obtido com o modelo “Proporção entre matriculados e titulados”**

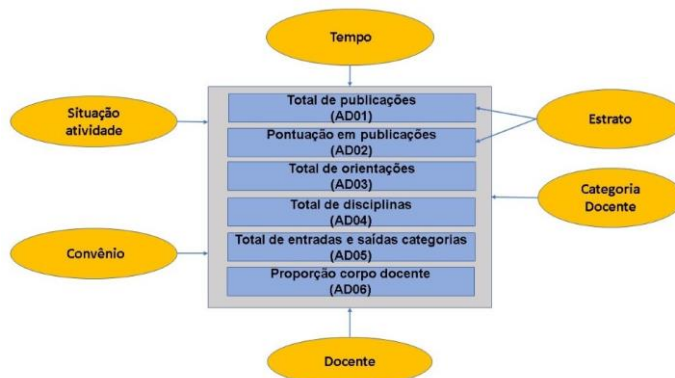


Fonte: Elaborado pelo autor

### Modelos para Atividade e Fluxo Docentes

A Figura 5 apresenta o modelo dimensional obtenção dos indicadores relacionados à atividade e fluxo docentes.

**Figura 5: modelo dimensional para atividade e fluxo docente**

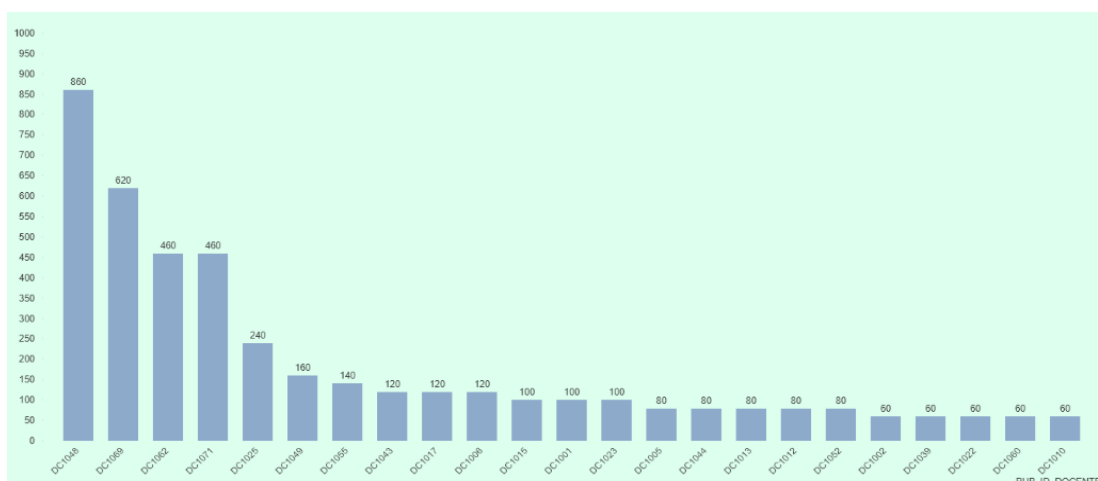


Elaborado pelo autor

O indicador AD01 permite a análise do total de publicações do corpo docente pelas dimensões “tempo”, “docente”, “categoria docente”, “convênio” e “estrato”. Já o AD02, utilizando as mesmas dimensões de análise, permite a análise da pontuação das publicações. Já os indicadores AD03 e AD04 possibilitam a análise do total de orientações e disciplinas, respectivamente. A análise pode ser feita por ano, por docente, por categoria docente e por convênio.

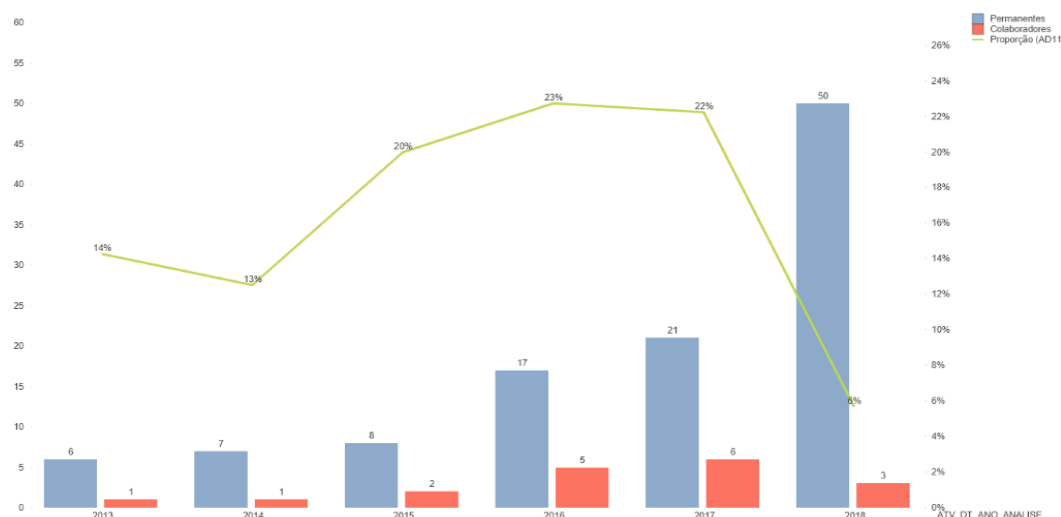
A dimensão “situação discente” permite a análise com atribuição dos parâmetros “matriculado”, “titulado”, “desligado”, “abandono”. A dimensão “tempo” permite a análise informando o ano e semestre desejado. A dimensão “convênio” permite a análise escolhendo o convênio a ser analisado. Já a dimensão “orientador” permite a análise por docente orientador.

**Figura 6: exemplo de indicador obtido com o modelo “Pontuação em publicações”**



Fonte: elaborado pelo autor

**Figura 7: Exemplo de indicador obtido com os modelos “Total de orientações” e “Total de disciplinas”**



Fonte: elaborado pelo autor

Os modelos aqui propostos se mostraram capazes de atender também às simulações para atividades de aulas e orientações, em andamento e concluídas, realizadas pelo corpo docente, listando docente a docente, calculando, após a inserção dos parâmetros, quais atividades os docentes deveriam realizar para atingir os objetivos do Programa.

Outro exemplo de indicador obtido é o total de publicações (AD01), considerando a dimensão “estrato” e tempo para filtro. A Figura 6 traz exemplo de indicador obtido pelo indicador AD02 (pontuação em publicações), com parâmetros “A1”, “A2” e “B1” para a dimensão “estrato”, por docente. Aqui, os docentes foram ranqueados pela pontuação obtida. Não foram atribuídos parâmetros para as demais dimensões.

Por fim, a Figura 7 traz exemplo de indicador obtido pelos modelos AD03 e AD04 (total de orientações e total de disciplinas), por categoria docente e ano.

## CONCLUSÃO

O objetivo desse trabalho foi propor modelos e indicadores para apoio à decisão para PPG com cursos de mestrado profissional da Área de “Administração Pública e de Empresas, Ciências Contábeis e Turismo”. Na busca por respostas às questões de pesquisa colocadas (QP1 e QP2), métodos estatísticos e relacionados à sistemas de apoio à decisão (SAD/BI) foram utilizados. Para a QP1, foi possível verificar que três itens (I12, I32, I41) da ficha de avaliação da CAPES contribuem fortemente para o sucesso de um PPG com mestrado profissional. Para a QP2, modelos dimensionais foram criados como possível solução para apoio à gestão. Além disso, foram criados painéis com dados de um PPG para avaliação dos modelos.

Este trabalho pode contribuir para o entendimento do funcionamento dos processos de avaliação da CAPES e como os itens utilizados na avaliação podem influenciar a nota final. Além disso, os modelos dimensionais propostos foram validados com dados de um PPG e foi possível verificar que a visualização de dados pode apoiar o processo de tomada de decisão.

Em trabalhos futuros, tem-se o objetivo de avaliar os indicadores obtidos com índices bibliométricos amplamente utilizados. Além disso, pretende-se realizar estudos em nível institucionais.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Beal, A. (2012). *Gestão estratégica da informação: como transformar a informação ea tecnologia da informação em fatores de crescimento e de alto desempenho nas organizações*. Atlas.
- Borges, V. A., Nogueira, B. M., & Barbosa, E. F. (2016). A multidimensional data model for the analysis of learning management systems under different perspectives. *2016 IEEE Frontiers in Education Conference (FIE)*, 1–8. <https://doi.org/10.1109/FIE.2016.7757743>
- Burigo, C. C. D., Jacobsen, A. de L., & Wiggers, L. (2016). Mestrado Profissional em Administração Universitária: Desafios e Perspectivas no Processo da Gestão Universitária. *Revista de Gestão e Tecnologia*, 6(2), 68–78.
- Cecilia, M., & Lopes, B. (2015). Administrando o Imensurável : Uma Crítica às Organizações Acadêmicas. *Cadenos EBAPE.BR*, 40–51.
- Comim, J., Ensslin, S. R., & Valmorbidia, S. M. I. (2017). Importância e Uso de Indicadores de Desempenho para Gestão Universitária, Segundo a Percepção de Gestores Públicos. *Congresso Internacional de Desempenho Do Setor Público, 2014*, 2484–2507.
- Dawson, S., Gašević, D., Siemens, G., & Joksimović, S. (2014). Current State and Future Trends: A Citation Network Analysis of the Learning Analytics Field. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, 231–240. <https://doi.org/10.1145/2567574.2567585>
- Field, A. (2009). *Descobrendo a estatística usando o SPSS*.
- Gutiérrez, F., Seipp, K., Ochoa, X., Chiluzza, K., De Laet, T., & Verbert, K. (2020). LADA: A learning analytics dashboard for academic advising. *Computers in Human Behavior*, 107(November), 0–1. <https://doi.org/10.1016/j.chb.2018.12.004>
- Jayakody, J., & Perera, I. (2016). Enhancing competencies of less-able students to achieve learning outcomes: Learner aware tool support through Business intelligence. *Proceedings of 2016 IEEE International Conference on Teaching, Assessment and*

- Learning for Engineering, TALE 2016, December*, 154–160.  
<https://doi.org/10.1109/TALE.2016.7851787>
- Kimball, R. (1997). A dimensional modeling manifesto. *DBMS*, 10(9), 58–70.
- Kimball, R., & Ross, M. (2013). *The data warehouse toolkit: The definitive guide to dimensional modeling*. John Wiley & Sons.
- Laudon, K., & Laudon, J. (2011). *Sistemas de Informações Gerenciais (9ª)*. Pearson Education do Brasil.
- Mesquita, P. S. B., & Nogueira, R. T. (2015). Um modelo estatístico para gestão de programas de pós-graduação. *Gestão Da Produção, Operações e Sistemas*, 10(2), 173–186. <https://doi.org/10.15675/gepros.v10i2.1218>
- Nieto, Y. V., Diaz, V. G., & Montenegro, C. E. (2016). Academic decision making model for higher education institutions using learning analytics. *2016 4th International Symposium on Computational and Business Intelligence (ISCBI)*, 27–32.  
<https://doi.org/10.1109/ISCBI.2016.7743255>
- Rodrigues, A. C. D. A. L., & Villardi, B. Q. (2017). Formação do Docente para a Gestão Universitária: Uma Análise Indutiva dos Professores Gestores da Pós-Graduação Stricto Sensu da UFRRJ. *Revista Foco*, 10(2), 208.
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., & Ferguson, R. (2011). Open Learning Analytics : an integrated & modularized platform. *Knowledge Creation Diffusion Utilization*, 1–20.
- Turban, E., Rainer, R. K., & Potter, R. E. (2005). *Administração de tecnologia da informação: teoria e prática*. Elsevier.
- Turban, E., Sharda, R., Aronson, J. E., & King, D. (2009). *Business Intelligence: um enfoque gerencial para a inteligência do negócio*. Bookman Editora.
- Villegas-Ch, W., Palacios-Pacheco, X., & Luján-Mora, S. (2020). A business intelligence framework for analyzing educational data. *Applied Sciences (Switzerland)*, 12(14), 1–21.  
<https://doi.org/10.3390/su12145745>