



## **Análise discriminante e probabilidade de evasão dos alunos de engenharia de produção na UFSJ**

### **Discriminant analysis and evasion probability of production engineering students at UFSJ**

Luiz Flávio Felizardo<sup>1</sup>  
Vanessa de Souza Silva<sup>2</sup>  
Gisleine do Carmo<sup>3</sup>  
Daniel Rocha Gualberto<sup>4</sup>  
Luiz Marcelo Antonialli<sup>5</sup>

#### **Resumo**

A evasão estudantil no ensino superior é um problema que afeta diversas instituições de ensino, podendo causar desperdícios acadêmicos, sociais e econômicos. Neste contexto, o presente estudo teve como objetivo identificar as características que mais discriminaram o grupo de alunos evasores dos não evasores no curso de Engenharia de Produção da Universidade Federal de São João del-Rei (UFSJ). Para alcançar tais resultados foram extraídos da base de dados da instituição informações a respeito dos alunos que ingressaram

---

<sup>1</sup> Doutorado em Administração, Universidade Federal de São João del-Rei (UFSJ), Praça Frei Orlando, 170, Centro, CEP: 36307-352, São João del-Rei-MG. E-mail: [felizardo@ufsj.edu.br](mailto:felizardo@ufsj.edu.br)

Orcid: <https://orcid.org/0000-0002-6324-7313>

<sup>2</sup> Mestre em Propriedade Intelectual e Transferência de Tecnologia para a Inovação (PROFNIT), Praça Frei Orlando, 170, Centro, CEP: 36307-352, São João del-Rei-MG. E-mail: [vanessaecd@gmail.com](mailto:vanessaecd@gmail.com)

Orcid: <https://orcid.org/0000-0003-0903-9254>

<sup>3</sup> Mestre em Administração, Universidade Federal de Lavras (UFLA), Trevo Rotatório Professor Edmir Sá Santos, s/n, CEP: 37203-202, Lavras-MG. E-mail: [gisleinecarmo95@gmail.com](mailto:gisleinecarmo95@gmail.com)

Orcid: <https://orcid.org/0000-0002-7260-5063>

<sup>4</sup> Mestre em Ciência da Computação, Universidade Federal de São João del-Rei, Praça Frei Orlando, 170, Centro, CEP: 36307-352, São João del-Rei-MG. E-mail: [daniel.gualberto@ufsj.edu.br](mailto:daniel.gualberto@ufsj.edu.br)

Orcid: <https://orcid.org/0000-0001-6907-6247>

<sup>5</sup> Doutor em Administração, Universidade Federal de Lavras (UFLA), Trevo Rotatório Professor Edmir Sá Santos, s/n, CEP: 37203-202, Lavras-MG. E-mail: [lmantonialli@ufla.br](mailto:lmantonialli@ufla.br)

Orcid: <https://orcid.org/0000-0003-1220-6164>

no curso mencionado entre os anos de 2014 e 2018. Posteriormente foi realizada a exportação destes para o software *Statistical Package for the Social Sciences* (SPSS) e utilizada a técnica de análise discriminante. Os resultados da pesquisa demonstraram que existe uma considerável quantidade de evasões acontecendo no curso analisado. Foi possível ainda, encontrar uma função discriminante com precisão real de mais de 90% que permite a preemptiva identificação de alunos mais propensos a abandonar o curso e conseqüentemente facilitar o trabalho da gestão ou da coordenação para que possam agir de forma diferenciada com tais alunos, a fim de evitar desistências.

**Palavras-chave:** Evasão. Ensino Superior. Engenharia de Produção. Análise Discriminante.

### **Abstract**

Student dropout in higher education is a problem that affects several educational institutions, and can cause academic, social and economic waste. In this context, the present study aimed to identify the characteristics that most discriminate the group of evasers from the group of non evasers in the Production Engineering course at the Federal University of São João del-Rei (UFSJ). To achieve such results, information regarding the students who entered the mentioned course between the years 2014 and 2018 was extracted from the institution's database. Subsequently, these were exported to the Statistical Package for the Social Sciences (SPSS) software and the discriminant analysis technique was used. The results of the research showed that there is a considerable amount of dropouts happening in the analyzed course. It was also possible to find a discriminant function with a real accuracy of over 90%, which allows the preemptive identification of students more likely to drop out of the course and consequently facilitate the work of management or coordination so that they can act differently with such students in order to avoid dropouts.

**Keywords:** Dropout. Higher Education. Engenharia de Produção. Discriminant Analysis.

### **Introdução**

A evasão é um fenômeno complexo que deve ser compreendido no contexto socioeconômico, político e cultural, no sistema educacional e nas instituições de ensino. Seu conceito em sentido amplo pode ser entendido como estudantes que iniciam, mas não concluem os cursos aos quais estavam matriculados. Esse evento pode ser um processo de exclusão determinado por fatores e variáveis internas e externas às instituições de ensino

conforme mencionado por Fritsch, Rocha e Vitelli (2015). De acordo com Silva Filho, et.al, (2007) a evasão estudantil no ensino superior é um problema que afeta o resultado dos sistemas educacionais, além de causar desperdícios acadêmicos, sociais e econômicos. É certamente, um dos mais comuns problemas que afligem as instituições de ensino em geral. Nas instituições federais este evento propicia que recursos públicos investidos não tenham seu devido retorno e no setor privado pode representar perda de receitas. Pode ainda significar uma fonte de ociosidade de professores, funcionários, equipamentos e espaço físico.

Saccaro, França e Jacinto (2019) apontam diversos estudos que revelam a existência de maiores taxas de evasão nos cursos das áreas de Ciência, Matemática e Computação e de Engenharia, Produção e Construção. Eles ressaltam que esses dados geram preocupação, uma vez que cursos das áreas de setores como engenharia e ciências naturais estão diretamente relacionados com parte significativa da geração de inovações tecnológicas. A baixa quantidade de pessoal qualificado nesta área pode impactar negativamente na economia brasileira, fazendo com que a sociedade, de uma forma geral, não possa contar com as externalidades positivas que poderiam ser geradas por esses profissionais.

O problema da evasão escolar nos cursos de graduação nas áreas de engenharia possui causas distintas dentro de uma série de fatores situados em diversos contextos. Estes fatores podem estar associados com aspectos sociais e econômicos, outros com características pessoais e outros ainda com aspectos metodológicos. Em conjunto apresentam aspectos que podem ser evidenciados no processo de ensino-aprendizagem, tais como as dificuldades pessoais ou as necessidades de conhecimento e habilidades prévias segundo Rigo, Barbosa e Cambruzzi (2011). Neste contexto, o presente estudo tem como objetivo identificar as características que separam o grupo de alunos evasores dos não evasores no curso de Engenharia de Produção da Universidade Federal de São João del-Rei (UFSJ) encontrando a função discriminante que calcula o fator (*score*) e a probabilidade de evasão.

## Metodologia

Neste trabalho foi utilizada a metodologia quantitativa baseada em princípios epistemológicos positivistas e ontologia objetiva. A partir da organização e tabulação dos dados foram utilizadas técnicas e testes estatísticos para orientar a análise e a apresentação dos resultados de acordo com Martins e Theóphilo (2016). Com relação aos objetivos, conforme mencionado, a pesquisa tentou identificar as características que mais discriminam o grupo de alunos evasores dos não evasores no curso de Engenharia de Produção da

Universidade Federal de São João del-Rei (UFSJ) para encontrar uma função capaz de calcular a probabilidade de evasão dos estudantes no curso mencionado, dessa forma, segundo Kauark, Manhães e Medeiros (2010) esta pesquisa pode ser considerada uma pesquisa de caráter descritivo.

Inicialmente foram extraídos os dados dos alunos da base de dados da instituição, que utiliza diversas tabelas organizadas de forma relacional. Para essa extração foi utilizada a linguagem *Structured Query Language* (SQL) e uma grande quantidade de linhas de códigos e comandos do SQL foi demandada até que se conseguisse os dados desejados em uma planilha do Microsoft® Excel para organização e filtragem adicional. Já no Excel foram removidos dados adicionais que não poderiam ou não deveriam ser utilizados nessa pesquisa, como alunos de outros cursos, variáveis não métricas dentre outros. Ao final dessa etapa foram selecionados apenas os alunos que foram admitidos pelo Sistema de Seleção Unificada (SISU) no curso de Engenharia de Produção da UFSJ a partir de 2014, ano em que a UFSJ adotou o SISU de forma completa (UFSJ, 2013), até dezembro de 2018. Esse recorte foi importante, pois as notas do Exame Nacional do Ensino Médio (ENEM) dos estudantes também foram incluídas na análise. Após a realização dessas etapas foi possível identificar uma população de 276 matrículas. Todos os dados foram coletados após o meio do ano de 2019, porém todas as informações são referentes aos dados dos alunos precisamente antes de iniciarem os estudos no ano de 2019, ou seja, foi utilizado um *Backup* dos dados dos alunos exatamente antes do início do primeiro período de 2019 para que os dados ficassem homogêneos e não fossem coletadas informações de alunos que fizeram apenas parte do primeiro do ano de estudos.

Em seguida foi feita a exportação de todas essas informações dos 276 matriculados para o software originalmente conhecido como *Statistical Package for the Social Sciences* (SPSS), por meio das funcionalidades de salvar em arquivos de textos tabulados do Excel e ler dados de texto do SPSS. A partir daí foi utilizada a técnica de análise discriminante do software para investigação dos dados. A aplicação da técnica envolveu todas as variáveis métricas de modo a definir se era possível encontrar uma função discriminante canônica válida, capaz de separar e identificar pelos outros atributos alunos com tendências de evasão. Para verificar a validade da função foi analisada sua correlação canônica baseada nos autovalores, sua significância de acordo com Wilks' Lambda e seu grau de precisão. Após validação, esse trabalho também apresenta os pesos de cada coeficiente, a função final e elabora um termômetro para facilitar a visualização da posição dos grupos de alunos no termômetro. A listagem a seguir apresenta o sumário das etapas metodológicas utilizadas e mencionadas no texto.

### Sumário das Etapas Metodológicas:

- Extração dos dados da base da instituição;
- Organização e filtragem dos dados no Excel;
- Exportação das informações para o SPSS;
- Execução da análise discriminante no SPSS;
- Verificação da correlação canônica baseada nos autovalores;
- Checagem da significância da função pelo Wilks' Lambda;
- Análise do grau de precisão do modelo;
- Apresentação da função discriminante canônica padronizada encontrada;
- Apresentação dos coeficientes da função discriminante canônica final;
- Elaboração do termômetro de evasão.

Com relação às etapas de validação da função discriminante é importante mencionar que a correlação canônica baseada nos autovalores precisa estar o mais próximo possível de 1, pois de acordo com Hardoon (2004) isso indica que os resultados dos testes padronizados e os resultados do teste de aptidão estão positivamente correlacionados. Ademais, o pesquisador precisa ainda determinar se o modelo canônico captura suficientemente a relação entre os conjuntos de variáveis dependentes e independentes para justificar a interpretação. Conforme Sherry (2005), o método mais comum utilizado para este fim é o Wilks' Lambda, que também foi utilizado aqui, nele, a denominada significância apresenta o valor associado à probabilidade dos resultados da amostra e quanto mais próxima de zero mais relação existe entre os conjuntos. E, por fim, o modelo precisa acertar o grupo para o qual cada elemento pertence a fim de ser considerado um bom modelo. Para mensurar esse nível de precisão foi utilizada a acurácia da função, que contabiliza as taxas de acertos de cada grupo e calcula ponderadamente a precisão geral da função e retorna o resultado em porcentagem.

### Análise e Resultados

Conforme mencionado no tópico anterior, este trabalho utilizou-se de uma população de 276 alunos, todos os aprovados e selecionados no SISU para o curso de Engenharia de Produção da UFSJ nos anos de 2014, 2015, 2016, 2017 e 2018. Dos 276 matriculados é possível identificar que 83 alunos (30,1%) evadiram do curso em algum momento e 193 (69,9%) não evadiram até o instante da coleta, a figura 1 apresenta esses dados de forma gráfica para facilitar a visualização. O valor próximo de 30% em evasões é um valor

considerável, que pode aumentar ainda mais se for levado em consideração que a maioria dos alunos da população ainda não concluíram o curso.

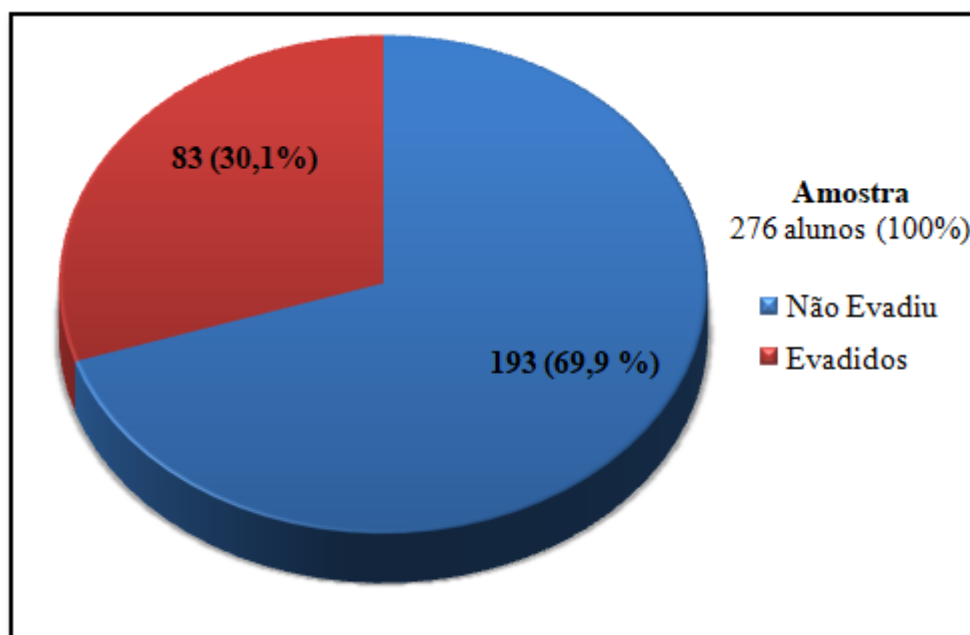


Figura 1 – População da pesquisa

A alternativa proposta por esse trabalho é identificar, com base na análise das características encontradas no banco de dados dos alunos, quais seriam aqueles com maior probabilidade de evasão para que a gestão da instituição ou a própria coordenação do curso possa se orientar e executar atividades de forma a prevenir tais desistências. Na função discriminante deste estudo foram utilizados dezessete atributos dos alunos como variáveis independentes e a variável evasão como a variável de grupo. As variáveis independentes são: idade, ano de admissão, tempo entre ensino médio e a admissão, quantidade de trancamentos, coeficiente de rendimento, média de presença nas aulas, percentual de aprovação nas disciplinas, média da avaliação dada pelos alunos para os docentes, notas para cada uma das provas do ENEM, duas médias diferentes para as notas do ENEM e a colocação no SISU para o curso ao ser admitido. Na figura 2 é possível observar alguns exemplos de alunos e suas respectivas variáveis, cada aluno está apresentado em uma linha do exemplo e os dois em destaque são alunos que evadiram do curso.



Aluno Evadiu	Idade	Ano de Admissão	Tempo Entre Ensino Médio	Quantidade de Trancamentos	Coefficiente de Rendimento (CR)	Média de Presença	Quantidade de Disciplinas Cursadas	Percentual de Aprovação	Avaliação para os Docentes	Nota Enem - Redação	Nota Enem - Linguagens	Nota Enem - Humanas	Nota Enem - Natureza	Nota Enem - Matemática	Nota Enem - Sem Pesos	Nota Enem - Com Pesos	Colocação do Candidato no Curso
0	26	2016	6	0	66	90	59	86	70	720	663	682	588	810	692	689	16
0	25	2016	5	0	64	91	60	85	49	680	658	681	621	821	692	697	8
0	23	2016	2	0	74	88	40	98	80	760	595	722	684	744	701	686	22
0	22	2016	1	0	84	94	42	100	74	940	573	636	658	749	711	683	27
0	21	2016	1	0	87	97	38	100	91	960	594	644	646	758	720	690	15
0	20	2016	0	2	64	86	30	83	81	820	621	631	604	808	697	687	21
1	21	2016	0	0	19	64	10	20	80	900	576	660	649	756	708	682	30
1	20	2016	0	0	42	72	20	60	91	720	615	665	658	789	689	688	18
0	27	2016	6	0	75	94	54	100	95	880	645	617	569	787	700	682	2

Figura 2 – Alguns exemplos de dados utilizados

Cada uma das variáveis é apresentada a seguir:

- **Aluno Evadiu** – Variável dependente, ou também chamada de variável de grupo e variável categórica. Indica se o aluno já evadiu da instituição ou não com os valores 1 e 0 respectivamente. Essa variável é a peça chave para verificar as diferenças dos grupos de evasores e não evasores por meio da análise discriminante uma vez que seu uso é apropriado quando se sabe o perfil geral de cada grupo;
- **Idade** – Apresenta a idade do discente na data imediatamente antes do início do 1º semestre letivo de 2019;
- **Ano de Admissão** – Aqui é o ano de entrada do aluno na instituição, o valor dessa variável será sempre um dos anos entre 2014 a 2018, pois este foi o período selecionado para esta pesquisa, ela é considerada uma variável métrica, pois para a função discriminante ela representa basicamente o tempo de curso que o aluno teve para evadir uma vez dentro da instituição;
- **Tempo Entre Ensino Médio** – Essa variável é a quantidade de anos que o aluno levou entre concluir seu ensino médio até a entrada no curso analisado. Exemplo, se o aluno terminou o ensino médio e logo em seguida entrou no curso, o valor dessa variável para este aluno será zero. Se o aluno fez um ano de cursinho pré-vestibular, por exemplo, o valor será um, e assim por diante;

- **Quantidade de Trancamentos** – Aqui é resgatado do banco de dados a quantidade de semestres que o aluno trancou sua matrícula durante o curso, até o momento da coleta;

**Coefficiente de Rendimento (CR)** – Conforme artigo 23 da resolução acadêmica da UFSJ (2018), o coeficiente de rendimento mede o grau de aproveitamento discente e consiste no somatório dos produtos da nota obtida pela carga horária de cada unidade curricular cursada, dividido pela carga horária total cursada, conforme a seguinte equação:

$$CR = \frac{\sum_{i=1}^N N_i \times CH_i}{\sum_{i=1}^N CH_i}$$

onde  $N_i$  é a nota obtida pelo discente na unidade curricular  $i$ , independente de aprovação ou reprovação,  $CH_i$  é a carga horária correspondente à unidade curricular  $i$  e  $N$  é o número de unidades curriculares consideradas de forma que:

I – são consideradas todas as disciplinas obrigatórias, optativas e em bloco, e as unidades curriculares do tipo Estágio e Trabalhos Acadêmicos;

II – disciplinas eletivas são contabilizadas até atingir a carga horária prevista no PPC, tendo como critérios de prioridade, em ordem decrescente, primeiro a nota final obtida e depois a carga horária da disciplina.

Para os discentes que possuam conceitos em seus históricos, são atribuídas notas conforme o que se segue: A = 10; B = 8; C = 6; D = 3; E = 0; SUF = 10; INS = 3.

- **Média de Presença** – Variável que descreve a média de presença do aluno nas disciplinas cursadas. Vale lembrar ainda que, para os cursos presenciais, é exigida, além de nota, frequência mínima de 75% (setenta e cinco por cento) para aprovação nas unidades curriculares do tipo disciplina, conforme normas e legislação vigentes;
- **Quantidade de Disciplinas Cursadas** – Aqui é contabilizado o total de disciplinas que o aluno cursou, independente se foi aprovado ou não, apenas o total é considerado, mesmo que ele tenha desistido da disciplina e evadido enquanto a cursava;
- **Percentual de Aprovação** – Como complemento da variável anterior, nessa variável é feito o cálculo do percentual de disciplinas que cada aluno teve em



aprovações. Para aprovação em uma unidade curricular, é obrigatória a obtenção de nota final igual ou superior a 60% além da frequência mínima em 75%, conforme já mencionado;

- **Avaliação sobre os Docentes** – Essa variável reflete a média das notas que os discentes atribuem para os docentes com o qual tiveram contato no semestre anterior. A escala vai até 100 e os alunos avaliam os docentes durante a matrícula de cada semestre. São diversos tópicos avaliados para cada docente e as notas são atribuídas através de um questionário de respostas obrigatórias que todos os alunos devem responder antes de realizarem a matrícula para o próximo semestre letivo. A figura 3 apresenta as perguntas que os discentes devem responder, onde a nota 1 representa que o professor foi muito aquém do que expressa o referencial mínimo de qualidade e a nota 5 ele foi muito além;

22.2.1. Quanto ao trabalho do docente					
Disciplina:	[REDACTED]				
Professor(a):	[REDACTED]				
	1	2	3	4	5
Demonstra respeito com os alunos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Demonstra interesse e cooperação na aprendizagem do estudante, valorizando os seus questionamentos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
É claro e preciso na comunicação do conteúdo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Demonstra conhecimento do conteúdo da Unidade Acadêmica	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Amplia o conteúdo da disciplina, ilustrando as aulas com resultados de pesquisa e/ou experiências profissionais	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Demonstra planejamento das aulas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Utiliza metodologias, técnicas e recursos compatíveis com os objetivos de ensino-aprendizagem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Aproveita adequadamente o tempo da aula	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Indica a busca de material impresso, referências disponíveis na Biblioteca e outras fontes de pesquisa	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Aplica instrumentos de avaliação (ex: provas, seminários, trabalhos etc.) correspondentes aos conhecimentos desenvolvidos na Unidade Acadêmica e em número adequado	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Retoma e discute os resultados das avaliações	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Articula o conteúdo da disciplina com a formação geral e/ou profissional do estudante	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Estimula a reflexão e crítica sobre os aspectos sociais, e/ou científicos, e/ou tecnológicos, e/ou políticos, e/ou econômicos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

Figura 3 – Avaliação docente solicitada para preenchimento dos discentes

- **Nota ENEM - Redação** – Esta e as próximas variáveis apresentam informações dos alunos com respeito a conhecimentos adquiridos previamente a sua entrada na instituição, e estão ligadas as notas do ENEM e sua colocação no processo seletivo do SISU. A primeira variável, como o nome sugere, é a nota do aluno para a

prova de redação do ENEM e pode variar de 0 até 1000;

- **Nota ENEM - Linguagens** – Nota obtida pelo aluno na prova do ENEM em Linguagens, Códigos e suas Tecnologias, que envolve componentes curriculares como Língua Portuguesa, Língua Estrangeira e outros;
- **Nota ENEM - Humanas** – Nota obtida pelo aluno na prova do ENEM em Ciências Humanas e suas Tecnologias, que se refere a componentes curriculares como História, Geografia, Filosofia e Sociologia;
- **Nota ENEM - Natureza** – Nota obtida pelo aluno na prova do ENEM em Ciências da Natureza e suas Tecnologias envolvendo os componentes curriculares como Química, Física e Biologia;
- **Nota ENEM - Matemática** – Nota obtida pelo aluno na prova de Matemática e suas Tecnologias do ENEM;
- **Nota ENEM – Sem Pesos** – Essa nota é referente a média simples de todas as 5 notas mencionadas anteriormente.
- **Nota ENEM – Com Pesos** – Para o cálculo dessa nota é levado em consideração o peso de cada disciplina, ou seja, é a média ponderada das 5 notas das provas do ENEM, essa média foi considerada pois cada curso pode apresentar pesos diferentes para cada prova; No caso de Engenharia de Produção os pesos são, Redação possui peso 1, Ciências da Natureza e suas Tecnologias peso 3, Ciências Humanas e suas Tecnologias com peso 1, Linguagens, Códigos e suas Tecnologias peso 3 e Matemática e suas Tecnologias possui peso 3.
- **Colocação do Candidato no Curso** – Por fim, foi incluída também a colocação que cada aluno obteve no processo seletivo quando foi selecionado para o curso, e é exatamente isso que esta variável representa.

Com todas as variáveis devidamente explicadas inicia-se o momento de análise da execução da função discriminante no software SPSS. Inicialmente verifica-se a tabela de autovalores ou *Eigenvalues* em inglês, de forma a certificar que o índice de correlação canônica está aceitável, quanto mais próximo de 1 melhor, e no caso deste trabalho foi encontrado o valor de 0,869, o que é considerado um valor alto conforme mencionado na metodologia. Após essa análise também foi realizada a checagem da significância da função pelo Wilks' Lambda gerado pelo software. Nesse valor o modelo também passou no teste ao apresentar uma significância de 0,000, ou seja, significativo a 1%. Esses e demais valores relacionados podem ser observados na Figura 4, retirada diretamente do próprio software.

**Eigenvalues**

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	3,078 <sup>a</sup>	100,0	100,0	,869

a. First 1 canonical discriminant functions were used in the analysis.

**Wilks' Lambda**

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,245	380,236	7	,000

**Figura 4 – Correlação Canonica e Wilks' Lambda**  
 Fonte: Elaborado pelos autores (2020)

Com relação a precisão do modelo, a figura 5, também captada da saída do software SPSS, apresenta a classificação dos resultados. A tabela mostra que dos 193 alunos que não evadiram, o modelo classificou 188 corretamente (97,4%) e apenas 5 (2,6%) como evadidos, por terem fortes características de evasão. É importante frisar que o modelo considera como evadido os alunos que tem um *score* que representa mais de 50% de chance de evasão. Já com relação aos alunos que evadiram, 83, o modelo classificou corretamente 76 alunos como evasores (91,6%) e 7 (8,4%) como não evasores. Assim, a precisão, ou acurácia do modelo, foi estabelecida em 95,7%.

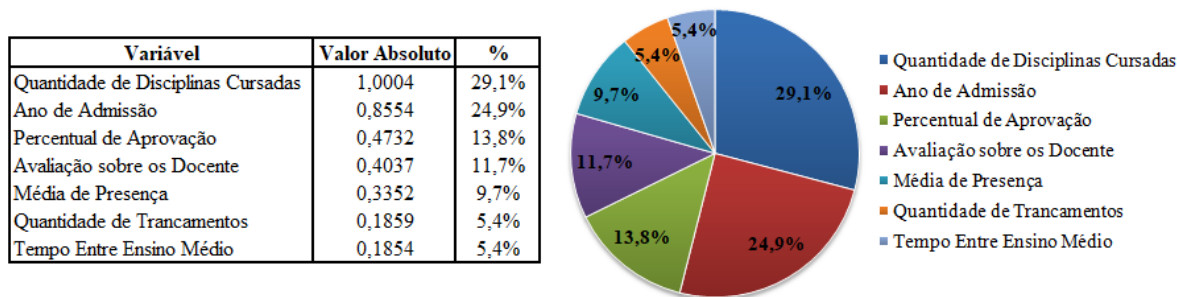
**Classification Results<sup>a</sup>**

		Predicted Group Membership		Total
		0	1	
Original	Count	0	1	
		188	5	193
		7	76	83
	%	97,4	2,6	100,0
		8,4	91,6	100,0

a. 95,7% of original grouped cases correctly classified.

**Figura 5 – Classificação dos Resultados e acurácia da função**  
 Fonte: Elaborado pelos autores (2020)

A função discriminante padronizada, que mostra o peso de cada variável está relacionada na figura 6. Por ela é possível visualizar os valores absolutos padronizados de cada uma das variáveis que apareceram na função discriminante final e perceber que a variável que mais discrimina a função com relação à evasão é a quantidade de disciplinas cursadas, seguida pelo ano de admissão, que sozinhos representam mais de 50% de discriminação. Também é possível notar que, para o curso de Engenharia de Produção da UFSJ, as variáveis relacionadas as notas do ENEM e a colocação do aluno no SISU não tiveram nenhuma representatividade na função encontrada.



**Figura 6 – Valor e peso dos coeficientes da função canônica padronizada**  
 Fonte: Elaborado pelos autores (2020)

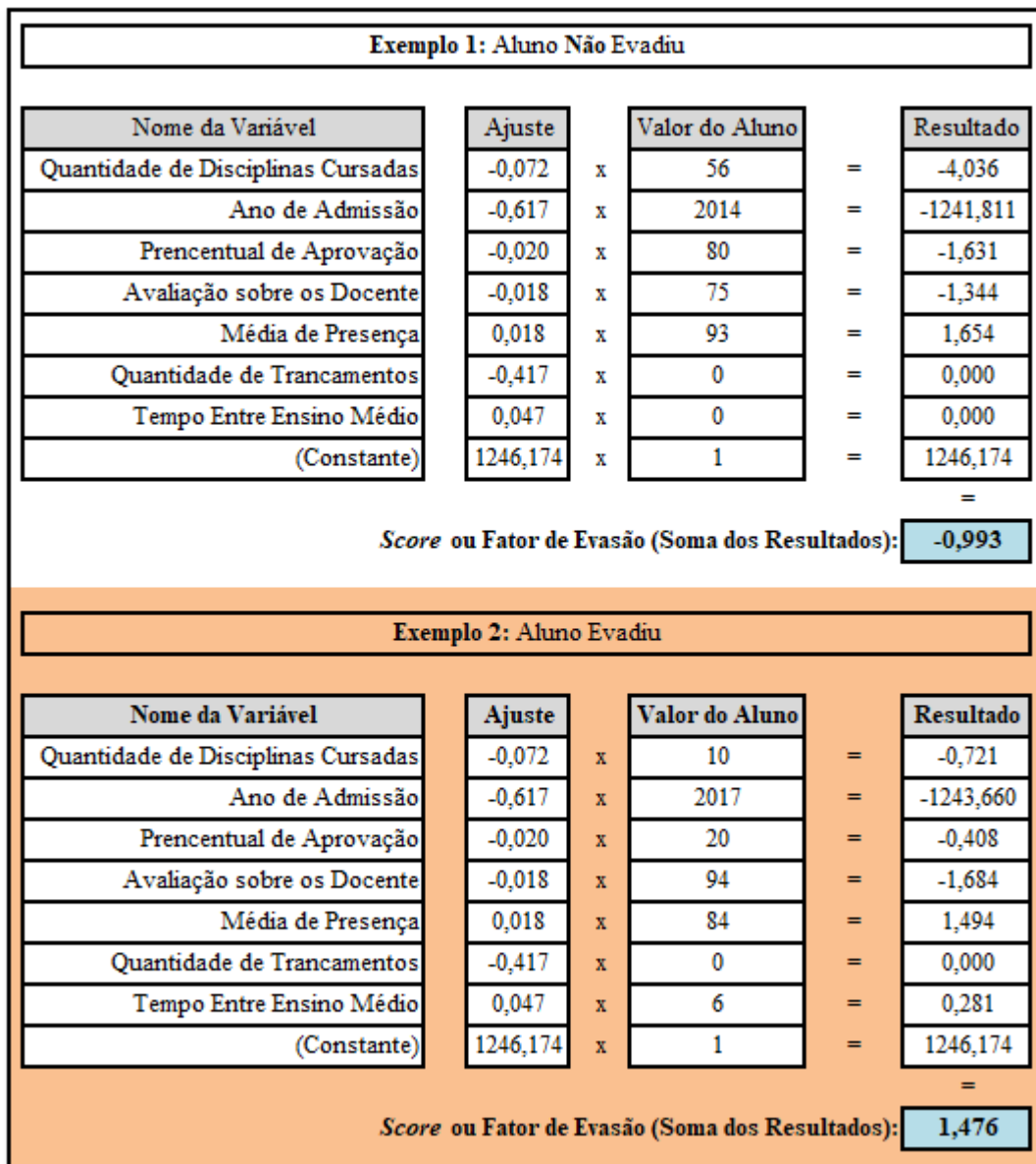
Apesar de a função padronizada facilitar a visualização dos pesos de cada variável, ela não pode ser utilizada dessa forma para calcular o *discriminant score* nem a probabilidade de evasão de cada aluno, isso acontece especialmente porque cada variável encontra-se em uma escala diferente. O ano de admissão, por exemplo, varia apenas entre 2014 até 2018, já o percentual de aprovação varia de 0 a 100. Para calcular o *score* e, conseqüentemente essa probabilidade serão utilizados, então, os coeficientes da função canônica discriminante não padronizada. A figura 7 exibe esses coeficientes, ou ajustes, de cada uma das sete variáveis da função, estes devem ser os fatores multiplicadores que posteriormente serão somados ou subtraídos a próxima variável até que se completem os quadrados da imagem para cada um dos alunos, a função também apresenta uma constante de ajuste, que não pode ser ignorada.

Nome da Variável	Ajuste		Valor do Aluno	=	Resultado
Quantidade de Disciplinas Cursadas	-0,072	x		=	
Ano de Admissão	-0,617	x		=	
Prencetual de Aprovação	-0,020	x		=	
Avaliação sobre os Docente	-0,018	x		=	
Média de Presença	0,018	x		=	
Quantidade de Trancamentos	-0,417	x		=	
Tempo Entre Ensino Médio	0,047	x		=	
(Constante)	1246,174	x		=	
<b>Score ou Fator de Evasão (Soma dos Resultados):</b>					

**Figura 7 – Cartão com os coeficientes da função canônica e seus cálculos**  
 Fonte: Elaborado pelos autores (2020)

A partir dos dados analisados, foram encontrados valores para o *score*, ou fator de evasão, que variam entre aproximadamente -3,5 até 5,5, onde os valores mais baixos representam alunos com menores características e possibilidade de evasão, e, valores mais altos indicam alunos com maior probabilidade de evadir, de acordo com o modelo. Na figura

8 são apresentados os cálculos de dois alunos, os dados pessoais desses alunos foram preservados por questões de ética, mas todos os valores utilizados são informações reais de dois alunos escolhidos aleatoriamente da população. O exemplo 1 apresenta um aluno que não evadiu da instituição e o exemplo 2, em destaque, informa os dados de outro, que já evadiu, observa-se grande diferença nos *scores* entre eles, -0,993 e 1,476 respectivamente.

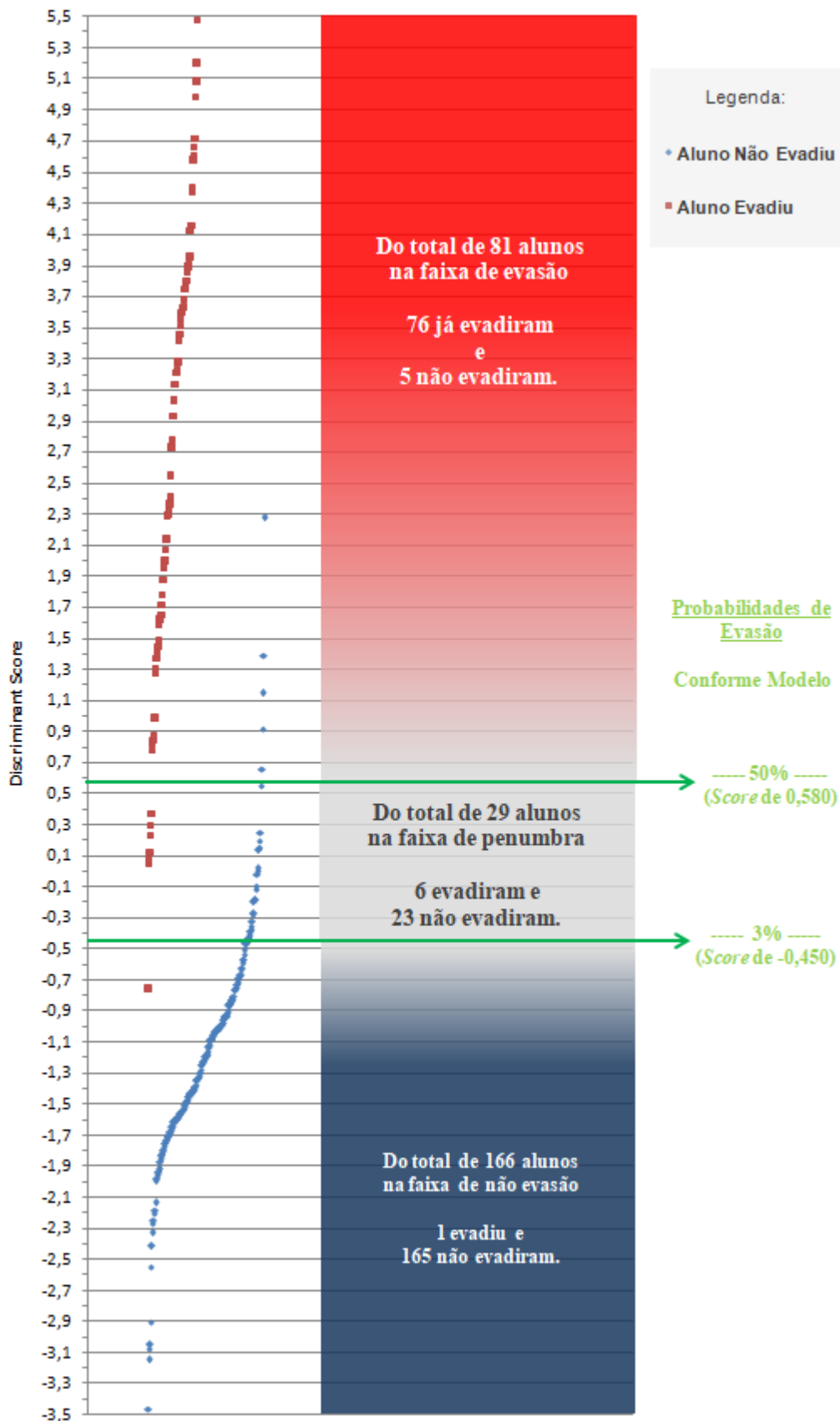


**Figura 8 – Exemplos de aplicação da função canônica**  
 Fonte: Elaborado pelos autores (2020)

De forma a exibir a posição de todos os alunos analisados elaborou-se um termômetro geral apresentado na figura 9. Este termômetro indica onde está localizado cada um dos alunos, sendo que, as marcações em vermelho são os alunos que já evadiram e em azul cada aluno que não evadiu. O eixo Y representa o *discriminant score* e a imagem também aborda

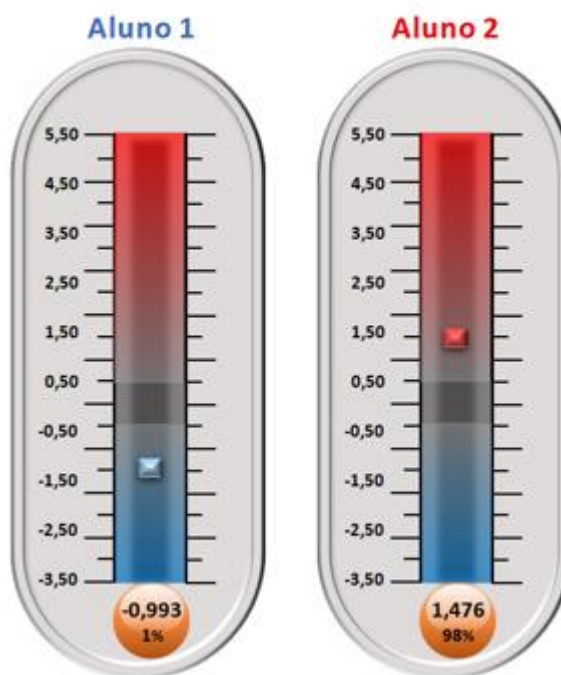
dois delimitadores importantes para a análise definido pelos autores, o primeiro refere-se ao local onde o *score* é aproximadamente 0,580, esse representa o marco onde os alunos tem 50% de chance de evasão, de acordo com o modelo. O segundo marco é o *score* aproximado de -0,45, que demarca onde alunos possuem 3% de chance de evadir. A faixa entre os dois marcos é considerada a faixa de penumbra pelos autores, os 28 alunos não evadidos que estão dentro dessa faixa ou acima dela devem ser prioridade no acompanhamento dos gestores ou coordenadores de curso a fim de que se possam evitar futuras evasões.

Por fim, continuando o exemplo dos dois alunos apresentados anteriormente, é possível localizá-los e apontá-los no termômetro a partir do seu fator de evasão. Isso é representado na figura 10, onde além do gráfico a figura realça o fator de evasão e a probabilidade de evasão de cada um dos dois alunos de acordo com o que foi calculado pela função discriminante.



**Figura 9 – Termômetro geral de evasão**  
 Fonte: Elaborado pelos autores (2020)





**Figura 10 – Termômetro individual dos alunos utilizados como exemplo**

Fonte: Elaborado pelos autores (2020)

Para demonstrar a eficácia do método foi conduzida ainda uma análise final com os dados mais recentes disponíveis da base de dados dos estudantes aqui verificados. Como até a publicação deste estudo nenhuma medida de prevenção de evasão direcionada aos alunos com maior probabilidade de desistirem do curso foi tomada, foi verificado a quantidade dos alunos que evadiram desde o início de 2019 até o início de 2022. Os seguintes resultados foram encontrados. Dos quinze alunos que evadiram a partir de 2019, onze já estavam com um score maior que  $-0.450$ , ou seja, dentro ou acima da faixa de penumbra mencionada na figura 10 e assim já eram considerados como possíveis desistentes. Além disso, de todos os estudantes em 2019 com score positivo (isto é, maior que  $0.000$ ), até o início de 2022 setenta por cento destes já havia abandonado o curso, vinte por cento ainda estavam matriculados com maiores chances de evasão, e apenas dez por cento conseguiu finalizar o curso e se formar, dando uma precisão real de 90% do modelo, mesmo propositalmente sem considerar nenhuma alteração nos cálculos desde 2019 e nem mesmo uma situação totalmente atípica de uma pandemia global que possivelmente alterou o cenário de desistências dos cursos de graduação em todo o mundo.

## Conclusão

Após todas as análises é possível concluir que existe uma considerável quantidade de evasões acontecendo em Engenharia de Produção da UFSJ, mesmo sendo um curso com ingresso bem concorrido e a instituição sendo uma prestigiada Universidade Federal do País. Apesar disso, com os resultados do artigo, foi possível encontrar um padrão de discrepância que permite a preemptiva identificação de alunos mais propensos a abandonar o curso e consequentemente facilitar o trabalho da gestão ou da coordenação para que possam agir de forma diferenciada com tais alunos a fim de evitar essas desistências. É válido mencionar que, apesar de não ser passível de generalizações entre diferentes instituições de ensino ou mesmo outros cursos, o termômetro pode ser reelaborado adaptando-o a partir dos dados disponíveis e de acordo com a realidade, por exemplo, uma instituição privada pode adicionar dados financeiros do aluno e de sua família, dado que possivelmente pode ser relevante para o abandono, mesmo para instituições públicas que detém tal informação. Como limitação, aponta-se que essa foi uma análise estritamente quantitativa, o que impossibilitou a identificação de causas mais subjetivas da evasão. Assim, sugere-se, para trabalhos futuros que este estudo seja analisado em conjunto com técnicas qualitativas para investigar, por exemplo, por que alunos com características de não evasão abandonaram seus cursos. Outra aplicação futura da pesquisa pode ser a elaboração de softwares que automatizem o processo aqui utilizado, gerando relatórios semestrais, alertas e gráficos individuais de cada estudante para as coordenações.

## Referências

- FRITSCH, R.; DA ROCHA, C. S.; VITELLI, R. F.. A evasão nos cursos de graduação em uma instituição de ensino superior privada. **Revista Educação em Questão**, v. 52, n. 38, p. 81-108, 2015.
- HARDOON, D. R.; SZEDMAK, S.; SHAW-TAYLOR, J. Canonical correlation analysis: An overview with application to learning methods. **Neural computation**, v. 16, n. 12, p. 2639-2664, 2004.
- KAUARK, F. S.; MANHÃES, F. C.; MEDEIROS, C. H. Metodologia da pesquisa: um guia prático. 2010.
- MARTINS, G. A.; THEÓPHILO, G. A. **Metodologia da investigação científica para ciências sociais aplicadas**. 3 ed. São Paulo: 2016

RIGO, Sandro José; BARBOSA, Jorge; CAMBRUZZI, Wagner. Educação em Engenharia e Mineração de Dados Educacionais: oportunidades para o tratamento da evasão. **EaD & Tecnologias Digitais na Educação**, v. 2, n. 3, p. 30-40, 2014.

SACCARO, A.; FRANÇA, M. T. A.; JACINTO, P. A.. Fatores Associados à Evasão no Ensino Superior Brasileiro: um estudo de análise de sobrevivência para os cursos das áreas de Ciência, Matemática e Computação e de Engenharia, Produção e Construção em instituições públicas e privadas. **Estudos Econômicos** (São Paulo), v. 49, n. 2, p. 337-373, 2019.

SHERRY, A.; HENSON, R. K. Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. **Journal of personality assessment**, v. 84, n. 1, p. 37-48, 2005.

SILVA FILHO, R. L. L., et al. A evasão no ensino superior brasileiro. **Cadernos de pesquisa**, v. 37, n. 132, p. 641-659, 2007.

UFSJ. Resolução Nº 012, de 4 de abril de 2018. Institui e regulamenta procedimentos acadêmicos no âmbito dos Cursos de Graduação da UFSJ e dá outras providências. 2018. Disponível em: <<https://ufsj.edu.br/colet/resolucoes.php/>>. Acesso em: 20 de jun. de 2019.

UFSJ. UFSJ extingue Vestibular tradicional. Página de Notícias, 2013. Disponível em: <[https://ufsj.edu.br/noticias\\_ler.php?codigo\\_noticia=3724/](https://ufsj.edu.br/noticias_ler.php?codigo_noticia=3724/)>. Acesso em: 20 de jun. de 2019.

Submetido em: 16.12.2022

Aceito em: 18.01.2023