

Análise multivariada de componentes principais dos fatores que interferem no (não) credenciamento de pesquisadores em programas de pós-graduação

João de Melo Maricato

jmmaricato@gmail.com

Faculdade de Comunicação e Biblioteconomia, Universidade Federal de Goiás, Goiânia, GO, Brasil

Ítala Moreira Alves

italaalves0@gmail.com

Instituto Federal Goiano, Goiânia, GO, Brasil

Dalton Lopes Martins

dmartins@gmail.com

Faculdade de Comunicação e Biblioteconomia, Universidade Federal de Goiás, Goiânia, GO, Brasil

Resumo: Esta pesquisa tem como objetivos contribuir para o objeto empírico “fatores que interferem no (não) credenciamento de pesquisadores em programas de pós-graduação”, e, sobretudo, explorar e compreender as potencialidades do uso da análise multivariada de componentes principais para a análise de dados no campo da ciência da Informação, em especial, nas áreas de bibliometria e da cientometria. A pesquisa foi feita com 578 docentes doutores das instituições públicas de ensino superior do estado de Goiás, com o retorno de 122 respostas válidas. Utilizou-se como metodologia de análise dos resultados alguns elementos de estatística descritiva e, também, análise multivariada de componentes principais por meio do Software SPSS. A análise mostra que das 22 variáveis utilizadas para realização desta pesquisa 4 fatores nos ajudam a explicar a variabilidade dos dados e apontam para condições fundamentais de entendimento das causas e questões que levam a esses resultados: fator infraestrutura, fator pessoal, fator motivacional e fator político. Estes aspectos estão próximos da realidade de gestores, o que facilita no processo de planejamento e incentivo aos sistemas de pós-graduação. Entende-se que análise multivariada pode ser utilizada em diversos tipos de problemas, tais como a caracterização de perfis de grupos, instituições, pesquisadores, artigos, citações, entre outros objetos de análise em relação a diferentes conjuntos de variáveis utilizados para descrever esses objetos. Tanto para a realização de estudos comparativos de indivíduos quando de variáveis utilizadas para descrevê-los.

Palavras-chave: Análise multivariada. Bibliometria. Cientometria. Pós-graduação.

Multivariate analysis of major components of factors affecting the (not) researchers accreditation in graduate programs

Abstract: This research aims to contribute to the empirical object "factors that interfere with (not) accreditation of researchers in graduate programs," and above all to explore and understand the potential use of multivariate principal component analysis for the analysis of data in the field of the science of information, particularly in the areas of bibliometrics and scientometrics. The survey was conducted with 578 doctor's teachers of public institutions of higher education in the state of Goiás, with the return of 122 valid responses. Was used as a methodology for analyzing results some descriptive statistics elements and multivariate principal component analysis using the SPSS software. The analysis shows that of the 22 variables used for this research four factors

help us to explain the variability of the data and point to fundamental ability to understand the causes and issues that lead to these results: infrastructure factor, personal factor, motivational factor and political factor. These aspects are close to the reality of managers, which facilitates the process of planning and encouraging postgraduate systems. It is understood that the multivariate analysis can be used in various types of problems such as the characterization of group profiles, institutions, research, articles, quotes, and other objects of analysis for different sets of variables used to describe those objects. For both the comparative studies of individuals as variables used to describe them.

Keywords: Multivariate analysis. Bibliometrics. Scientometrics. Graduate studies.

Análisis multivariado de componentes principales de los factores que interfieren en la (no) validación de investigadores en programas de posgrado

Resumen: Esta investigación tiene como objetivo contribuir al objeto empírico “factores que interfieren en la (no) validación de investigadores en programas de posgrado”, y, sobre todo, explorar y comprender las potencialidades del uso del análisis multivariado de los componentes principales para el análisis de datos en el campo de la Ciencia de la Información, en especial, en las áreas de bibliometría y cienciometría. La investigación fue aplicada a 578 docentes doctores de las instituciones públicas de enseñanza superior del estado de Goiás, teniendo un retorno de 122 respuestas válidas. Se utilizó como metodología de análisis de los resultados algunos elementos de estadística descriptiva y, también, de análisis multivariado de los componentes principales por medio del Software SPSS. El análisis muestra que de las 22 variables utilizadas para realización de esta investigación, cuatro factores nos ayudan a explicar la variabilidad de los datos y señalan para condiciones fundamentales para comprender las causas e interrogantes que llevan a esos resultados: el factor infraestructura, el factor personal, el factor motivacional y el factor político. Estos aspectos están próximos a la realidad de los gestores, lo que facilita el proceso de planificación e incentivo de los sistemas de posgrado. Se entiende que el análisis multivariado puede ser utilizado en diversos tipos de problemas, tales como la caracterización de perfiles de grupos, instituciones, investigadores, artículos, citas, entre otros objetivos de análisis en relación a los diferentes conjuntos de variables utilizados para describir esos objetos. Tanto para la realización de estudios comparativos de individuos como de variables utilizadas para describirlos.

Palabras-clave: Análisis multivariado. Bibliometría. Cienciometría. Posgrado.

1 Introdução

Diante das potencialidades dos cursos de pós-graduação para o desenvolvimento da ciência, tecnologia e inovação e, conseqüentemente, econômico e social, diversas questões relacionadas à sua estruturação e funcionamento merecem ser compreendidas. Um desses aspectos são os recursos humanos, especialmente, professores e pesquisadores, e os seus diversos interesses e motivações para atuar ou não em programas de pós-graduação. Não existe clareza, atualmente, sobre os fatores que interferem na decisão de professores e pesquisadores em atuar (ou não) em programas de pós-graduação. Diversos podem ser esses fatores: organizacionais, sociológicos, econômicos, comportamentais, ideológicos, relacionados à saúde, culturais, etc.

Considerando o exposto, a presente pesquisa tem como um dos objetivos explorar e compreender, por meio de análise multivariada, os fatores identificar fatores que influenciam na decisão, por parte de professores/pesquisadores, de credenciamento ou não em programas de pós-graduação *stricto sensu*, mesmo reunindo certas condições para tal.

Diversas podem ser as abordagens metodológicas e analíticas para a compreensão do fenômeno ora investigado. Além de buscar a compreensão do fenômeno em si, a presente pesquisa busca trazer contribuições por meio da abordagem metodológica utilizada.

A pesquisa, do ponto de vista de instrumental analítico, explora estatística descritiva, mas, enfatiza o uso da análise multivariada de componentes principais para compreender o fenômeno e investigar as possibilidades analíticas desse método.

Em estudo anterior a abordagem utilizada lançou mão da análise descritiva univariada dos dados. Com intuito de obter maior compreensão do fenômeno e principalmente explorar esse recurso analítico como potencial para a área, essa pesquisa tem como um ponto central a exploração da análise multivariada, ao realizar a análise simultânea de medidas múltiplas do objeto de análise. Parte-se da premissa de que essa opção analítica pode complementar as análises univariadas, possibilitando análises mais aprimoradas para a compreensão do fenômeno, sobretudo pelo fato de levar em consideração múltiplas variáveis e suas relações entre si, ampliando a complexidade do estudo e, imagina-se, chegando mais próximo de uma melhor modelagem do fenômeno social em estudo.

Diante do exposto, a presente pesquisa tem como objetivo trazer contribuições para o objeto empírico “fatores que interferem no (não) credenciamento de pesquisadores em programas de pós-graduação”, e, sobretudo, explorar e compreender as potencialidades do

uso da análise multivariada de componentes principais para a análise de dados desta natureza, discutindo, adicionalmente, as suas possibilidades de aplicações em estudos bibliométricos e cientométricos.

2 Análises multivariadas e suas aplicações em estudos bibliométricos e cientométricos

Análises multivariadas estão relacionadas as técnicas estatísticas que analisam simultaneamente múltiplas medidas sobre indivíduos ou objetos de investigação, podendo ser consideradas extensões ou complementares às análises univariadas ou bivariadas. As análises multivariadas pretendem executar, de uma única vez, uma análise de múltiplas variáveis, enquanto que em outras análises univariadas seria necessário executar diversas análises em separado para tentar compreender o fenômeno, analisando no máximo a relação entre duas variáveis. Existem, também, técnicas multivariadas que foram desenvolvidas para lidar com aspectos multivariados por excelência, como, por exemplo, a identificação de uma estrutura inerente a um conjunto de variáveis ou uma análise que visa distinguir grupos fundamentados em um conjunto de variáveis. Uma das maiores vantagens em se utilizar a análise multivariada é utilizar um método analítico mais sofisticado que “fornece previsões melhores de y do que um modelo com uma única variável explicativa. Tal modelo pode analisar também os relacionamentos entre variáveis enquanto controla outras variáveis”. (AGRESTI; FINLAY, 2012)

Uma das principais motivações para o uso dessa técnica estatística é o fato de ao se realizar pesquisas que envolvem a coleta de uma grande quantidade de dados, com o uso de questionários, por exemplo, tem-se dificuldade em explorar a relação entre as variáveis de resposta e encontrar possíveis elementos que ajude a explicar melhor o fenômeno em análise. De fato, como mencionam Fávero et al. (2009, p. 03):

“a quantidade de dados que podem ser extraídos de uma determinada pesquisa pode ser extremamente elevada, dificultando a determinação de como se darão as inter-relações entre as variáveis e, principalmente, como será a definição do modelo mais apropriado para se chegar às respostas desejadas.”

Em princípio, qualquer análise simultânea que de mais do que duas variáveis poder ser considerada multivariada. No entanto, na literatura existe confusão sobre o significado dessas análises. Para que uma análise possa ser realmente considerada como multivariada “... todas as variáveis devem ser aleatórias e inter-relacionadas de tal maneira que seus diferentes efeitos não podem ser significativamente interpretados em separado” (HAIR, 2009, p. 23).

O uso do termo análise multivariada *stricto sensu* “se refere a um conjunto de técnicas em particular, incluindo, entre outras, Análise de Agrupamento (Cluster Analysis), Escalonamento Multidimensional (Multidimensional Scaling), Análise Fatorial (Factor Analysis) e Análise de Correspondência (Correspondence Analysis). ” Essas técnicas de análises multivariadas podem auxiliar no descobrimento de informações e conhecimentos não imediatamente aparentes numa análise mais simples. (PEREIRA, 1999).

Fávero et al. (2009, p. 05) fazem importante distinção e facilitam o entendimento do que é análise univariada, bivariada e multivariada:

“A análise multivariada de dados é uma extensão natural das análises univariadas e bivariadas. Enquanto a análise univariada estuda os comportamentos e as distribuições de uma variável isolada, a análise bivariada estuda as associações e correlações e elabora as análises de variância relativas a duas variáveis simultaneamente. A análise multivariada, por sua vez, é utilizada para estudar modelos em que todas as variáveis sejam aleatórias e inter-relacionadas, de modo que seus diferentes efeitos não possam ser interpretados de forma separada.”

As análises multivariadas podem ser aplicadas a diversos dados e escalas de medida. Podem ser aplicadas, com limitações, a alguns dados não-métricos (qualitativos). No entanto, aplicações mais eficazes das análises multivariadas são para as escalas de medidas métricas (quantitativos), que são utilizadas:

[...] quando indivíduos diferem em quantia ou grau em relação a um atributo em particular. Variáveis metricamente medidas refletem quantidade ou grau relativo e são apropriados para atributos envolvendo quantia ou magnitude, como o nível de satisfação, ou compromisso com um emprego. As duas escalas de medida métricas são as escalas intervalares e de razão. [...] As escalas intervalares e escalas de razão (ambas métricas) fornecem o mais alto nível de precisão de medida, permitindo que quase todas as operações matemáticas sejam executadas. Essas duas escalas têm unidades constantes de medida e, portanto, diferenças entre quaisquer dois pontos adjacentes em qualquer parte da escala são iguais (HAIR, 2009, p. 25).

Os estudos bibliométricos e cientométricos historicamente focam suas análises na descrição de dados, havendo, normalmente, reduzida quantidade de parâmetros a serem analisados. Assim, as análises dos dados por meio de rankings e gráficos são predominantes. Normalmente os dados são analisados a partir de simples contagem de publicações ou de citações ou relações. Além da estatística univariada (por exemplo, o emprego de média, mediana, moda), observa-se, com menor frequência, emprego de estatística bivariada (como, por exemplo, as análises de correlação entre duas variáveis). No entanto, o uso de estatística multivariada ainda é pouco difundido na área.

O crescimento e aumento da complexidade da ciência, assim como dos aspectos sociais e cognitivos em que o conhecimento está estruturado tem sugerido o uso de outros métodos de análise. Os avanços em hardware e software ocorridos nas últimas décadas (especialmente os pacotes estatísticos) somado a ampliação das possibilidades coleta de dados quer sejam por entrevistas, questionários ou diretamente das bases de dados aumentaram a facilidade dos usos de técnicas de análise multivariadas e a transformação de dados brutos em conhecimento.

Análises multivariadas, em estudos bibliométricos e cientométricos, vem crescendo ao longo do tempo, destacando-se o uso de análises multivariadas nos estudos de citação e cocitação, coautoria, copalavras, cocalificação, etc. Algumas técnicas mais utilizadas na área são as de Análise de Clusters, Escalonamento Multidimensional, Análise de Correspondência e Análise Fatorial. Esses métodos pretendem simplificar a complexidade de padrões de associação de muitas variáveis. Essa simplificação ocorre por meio de uma representação ou projeção de um objeto com um alto dimensionamento espacial, em um espaço com uma dimensão numericamente menor, favorecendo melhor interpretação dos dados (OSAREH, 2012; VANZ; STUMPF, 2010).

Na presente pesquisa é utilizada a análise multivariada denominada componentes principais. É um tipo de análise fatorial cuja abordagem estatística pode ser utilizada para analisar inter-relações entre um grande número de variáveis de modo a explicar as variáveis em termos de suas dimensões inerentes comuns (fatores). O objetivo desse tipo de análise é: “encontrar um meio de condensar a informação contida em várias variáveis originais em um conjunto menor de variáveis estatísticas (fatores) com uma perda mínima de informação. Pelo fato de fornecer uma estimativa empírica”. Em suma, essa técnica estatística é preferida quando a meta principal é a redução de dados. (HAIR, 2009, p. 33, p. 143). Como consequência dessa redução de dados e um importante uso da análise fatorial, sobretudo considerando os interesses e aplicações da cientometria, é que ela “possibilita ao pesquisador a criação de indicadores inicialmente não observáveis compostos do agrupamento de variáveis” (FÁVERO et al., 2009, p. 235).

3 Metodologia

O universo da pesquisa foi construído a partir dos docentes doutores das instituições públicas de ensino superior do Estado de Goiás: Instituto Federal Goiano (IFGoiano), Instituto

Federal de Goiás (IFG), Universidade Estadual de Goiás (UEG) e Universidade Federal de Goiás (UFG).

A primeira etapa da pesquisa foi o levantamento de todos os professores doutores que faziam parte do quadro das quatro universidades públicas do estado de Goiás (IFGoiano, IFG, UEG e UFG) no ano de 2012.

Com a lista de professores dessas instituições, levantou-se a produção científica gerada por esses. A produção foi levantada por meio do Currículo Lattes de cada pesquisador no mês de março de 2013. Posteriormente foram selecionados os professores/pesquisadores que publicaram ao menos um artigo científico no período de 5 anos (2009 a 2013).

Dentre o conjunto de professores que publicaram ao menos um artigo científico no período, identificaram-se aqueles que não estavam credenciados em programas de pós-graduação. O levantamento foi realizado por meio dos dados disponíveis sobre os programas na CAPES e complementados com informações das informações disponíveis nas páginas da Web dos próprios programas de pós-graduação. Esse levantamento foi realizado nos meses de março e abril de 2014.

A partir dos procedimentos adotados, chegou-se ao número de 578 professores/pesquisadores que reuniam as características desejadas. Foram enviados questionários a esses professores para identificar aspectos ou fatores que influenciam na decisão de atuarem ou não atuarem em programas de pós-graduação, mesmo reunindo certas condições.

O questionário foi elaborado utilizando-se os princípios de Herzberg (1966) sendo aplicado no modelo de Satisfação de Nyquist, Hitchcock e Teherani (2000) utilizando, também, dos conceitos apresentados por Hagedon (2000). Além disso, buscou-se levantar, na literatura, características e possíveis fatores que poderiam influenciar na decisão de um professor/pesquisador participar ou não de programas de pós-graduação.

Esse instrumento de coleta de dados foi elaborado utilizando a escala de Likert, com 5 níveis, para compreender os níveis de influência, sendo: Influencia totalmente (5); Influencia fortemente (4); Influencia moderadamente (3); Influencia fracamente (2); e, Não influencia (1). Foram elaboradas 22 questões fechadas e uma aberta contemplando os fatores descritos no Quadro 1.

Quadro 1 – Fatores organizacionais, fatores relacionados ao trabalho e fatores pessoais contemplados nas questões

Organizacionais	Trabalho	Pessoais
Problemas relacionados a relações pessoais	Precária política de pós-graduação do estado de Goiás	Interesse pelas atividades de pesquisa, orientação, publicação, etc.
Discordância com o sistema de avaliação da CAPES	Distância geográfica do programa de interesse	Responsabilidade social e profissional
Ausência de políticas da instituição de ensino	Falta de remuneração	Reconhecimento social
Dificuldade de adaptação com a cultura institucional	Falta de qualidade da infraestrutura (biblioteca, laboratórios, etc.)	Possibilidade de captação de recursos
Inexistência de linhas de pesquisa de meu interesse	Sobrecarga de trabalho atual	Realização pessoal
Dificuldade de atingir os critérios dos programas de credenciamento dos programas de pós-graduação	Inexistência de programas de interesse	Maior liberdade de pesquisa e desenvolvimento acadêmico
Escassez de recursos para pesquisa		Dificuldades relacionadas à família
		Dificuldades relacionadas à saúde
		Desinteresse nas atividades de pesquisa e pós-graduação

Fonte: dados da pesquisa

O questionário foi enviado eletronicamente (por *e-mail*) aos 578 aos professores/pesquisadores no dia primeiro de julho de 2014 e o recebimento das respostas deu-se até o dia 8 de agosto de 2014. Ao final do prazo de recebimento das respostas, chegou-se a 122 respostas válidas.

Constatou-se que 38 *e-mails* retornaram (não foram recebidos pelos destinatários), 38 docentes informaram que se credenciaram recentemente em programas de pós-graduação e dois aposentaram-se e desligaram-se da instituição. Os demais não deram retorno e não responderam o questionário.

Utilizou-se como metodologia de análise dos resultados alguns elementos de estatística descritiva e, também, análise multivariada de componentes principais. As análises multivariadas foram realizadas utilizando-se o Software SPSS.

4 Resultados e discussões

A partir da análise descritiva e univariada, pôde-se constatar que 94% dos respondentes afirmaram ter interesse em participar de programas de pós-graduação. Percebe-

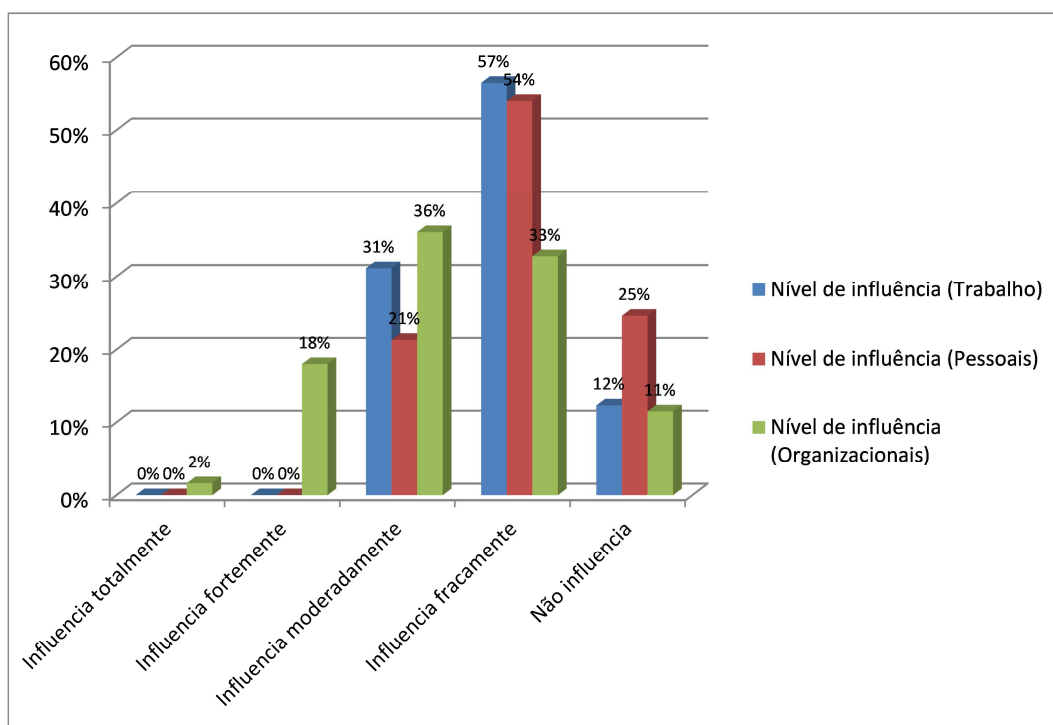
se que uma parcela importante dos respondentes ingressou recentemente na carreira docente em universidades públicas e que, apesar da quantidade significativa de pessoas que ainda não atuaram nesses programas, quase a totalidade tem o interesse em participar e, apenas 6% informaram não ter interesse em credenciar.

Analisando-se o conjunto de dados, observa-se que os fatores relacionados ao trabalho e os pessoais possuem muito baixa influência na decisão do pesquisador em não participar de programas de pós-graduação (Gráfico 1). Nenhum dos respondentes, a partir das médias das respostas, foi categorizado em “Influencia totalmente” ou “Influencia fortemente”.

Constatou-se com as análises que os fatores organizacionais em conjunto, demonstram serem aqueles que influenciam mais fortemente na decisão dos professores/pesquisadores em atuar ou não em programas de pós-graduação. Observou-se que fatores organizacionais influenciam fortemente ou totalmente 20% respondentes.

Além disso, esse fator é apontado como influência moderada para 36% dos respondentes, enquanto a “Influência moderada” fica abaixo no contexto dos fatores relacionados ao trabalho (31%) e pessoais (21%). Certamente, os fatores organizacionais estão sendo decisivos na influência dos professores decidirem em atuar ou não em programas de pós-graduação.

Gráfico 1 – Fatores que influenciam na decisão de professores/pesquisadores participarem ou não de programas de pós-graduação



Quando se analisa os fatores pessoais isoladamente, nenhum deles merece destaque. No entanto, quando se analisa os fatores relacionados ao trabalho isoladamente, constata-se que um fator se destaca entre os demais: “Sobrecarga de trabalho atual”. Este possui forte influência para a maioria dos respondentes, alcançou a mediana de 3,4 na escala de 1 a 5, indicando ser, isoladamente, um fator que influencia moderadamente a fortemente a decisão dos professores/pesquisadores em atuar em programas de pós-graduação. Diversos foram os relatos relacionados a acúmulo de responsabilidades e sobrecarga de trabalho, especialmente aquelas relacionadas a demandas administrativas e relacionadas aos cursos de graduação.

A participação em programas de pós-graduação exige do docente, qualificação, carga horária, volume/qualidade das publicações, orientações, às quais devem ser somadas a outras várias funções dentro da instituição. Apesar da existência do discurso de forte discordância com o sistema CAPES de avaliação, largamente abordado por alguns autores, como Gatti et al. (2003); Hortale (2003); Silva (2007); Sguissardi (2006), além da falta de liberdade para a pesquisa abordada por Luz (2005), constatou-se que esse fator, isoladamente, não foi apontado como importante na decisão de um professor/pesquisador ingressar ou não em programas de pós-graduação.

Analisando-se os fatores organizacionais, constata-se que a dificuldade de atingir os critérios exigidos pelos programas de pós-graduação para credenciamento é o principal fator apontado pelos professores/pesquisadores. Esse fator, na escala de 1 a 5 obteve a mediana 4, ou seja, esse fator influencia fortemente o grupo. Vários professores/pesquisadores relatam, de maneira complementar, que os cursos existentes são fechados e que os critérios para credenciamento são fortemente políticos e pouco objetivos. Confirmando essa tendência, observou-se que vários foram os relatos de dificuldades de relacionamento interpessoal.

Outro fator organizacional que merece ser destacado refere-se ao fator “Ausência de políticas da instituição de ensino”. Esse fator atingiu a mediana 3 para o grupo analisado (Influencia moderadamente). Analisando-se os dados qualitativos, constata-se que os relatos se referem à ausência de políticas inclui aspectos relacionados a exigência de produtividade sem infraestrutura adequada, falta de políticas institucionais e governamentais que deem apoio ao desenvolvimento das pesquisas, falta de infraestrutura física e de recursos humanos, burocracia estatal, dentre outros.

Para complementar a análise descritiva univariada, observam-se, agora, os resultados da análise multivariada de componentes principais. O objetivo principal de realizar a análise

multivariada com base em nosso conjunto é indicar possíveis relacionamentos entre as variáveis que nos permitam apontar fatores que auxiliem a explicar a causalidade dos resultados. Fala-se aqui de fatores sociais, sobretudo, considerando as dimensões nas quais as próprias perguntas e variáveis da pesquisa qualitativa foram organizadas. Na medida em que essas perguntas podem ser analisadas em grupo, pois representam tendências de resposta em conjunto, podendo-se identificar quais os principais fatores daqueles utilizados para questionar os participantes da pesquisa que explicam a variabilidade dos dados e nos ajudam a entender possíveis padrões de comportamento e atitude a partir das respostas fornecidas.

Na Tabela 1, apresentam-se alguns resultados estatísticos que permitem avaliar a pertinência da análise multivariada para os dados com os quais estão sendo trabalhados.

Tabela 1 – Teste de KMO e Bartlett

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		,663
Teste de esfericidade de Bartlett	Qui-quadrado aprox.	764,071
	df	231
	Sig.	,000

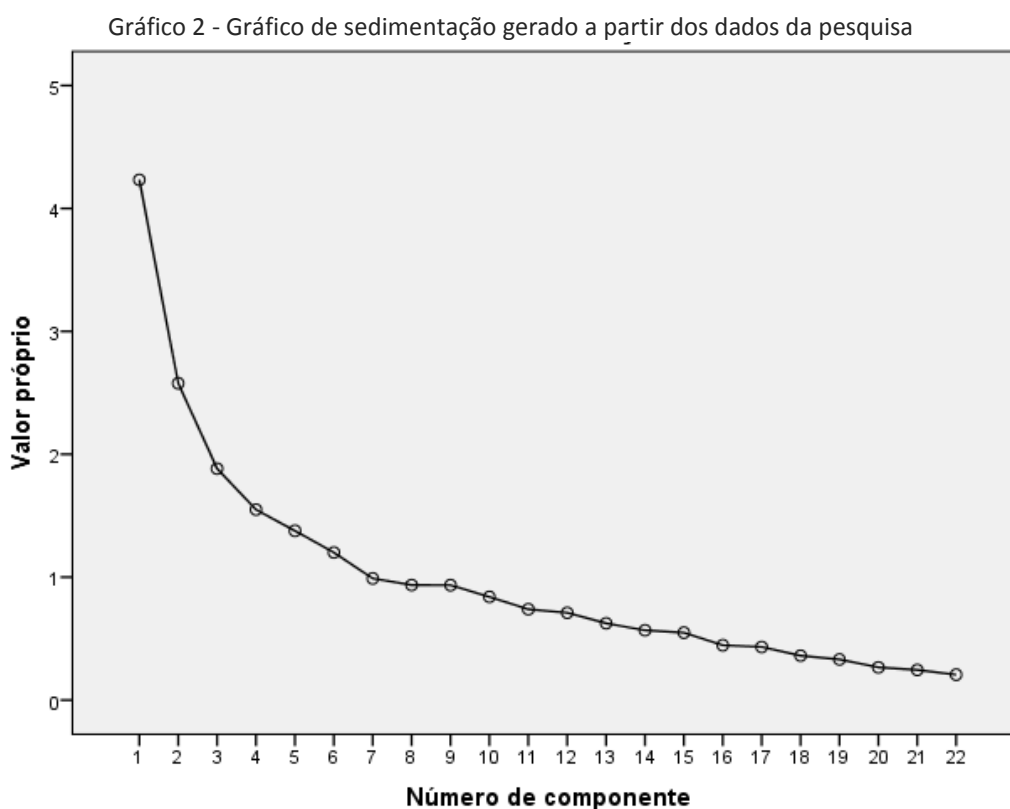
Fonte: dados da pesquisa.

A medida de Kaiser-Meyer-Olkin de adequação da amostragem é calculada para variáveis individuais e múltiplas e representa a razão da correlação ao quadrado para a correlação parcial ao quadrado entre as variáveis. Valores próximos de 0, indicam que as correlações parciais são grandes e há uma grande difusão na relação entre as variáveis, logo, a análise de fatores não é recomendada. Valores próximos a 1 indicam que as relações entre as variáveis são restritas, assim a análise de fatores se torna mais pertinente ao conjunto de dados. Recomenda-se que valores entre 0,7 e 0,8 são muito bons para esse tipo de análise.

Observa-se que o resultado da presente pesquisa retorna 0,663, o que pode ser considerado muito próximo do limite de um bom valor para prosseguir as análises. Já a significância apresentada na Tabela 1 pela sigla “sig” está relacionada ao teste de esfericidade de Bartlett e é desejável que esse valor seja menos que 0,05, o que é nosso caso. Esse teste indica que há relacionamento entre as variáveis, pois testa se a matriz de relacionamento entre elas é uma matriz identidade (indicando que não há relacionamentos entre as variáveis a não ser entre elas mesmas). Logo, conclui-se que os resultados analíticos da pertinência da

análise de fatores são favoráveis para considerarmos essa análise estatisticamente válida ao nosso conjunto de dados.

No Gráfico 2, vê-se a sedimentação que permite identificar quantos fatores são pertinentes e que se deve considerar como estratégia de agrupamento das variáveis. A teoria estatística de análise multivariada sugere que seja realizada a análise quando essa curva faz uma mudança significativa de declividade e passa decair de forma mais suave. Nos dados da pesquisa, percebe-se que o número 4 indica uma importante mudança de declividade se propagando até o número 7, onde muda novamente, porém de forma menos acentuada, e passa a decair com a mesma declividade até o final do conjunto de componentes. Dessa maneira, considera-se que se deve agrupar as variáveis num conjunto de 4 fatores e, dessa forma, procurar explicar a variabilidade dos dados.



Fonte: dados da pesquisa.

Por fim, a Tabela 2, mostra a estrutura dos 4 fatores que foram utilizados para a análise dos dados. Observa-se na Tabela 2 o número de 4 componentes que são formados pelo agrupamento das 22 perguntas, logo variáveis, utilizadas no questionário. Esses 4 componentes agrupam, em geral, diferentes variáveis e formam fatores. Quando se analisa as

variáveis que os compõem tem-se a possibilidade de entender melhor como as dimensões “trabalho”, “organizacional” e “pessoal” propostas para esta pesquisa se relacionam entre si.

Os resultados indicam que o primeiro componente é formado pela junção de variáveis da dimensão trabalho e organizacional. Quando analisadas em conjunto, essas variáveis falam sobre as condições estruturais do ambiente de pesquisa e fomento a cultura de pós-graduação, ressaltando a falta de qualidade de infraestrutura, inexistência de programas e linhas de pesquisa de interesse, precariedade e ausência de políticas, distância geográfica, cultura organizacional e recursos financeiros de forma geral. Esse fator, portanto, mostra que as pessoas respondem a essas perguntas de forma similar, ressaltando as condições estruturais como o principal fator que explica o nosso conjunto de dados.

O segundo componente é formado pela junção de variáveis da dimensão pessoal, organizacional e uma do trabalho relacionado à remuneração. Esse fator mostra a influência das condições pessoais incluindo o relacionamento entre as pessoas em seu ambiente de trabalho, o que pode ser visualizado pela questão cultural e da liberdade de pesquisa. As condições pessoais parecem aqui apontar importante fator relacionado a predisposição a pesquisar.

O terceiro componente é formado apenas por variáveis da dimensão pessoal e ressalta pela composição temática de fatores que em conjunto explicam o interesse ou não pela pesquisa por variáveis motivacionais e de reputação.

Já o último componente é formado por algumas variáveis que indicam a influência da política de avaliação da pesquisa da CAPES, dos programas de pós-graduação e da política no estado de Goiás. Têm-se aqui as condições políticas em questão.

Logo, a análise mostra que das 22 variáveis utilizadas para realização desta pesquisa, 4 fatores nos ajudam a explicar a variabilidade dos dados e apontam para condições fundamentais de entendimento das causas e questões que levam a esses resultados. Os fatores são descritos aqui como fator infraestrutura, fator pessoal, fator motivacional e fator político. É possível imaginar, a partir desta pesquisa, um indicador que pudesse avaliar a evolução das condições do sistema de pós-graduação no estado de Goiás a partir desses 4 fatores e estudar como políticas de fomento e incentivo possam impactar no seu desenvolvimento. Além disso, o conhecimento desses 4 fatores torna-se um elemento fundamental para gestores de políticas públicas e privadas, voltadas a ciência e tecnologia, realizarem planejamentos e considerarem formas de alterar o quadro atual apresentado por esta pesquisa.

Tabela 2 – Matriz de estruturas dos 22 fatores utilizados para a análise dos dados

	Componente			
	1	2	3	4
TRAB-Falta de qualidade da infraestrutura (biblioteca, laboratórios, etc)	,743			
TRAB-Inexistência de programas de interesse	,685			
TRAB-Precária política de Pós-graduação do estado de Goiás	,663			,476
ORG-Ausência de políticas da instituição de ensino	,631			
ORG-Inexistência de linhas de pesquisa de meu interesse	,556			
TRAB-Distância geográfica do programa de interesse	,548			
TRAB-Sobrecarga de trabalho atual				
PES-Dificuldades relacionadas a família		,716		
ORG-Problemas relacionados a relações pessoais		,665		
ORG-Dificuldade de adaptação com a cultura institucional	,438	,652		
PES-Dificuldades relacionadas à saúde		,610		
TRAB-Falta de remuneração	,403	,545		
PES-Maior liberdade de pesquisa e desenvolvimento acadêmico		,434		
PES-Desinteresse nas atividades de pesquisa e pós-graduação				
PES-atividades de pesquisa, orientação, publicação			,754	
PES-responsabilidade social e profissional			,691	
PES-realização pessoal			,685	
PES-captação de recursos			,664	
PES-reconhecimento social			,636	
ORG-Discordância com o sistema de avaliação da CAPES				,765
ORG-Dificuldade de atingir os critérios dos programas de credenciamento dos programas de Pós-Graduação				,734
ORG-Escassez de recursos para pesquisa	,583			,668

Método de extração: Análise do Componente principal.

Método de rotação: Promax com normalização de Kaiser.

Fonte: dados da pesquisa

5 Conclusão

A análise de componentes principais se aplica a tabelas de dados onde as linhas correspondem aos indivíduos a serem pesquisados e as colunas as variáveis quantitativas a serem levadas em consideração. Essa análise permite o estudo tanto da relação entre as variáveis quanto da relação entre os indivíduos.

Estudar os indivíduos significa identificar similaridades entre eles, ou seja, dado o conjunto de variáveis em questão de que maneira os indivíduos classificados por essas variáveis são próximos ou distantes. Já estudar as variáveis permite identificar os múltiplos relacionamentos lineares que existem entre elas, já que o método de componentes principais apenas lida com esse tipo de relacionamento entre variáveis. (HUSSON; LÊ; PAGÉS, 2011)

Um dos aspectos úteis a serem explorados pela análise de componentes principais é a possibilidade de agrupar e reduzir o número de variáveis significativas que deveriam ser utilizadas para análise de um problema. Ou seja, o método apresenta como as diferentes variáveis contribuem para as diferenças ou similaridades entre os indivíduos, de onde se pode chegar a conclusão que duas ou mais variáveis possuem o mesmo comportamento, logo utilizar apenas uma permitiria avaliar a mesma tendência no conjunto de dados.

Dessa forma, o objetivo central da análise de componentes principais é permitir tirarmos conclusões das relações lineares entre as variáveis detectando quais são as dimensões principais responsáveis pela variabilidade nos dados (HUSSON; LÊ; PAGÉS, 2011) o que explicaria, portanto, quais são as principais dimensões variáveis que devem ser levadas em consideração na caracterização do conjunto de dados em questão.

A análise de componentes principais mostrou quatro fatores fundamentais para o entendimento da estrutura e organização dos dados, dando destaque para a influência do fator infraestrutura, fator pessoal, fator motivacional e fator político. Entende-se que a análise dos dados dos fatores que influenciam na decisão de pesquisadores em atuar em programas de pós-graduação pode ser melhor compreendida com o uso da estatística multivariada, sendo complementar as análises univariadas e bivariadas. O apontamento desses fatores tornou 22 variáveis compreendidas a partir de quatro aspectos, o que de fato simplifica o entendimento e facilita a possibilidade de entender o fenômeno a partir de aspectos mais próximos da realidade de gestores e, portanto, facilita o próprio processo de planejamento e incentivo ao sistema de pós-graduação.

Entende-se que esse tipo de análise pode ser utilizado, também, em diversos tipos de problemas de interesse da cientometria e bibliometria, tais como a caracterização de perfis de grupos, instituições, pesquisadores, artigos, citações, entre outros objetos de análise em relação a diferentes conjuntos de variáveis utilizados para descrever esses objetos. Tanto para a realização de estudos comparativos de indivíduos quando de variáveis utilizadas para descrevê-los, a análise de componentes principais parece contribuir com as pesquisas na área.

Referências

- AGRESTI, Alan; FINLAY, Barbara. **Métodos estatísticos para as ciências sociais**. 4. ed. Porto Alegre: Penso, 2012. 664 p.
- CAPES. Coordenadoria de Pessoal de Nível Superior. Disponível em: <<http://www.capes.gov.br/>>. Acesso em: maio 2014.
- FÁVERO, Luiz; BELFIORE, Patrícia; SILVA, Fabiana Lopes da; CHAN, Betty Lilian. **Análise de dados: modelagem multivariada para tomada de decisões**. 5. ed. Rio de Janeiro: Elsevier, 2009. 646 p.
- PEREIRA, Júlio César Rodrigues. **Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde humanas e sociais**. São Paulo: Edusp, 1999.
- GATTI, Bernardete *et al.* Documento: o modelo de avaliação da CAPES. **Revista Brasileira de Educação**, n. 22, jan./abr. 2003.
- HAIR, Joseph F. *et al.* **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009.
- HORTALE, Virginia Alonso. Modelo de avaliação CAPES: desejável e necessário, porém, incompleto. **Cad. Saúde Pública**, Rio de Janeiro, v. 6, n.19, p.1837-1840, nov./dez. 2003.
- HUSSON, François, LÊ, Sébastien, PAGÉS, Jérôme. **Exploratory Multivariate Analysis by example using R**. CRC Press. 1ed. 2011. 228p
- LUZ, Model T. Prometeu acorrentado: análise sociológica da categoria produtividade e as condições atuais da vida acadêmica. **PHYSIS: Rev. Saúde Coletiva**, Rio de Janeiro, v. 1, n.15, p.39- 57, 2005.
- OSAREH, Farideh. The use and application of multivariate analysis techniques in bibliometric and scientometric studies. **International Journal of Information Science and Management**, v. 1, n. 2, p. 59-71, 2012.
- SGUISSARDI, Valdemar. A avaliação defensiva no “modelo CAPES de avaliação” - É possível conciliar avaliação educativa com processos de regulação e controle do Estado? **Perspectiva**, Florianópolis, v. 14, n. 1, jan./jun. 2006, p. 49-88.
- SILVA, Alcino Lázaro da. Pós-graduação e CAPES. **Rev. Col. Bras. Cir.**, v. 34, n.6, p. 360, nov./dez. 2007.
- VANZ, Samile Andréa de Souza; STUMPF, Ida Regina Chittó. Procedimentos e ferramentas aplicados aos estudos bibliométricos. **Informação & Sociedade: Estudos**, v. 20, n. 2, 2010.

Recebido/Recibido/Received: 2015-05-26

Aceitado/Aceptado/Accepted: 2015-08-27