



ConBRepro

XI CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO



01 a 03
de dezembro 2021

Classificação dos estados brasileiros em relação ao acesso à informação por meio de um modelo de análise fatorial multivariada

Mayara Scheuer

Programa de Pós-Graduação em Engenharia de Produção - UFPR

Marcos Augusto Mendes Marques

Programa de Pós-Graduação em Engenharia de Produção - UFPR

Marcelo Vitor Fiatkoski

Programa de Pós-Graduação em Engenharia de Produção - UFPR

Resumo: Sabendo que o acesso à informação é o maior gerador de desenvolvimento tecnológico de um país e que a infraestrutura dos estados influencia diretamente neste desenvolvimento, este artigo apresenta a classificação dos estados brasileiros com relação ao nível de acessibilidade de informação por meio de uma metodologia de classificação. A metodologia escolhida para a aplicação foi a análise fatorial multivariada. O desenvolvimento do método foi dividido nas etapas de: definição e redução de variáveis da matriz de dados, validação da possibilidade de fatoração do conjunto de dados por meio do teste de esfericidade de Bartlett e da medida de adequacidade da amostra de Kaiser-Meyer-Olkin (KMO), estimação dos fatores, classificação dos estados a partir dos escores fatoriais finais ponderados e por fim, a discussão dos resultados. Concluiu-se que os três estados que devem ser priorizados nos planos de ação de políticas de melhoria são os estados do Maranhão, Piauí e Alagoas, nesta ordem de prioridade. As variáveis que mais impactam na análise estão relacionadas a equipamentos e aos meios de acesso à informação, desta forma o plano de ação de melhoria deve contemplar ações de desenvolvimento focado nestas variáveis. De acordo com a metodologia aplicada, os três estados mais bem colocados na classificação foram os estados do Rio de Janeiro, São Paulo e Santa Catarina.

Palavras-chave: Análise multivariada, Análise Fatorial, Acesso à informação.

Classification of Brazilian states in relation to access to information using a multivariate factor analysis model

Abstract: Knowing that access to information is the largest generator of technological development in a country and that state infrastructure directly influences this development, this paper presents the classification of Brazilian states in relation to the level of accessibility of information through a classification methodology. The methodology chosen for the application was multivariate factor analysis. The development of the method was divided into the steps of: definition and reduction of data matrix variables, validation of the possibility of factorization of the data set by Bartlett's sphericity test and the Kaiser-Meyer-Olkin sample adequacy measure (KMO), factor estimation, state classification based on the final weighted factor scores and, finally, the discussion of the results. It was concluded that the three states that should be prioritized in improvement policy action plans are the states of Maranhão, Piauí and Alagoas, in this order of priority. The variables that most impact

the analysis are related to equipment and the means of access to information, so the improvement action plan should include actions in the development of these variables. According to the applied methodology the three best-ranked states were the states of Rio de Janeiro, São Paulo, and Santa Catarina.

Keywords: Multivariate Analysis, Factor Analysis, Access to Information

1. Introdução

O conhecimento vem sendo considerado o principal fator gerador de riqueza, e isto se deve principalmente ao advento da revolução tecnológica. Estamos na chamada era da informação e do conhecimento (NETO et. al 2018). Segundo Silva et. al (2018) vivemos a popularização do acesso à tecnologia por meio do uso constante de dispositivos eletrônicos, através dos computadores, smartphones e tablets. Portanto, percebe-se a relação direta entre conhecimento e o acesso à tecnologia. Em levantamento feito no ano de 2017 e divulgado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), apontou-se que o Brasil vem acompanhando significativamente esse crescimento do acesso à internet, sendo que em tal ano o número de internautas cresceu cerca de 10 milhões.

Apesar do aumento do acesso à internet e dos meios disponíveis para tal, isto ainda não é a realidade de diversas pessoas e regiões brasileiras (HENRIQUES, 2016). Portanto, é necessário que as empresas responsáveis pelo fornecimento dos meios de acesso à Internet possam identificar onde há a falta deste recurso, para que então possam estudar a possibilidade de fornecimento nestas áreas. No entanto, para que isso seja possível, é preciso que as fornecedoras tenham um direcionamento e uma classificação de prioridades de regiões e aspectos relacionados à disponibilidade de recursos. Para que desta forma, possa ser feito um planejamento estratégico eficiente e assertivo.

Uma forma de suporte para a tomada de decisão das ações e do planejamento estratégico é por meio da identificação do comportamento das regiões em relação à utilização da internet, ao acesso a equipamentos e os meios de acesso, como também as variáveis de maior influência para a utilização da rede. Isto é possível por meio de técnicas de análise estatística multivariada de dados.

O agrupamento de regiões também pode ser considerado uma forma de analisar o cenário brasileiro quanto ao acesso ao conhecimento. Este trabalho, particularmente, consiste no estudo da classificação dos vinte e sete estados brasileiros com relação à acessibilidade da informação, com enfoque nos grupos menos favorecidos. Analisando-se a existência, ou não, de algum fator geográfico ou social significativo nas regiões menos favorecidas e quais as variáveis que mais impactam para a classificação.

É neste contexto, onde se observa a necessidade de acesso a Internet - principalmente por, atualmente, ser umas das maiores fontes de acesso ao conhecimento e informação - sendo um fator essencial para o crescimento do país. O presente artigo tem como objetivo a aplicação de uma técnica estatística multivariada, mais especificamente análise fatorial multivariada, para diagnosticar o comportamento das variáveis relacionadas ao acesso ao conhecimento e a classificação de cada estado brasileiro em relação a sua acessibilidade. Desta forma, as empresas fornecedoras podem atuar de forma eficiente e assertiva no fornecimento de equipamentos e meios de acesso à informação das regiões mais necessitadas.

pode ser aplicado também em diferentes áreas e para outros fins, e não somente na classificação dos estados, mas como também na análise e classificação de fornecedores, priorização de políticas públicas em outras áreas como saúde, educação, etc. Desta maneira, o artigo tem como objetivo propor um método de classificação o qual pode ser aplicado à várias áreas da engenharia de produção e a partir deste modelo classificar os estados brasileiros em relação ao acesso à informação.

2. Referencial teórico

2.1 Análise multivariada

Quando o conjunto de dados de um ou mais objetos de interesse podem ser determinados por medidas simultâneas ou por diferentes variáveis de interesse, pode-se utilizar a análise multivariada de dados (MARQUES, 2005).

As técnicas de análise multivariada permitem a criação de conhecimento pelas organizações, isto porque encontra-se a partir destas técnicas as relações entre as variáveis, as quais muitas vezes não são percebidas intuitivamente sem a aplicação de alguma técnica. Segundo Hair et al. 2009, a análise multivariada refere-se às técnicas estatísticas que analisam juntamente múltiplos dados sobre objetos sob estudo. Seu objetivo consiste em medir, explicar e prever o grau de relação entre variáveis estatísticas.

O constructo da análise multivariada é a variável estatística, uma combinação linear de variáveis com pesos empiricamente determinados. As variáveis são especificadas pelo responsável do estudo e os pesos são estabelecidos a partir da técnica multivariada para atingir um objetivo específico (HAIR et. al, 2009). Uma variável estatística de n variáveis ponderadas e pode ser enunciada matematicamente como:

$$\text{Valor da variável estatística} = w_1X_1 + w_2X_2 + \dots + w_nX_n \quad (1)$$

Onde:

X_n = variável observada

w_n = peso estabelecido pela técnica multivariada.

Tem-se como resultado, a combinação do conjunto inteiro a partir de um único valor o qual melhor atinge o objetivo da análise multivariada.

2.2 Análise fatorial

Segundo Fachel (1976) a análise fatorial é a técnica que trata as relações internas de um conjunto de dados, resultando na substituição dos valores iniciais das variáveis correlacionadas por um conjunto reduzido de “fatores”, podendo ser não correlacionadas ou correlacionadas. Sendo que esse conjunto de fatores é responsável por explicar a maior parte da variância do conjunto original. HAIR et. al (2009), propõe que o objetivo desta técnica é buscar e definir os constructos fundamentais ou dimensões assumidas como inerentes às variáveis originais, com a menor perda de informação possível.

Para que a análise fatorial possa ser feita, é preciso verificar se os dados são válidos para esta técnica, ou seja, se são passíveis de fatoração. Verifica-se a possibilidade de fatoração por meio do teste de esfericidade de Bartlett e da medida de adequacidade da amostra de Kaiser-Meyer-Olkin (KMO).

2.2.1 Teste de esfericidade de Barlett

O teste de esfericidade de Bartlett mostra qual é o nível que a matriz de covariância é semelhante com a matriz-identidade (Damásio, 2012). Também testa a hipótese de que as variáveis do conjunto de dados não sejam correlacionadas na população, ou seja, admitir ausência de associação linear entre as variáveis. A hipótese básica diz que a matriz de correlação da população é uma matriz identidade a qual indica que o modelo fatorial é inapropriado (MARQUES, 2005).

$$\text{O teste é obtido através da equação: } x^2 = - \left[(n - 1) - \frac{2p+5}{6} \right] \ln|R| \quad (2)$$

$$\text{Com distribuição qui-quadrado com graus de liberdade } v = \frac{p(p-1)}{2} \quad (3)$$

Onde:

n = tamanho da amostra

p = número de variáveis

$|R|$ = determinante da matriz de correlação

Logo, para que este teste seja adequado é necessário rejeitar a hipótese básica. Desta forma, o p -valor do teste de Bartlett deve ser menor que o nível de significância estipulado, geralmente de 0,05 (DINI et al. 2014).

2.2.2 Medida de adequacidade da amostra de Kaiser-Meyer-Olkin (KMO).

“O teste Kaiser-Meyer-Olkin indica a proporção da variância dos dados que pode ser considerada comum a todas as variáveis” (REGIS, 2012).

A medida é representada por um índice (MSA) que mede a adequacidade da análise

fatorial, sendo obtida por meio da equação (4):
$$MSA = \frac{\sum \sum_{j \neq k} r^2_{jk}}{\sum \sum_{j \neq k} r^2_{jk} + \sum \sum_{j \neq k} q^2_{jk}}$$

Onde:

r^2_{jk} = quadrado dos elementos da matriz de correlação original (fora da diagonal), ou seja, representa o coeficiente de correlação linear simples entre as variáveis X_j e X_k

q^2_{jk} = quadrado dos elementos fora da diagonal da matriz anti-imagem da correlação, onde q_{jk} representa o coeficiente de correlação linear parcial entre as variáveis X_j e X_k

Valores de KMO acima de 0,6 são considerados apropriados para a análise fatorial, já valores menores que 0,6 consideram-se inapropriados para a análise (DINI et al. 2014).

3. Material e método

3.1 Matriz de dados

Os dados analisados neste trabalho foram obtidos através do sistema de pesquisas domiciliares. Este sistema foi inserido no Brasil a partir da Pesquisa Nacional por Amostras de Domicílios – PNDA, e tem como objetivo a produção de informações básicas para que seja possível estudar o desenvolvimento socioeconômico do País. A PNDA foi encerrada em 2016, com a divulgação das informações referentes ao ano de 2015. Produzindo então, resultados para Brasil, Grandes Regiões, Unidades de Federação e nove Regiões Metropolitanas, sendo elas, Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre (Instituto Brasileiro de Geografia e Estatística – IBGE). A pesquisa considerada neste artigo foi realizada por uma amostra probabilística de domicílios e a técnica de coleta utilizada foi entrevista pessoal assistida por computador – CAPI. A coleta teve início em 01 de outubro de 2015 e fim em 15 de dezembro de 2015, utilizando-se como unidade de investigação e de análise pessoal e domiciliar.

A partir dos dados coletados pelo IBGE por meio do sistema de pesquisas domiciliares, identificou-se as variáveis para a análise do cenário brasileiro. A matriz de dados inicial obtida no site do IBGE era composta por 41 linhas e 203 colunas, onde cada linha (observações) da matriz corresponde a um estado brasileiro e regiões metropolitanas e cada coluna representa as variáveis relacionadas ao acesso e utilização da Internet.

Para a aplicação da técnica de análise fatorial e para uma maior representatividade da realidade, tratou-se os dados da matriz de dados inicial. Inicialmente retiraram-se as regiões metropolitanas e dividiram-se os valores absolutos pelo número de habitantes de cada estado, desta forma gerou-se uma matriz com os dados per capita, evitando assim dados absolutos que poderiam interferir na classificação final. Após estruturar o conjunto de dados obteve-se uma matriz final composta por 27 linhas (observações) e 20 colunas (variáveis). O Quadro 1 apresenta as variáveis estudadas e suas categorias.

Quadro 1 – Variáveis finais estudadas

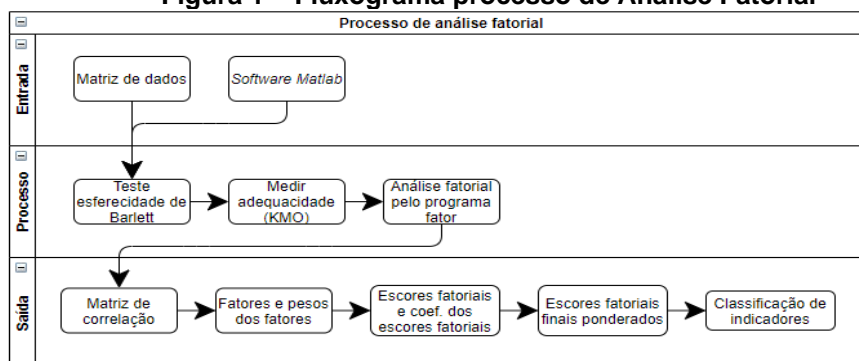
Domicílios particulares com Recepção de sinal digital de televisão aberta ou por assinatura	Acesso à Televisão
Domicílios particulares com Recepção de sinal de televisão por antena parabólica	
Domicílios particulares com Telefone móvel celular	Equipamento para acessar a Internet
Domicílios particulares com Tablet ou microcomputador	
Moradores em domicílios particulares com Banda larga móvel	Tipo de conexão com Internet
Moradores em domicílios particulares com Banda larga fixa e móvel	
Pessoas de 10 anos ou mais de idade Fundamental completo	Utilização da Internet
Pessoas de 10 anos ou mais de idade Médio completo	
Pessoas de 10 anos ou mais de idade Superior completo	
Estudantes de 10 anos ou mais de idade que utilizaram Internet, no período de referência dos últimos três meses	
Pessoas de 10 anos ou mais de idade com rendimento mensal de mais de 1 a 2 salários mínimos que utilizaram Internet, no período de referência dos últimos três meses	
Pessoas de 10 anos ou mais de idade com rendimento mensal de mais de 2 a 3 salários mínimos que utilizaram Internet, no período de referência dos últimos três meses	
Domicílios particulares permanentes com existência de utilização da Internet no domicílio	
Homens e mulheres de 10 anos ou mais de idade com Posse de telefone celular pessoal	Posse de Telefone celular
Pessoas de 10 anos ou mais de idade c/ telefone celular para uso pessoal fundamental completo	
Pessoas de 10 anos ou mais de idade c/ telefone celular para uso pessoal Médio completo	
Pessoas de 10 anos ou mais de idade c/ telefone celular para uso pessoal Superior completo	
Estudantes de 10 anos ou mais de idade com Posse de telefone celular para uso pessoal	
Pessoas de 10 anos ou mais de idade com rendimento mensal de mais de 1 a 2 salários mínimos com Posse de telefone celular para uso pessoal	
Pessoas de 10 anos ou mais de idade com rendimento mensal de mais de 2 a 3 salários mínimos com Posse de telefone celular pessoal	

Fonte: Os Autores (2021)

3.2 Método

A partir da matriz de dados final obteve-se a matriz de correlação dos dados, definindo-se como critério para a seleção do número de fatores o método das componentes principais. Posteriormente, obtêm-se os coeficientes dos escores fatoriais, os escores fatoriais, os escores fatoriais finais ponderados e ordenados. Resultando assim, na classificação dos estados. Esse processo está representado na Figura 1.

Figura 1 – Fluxograma processo de Análise Fatorial



Fonte: Os Autores (2021)

4. Resultados

O processo de obtenção dos resultados dividiu-se em três etapas: verificar se os dados obtidos eram passíveis de fatoração, estimar os fatores e a classificação dos estados e agrupar os estados a partir de sua infraestrutura em relação ao acesso à informação.

4.1 Verificação da viabilidade de fatoração dos dados

Para verificar se os dados eram passíveis de fatoração testou-se a esfericidade de Bartlett e a medida de adequacidade da amostra de Kaiser-Meyer-Olkin (KMO) a partir de funções programas do *software Matlab*. No teste de esfericidade a matriz teve seu p-valor menor que 0,05, ou seja, rejeitou-se a hipótese nula de que a matriz de dados é similar a uma matriz-identidade, ou seja, as correlações entre as variáveis são significativas, como mostra a Tabela 1. Assim como, ao medir a adequacidade da amostra, obteve-se um resultado de 0,68, considerando desta forma aceitável a análise. Portanto, conclui-se que a matriz é passível de fatoração.

Tabela 1 – Resultado teste de Barlett

Sample-size	Variáveis	X2	df	p
27	20	12.588.462	190	0.000

Fonte: Os Autores (2021)

4.2 Análise fatorial

4.2.1 Estimação dos fatores

Segundo Kaiser (1960), utiliza-se apenas os fatores que possuem raízes características (autovalores) maiores que um, isto porque estas raízes refletem a importância relativa de cada fator. Considerando isso, adotou-se o critério de Kaiser para seleção do número de fatores, obtendo-se três fatores, os quais explicam 86% da amostra. O valor de cada autovalor e a explicação da variância acumulada estão apresentados na Tabela 2.

Tabela 2 – Explicação acumulada

	1 Fator	2 Fatores	3 Fatores
Autovalores	12,37	3,12	1,72
Variância acumulada	0,6185	0,7744	0,8603

Fonte: Os Autores (2021)

Para avaliar o quanto cada variável é representada por cada fator, calcularam-se suas comunalidades. Quanto mais perto de um, melhor a variável é explicada pelos fatores. Variáveis com comunalidades menores que 0,5 devem ser descartadas da análise, e refazer a análise de fatores (ENGEL et al. 2013). Na amostra, todas as variáveis seguiram na análise, visto que elas possuíam comunalidade maior que 0,5, como mostra a Tabela 3.

Tabela 3 – Análise fatorial, método das componentes principais

Variável	Pesos Estimados			Com.	Var. Exp.
	F1	F2	F3		
1	0,9415	-0,0465	0,1731	0,92	0,08
2	-0,4826	0,6347	-0,3783	0,78	0,22
3	0,9646	-0,0442	-0,1141	0,95	0,05
4	0,9406	0,0323	-0,0182	0,89	0,11
5	0,7362	-0,3534	-0,0366	0,67	0,33
6	0,9192	-0,128	-0,0379	0,86	0,14
7	0,6106	0,5023	0,4613	0,84	0,16
8	0,7704	-0,1	0,4294	0,79	0,21
9	0,8941	-0,2676	-0,0727	0,88	0,12
10	0,6154	-0,5098	-0,4489	0,84	0,16

Variável	F1	F2	F3	Com.	Var. Exp.
11	0,0677	0,7996	-0,4142	0,82	0,18
12	0,835	0,4	0,1537	0,88	0,12
13	0,9689	0,0507	-0,0858	0,95	0,05
14	0,917	0,1977	-0,3112	0,98	0,02
15	0,7515	0,5388	0,1762	0,89	0,11
16	0,8534	-0,0519	0,28	0,81	0,19
17	0,8927	-0,2723	-0,0858	0,88	0,12
18	0,5391	-0,5527	-0,4945	0,84	0,16
19	0,5488	0,5753	-0,5036	0,89	0,11
20	0,8607	0,3598	0,1029	0,88	0,12

Fonte: Os Autores (2021)

Ao agruparmos as variáveis a partir do fator que melhor representa um determinado grupo (maior peso dentro do fator). Notou-se que as variáveis relacionadas a situações de rendas menores estão sendo melhor representadas pelo fator 2, sendo as variáveis 2, 11, 18 e 19. Já as outras variáveis estão sendo representadas todas pelo fator 1. Os fatores são variáveis latentes, ou seja, combinações lineares das variáveis originais, portanto nem sempre é possível nomeá-los ou encontrar relação clara entre as variáveis. Desta maneira, percebe-se que o fator 3 é uma variável latente, por cumprir os requisitos supracitados.

4.2.2 Escores Fatoriais

Após a análise de fatores e comunalidades, obtêm-se os escores finais dos fatores. Para isso, calcularam-se os coeficientes dos escores fatoriais (Tabela 4), pois eles estimam o peso relativo de cada variável no fator, quanto maior o valor absoluto do coeficiente maior o peso no cálculo dos escores finais será a variável correspondente (SUPORTE AO MINITAB 18). Nota-se que a variável mais importante da amostra é a variável 19, contida no fator 3.

Tabela 4 – Coeficiente dos Escores Fatoriais

Variável	Fator 1	Fator 2	Fator 3
1	0,0761	-0,0149	0,1008
2	-0,039	0,2036	-0,2202
3	0,078	-0,0142	-0,0664
4	0,076	0,0104	-0,0106
5	0,0595	-0,1134	-0,0213
6	0,0743	-0,0411	-0,0221
7	0,0494	0,1611	0,2685
8	0,0623	-0,0321	0,2499
9	0,0723	-0,0859	-0,0423
10	0,0498	-0,1635	-0,2613
11	0,0055	0,2565	-0,2411
12	0,0675	0,1283	0,0895
13	0,0783	0,0163	-0,05
14	0,0741	0,0634	-0,1811
15	0,0607	0,1728	0,1025
16	0,069	-0,0166	0,163
17	0,0722	-0,0874	-0,0499
18	0,0436	-0,1773	-0,2878
19	0,0444	0,1846	-0,2931
20	0,0696	0,1154	0,0599

Fonte: Os Autores (2021)

Após o cálculo dos coeficientes dos escores fatoriais, foi possível determinar os escores fatoriais (Tabela 5). O escore fatorial é a uma medida composta para cada observação, em que se utilizam os pesos fatoriais em conjunção com a variável original para obtenção do valor (SISTEMA DE APOIO A APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS, 2021).

Tabela 5 – Escores fatoriais estimados das observações

Fator 1	Fator 2	Fator 3
-0,5087	0,904	-0,1803
-0,8537	-0,6085	0,4318
-0,6091	-14474	14615
0,0634	-14599	-0,9834
-10284	-0,2026	11768
-0,2636	-14828	0,8172
-0,5323	0,5992	-11254
-18393	-0,4001	17733
-13591	0,0827	-0,6048
-0,8959	0,7018	0,4405
-0,2084	-0,4804	-0,8564
-0,4154	-0,3226	-11809
-0,4985	0,2866	-0,0521
-12844	0,0835	-0,6772
-0,3724	-0,1969	-10379
-0,6343	0,0077	-0,3291
0,4581	10952	-0,7464
0,7053	0,5802	0,2386
15916	0,4016	13749
16223	-0,3051	16800
0,7884	0,6581	0,4777
11921	11589	12708
0,9626	13941	-0,1616
0,6343	0,4235	-15028
0,3681	0,9027	0,4096
0,7862	0,7555	-11621
21311	-31290	-0,9524

Fonte: Os Autores (2021)

A partir dos escores fatoriais, estimados das observações, foi possível determinar os escores fatoriais finais ponderados.

Podemos também chamá-los de indicadores sintéticos, os quais são calculados pela equação: $Indicador = \hat{F}_1 x \% Var. Expl._1 + \hat{F}_2 x \% Var. Expl._2 + [\dots] + \hat{F}_n x \% Var. Expl._n$

Onde:

F = Escores estimados dos fatores das observações

$Var. Expl.$ = Variância explicada pelo fator

Neste estudo, utilizaram-se três fatores, sendo a variância explicada do fator 1 61,85%, a variância explicada pelo fator 2 15,54% e a variância explicada pelo fator 3 8,59% da amostra. Logo, a variância acumulada dos três fatores é igual a 86,03%. A partir destes escores é possível ranquear os estados brasileiros. Por fim, ordenaram-se os escores fatoriais finais para a classificação dos estados brasileiros. A Tabela 6 classifica os estados

de forma decrescente, ou seja, o estado que possuir um menor escore final terá sua infraestrutura pior que o estado que possuir um escore final maior que o seu.

Como nenhuma variável possui maior peso somente no fator 3, efetuou-se uma nova análise com dois fatores para efeito comparativo e análise do comportamento da amostra. Seguiram-se os passos demonstrados anteriormente. Porém o critério escolhido para seleção do número de fatores foi feito empiricamente e considerou-se como critério que o número de fatores escolhido deveria explicar pelo menos 70% da variância total.

Tabela 6 – Classificação dos estados brasileiros

Classificação	Escore final	Estado
1	-10,476	Maranhão
2	-0,8797	Piauí
3	-0,8396	Alagoas
4	-0,5857	Acre
5	-0,5666	Pará
6	-0,4767	Amazonas
7	-0,4194	Bahia
8	-0,4086	Paraíba
9	-0,4069	Ceará
10	-0,3502	Sergipe
11	-0,3325	Tocantins
12	-0,324	Amapá
13	-0,2774	Rio Grande do Norte
14	-0,2728	Roraima
15	-0,2682	Pernambuco
16	-0,1892	Rondônia
17	0,3293	Mato Grosso do Sul
18	0,3899	Minas Gerais
19	0,4035	Mato Grosso
20	0,5042	Goiás
21	0,5472	Espírito Santo
22	0,6313	Paraná
23	0,7486	Distrito Federal
24	0,7987	Rio Grande do Sul
25	10,271	Santa Catarina
26	11,001	São Paulo
27	11,651	Rio de Janeiro

Fonte: Os Autores (2021)

Com o novo critério selecionaram-se dois fatores, sendo que estes possuem a mesma variância explicada que o fator 1 e 2 obtidos anteriormente. Esse resultado era esperado, pois o método utilizado foi igual ao anterior. Calcularam-se os pesos e as comunalidades das variáveis. As variáveis seguiram na análise por possuírem suas comunalidade acima de 0,5. Por fim, calculou-se os escores fatoriais finais ponderados para a nova simulação (Tabela 7).

Tabela 7 – Classificação 2 dos estados

Classificação	Escore final	Estado
1	-12	Maranhão
2	-0,8277	Piauí
3	-0,7814	Alagoas
4	-0,6677	Pará
5	-0,6228	Acre
6	-0,6023	Amazonas
7	-0,4447	Ceará
8	-0,3942	Amapá
9	-0,3912	Bahia
10	-0,3072	Paraíba
11	-0,2637	Pernambuco
12	-0,261	Sergipe
13	-0,2359	Tocantins
14	-0,2038	Rio Grande do Norte
15	-0,1883	Roraima
16	-0,1737	Rondônia
17	0,3683	Mato Grosso
18	0,4541	Minas Gerais
19	0,4584	Mato Grosso do Sul
20	0,5267	Espírito Santo
21	0,5902	Paraná
22	0,604	Goiás
23	0,8126	Rio Grande do Sul
24	0,8304	Distrito Federal
25	0,9179	Santa Catarina
26	0,9558	São Paulo
27	10,47	Rio de Janeiro

Fonte: Os Autores (2021).

Para facilitar a análise, dividiu-se os estados em três categorias, seguindo em ordem decrescente de escores. Podemos perceber que não há divergências significativas entre as duas análises quando comparadas nas categorias citadas, e que apenas três estados transitam entre estas categorias (Figura 2)

Figura 2 – Comparação de análises fatoriais

	ANÁLISE FATORIAL - 2 FATORES			ANÁLISE FATORIAL - 3 FATORES		
	Classificação	Escore final	Estado	Classificação	Escore final	Estado
Categoria 1 Emergencial	1	-12	Maranhão	1	-10,476	Maranhão
	2	-0,8277	Piauí	2	-0,8797	Piauí
	3	-0,7814	Alagoas	3	-0,8396	Alagoas
	4	-0,6677	Pará	4	-0,5857	Acre
	5	-0,6228	Acre	5	-0,5666	Pará
	6	-0,6023	Amazonas	6	-0,4767	Amazonas
	7	-0,4447	Ceará	7	-0,4194	Bahia
	8	-0,3942	Amapá	8	-0,4086	Paraíba
	9	-0,3912	Bahia	9	-0,4069	Ceará
Categoria 2 Em alerta	10	-0,3072	Paraíba	10	-0,3502	Sergipe
	11	-0,2637	Pernambuco	11	-0,3325	Tocantins
	12	-0,261	Sergipe	12	-0,324	Amapá
	13	-0,2359	Tocantins	13	-0,2774	Rio Grande do Norte
	14	-0,2038	Rio Grande do Norte	14	-0,2728	Roraima
	15	-0,1883	Roraima	15	-0,2682	Pernambuco
	16	-0,1737	Rondônia	16	-0,1892	Rondônia
	17	0,3683	Mato Grosso	17	0,3293	Mato Grosso do Sul
	18	0,4541	Minas Gerais	18	0,3899	Minas Gerais
Categoria 3 Sob Controle	19	0,4584	Mato Grosso do Sul	19	0,4035	Mato Grosso
	20	0,5267	Espírito Santo	20	0,5042	Goiás
	21	0,5902	Paraná	21	0,5472	Espírito Santo
	22	0,604	Goiás	22	0,6313	Paraná
	23	0,8126	Rio Grande do Sul	23	0,7486	Distrito Federal
	24	0,8304	Distrito Federal	24	0,7987	Rio Grande do Sul
	25	0,9179	Santa Catarina	25	10,271	Santa Catarina
	26	0,9558	São Paulo	26	11,001	São Paulo
	27	10,47	Rio de Janeiro	27	11,651	Rio de Janeiro

Fonte: Os Autores (2021)

5 Conclusão

Após a aplicação da análise fatorial – para três fatores e dois fatores – considerando-se a construção dos fatores, foi possível concluir que as políticas de melhorias relacionadas à acessibilidade da informação devem ser direcionadas às variáveis que possuem maior peso no fator 1. Isto porque, tais variáveis explicam 61,85% da variabilidade do conjunto de dados. Dentro do fator 1, selecionou-se as cinco variáveis com o maior peso dentro deste fator, e é para estas que devem ser feitos planos de ação primeiramente, pois são elas que possuem maior impacto na análise, como apresentado na Tabela 8.

Tabela 8 – Variáveis com maiores pesos no Fator 1

Variável	Peso Fator 1	Variáveis explicadas
13	0,9689	Domicílios particulares permanentes que havia existência de utilização da Internet no domicílio
3	0,9646	Domicílios particulares com telefone móvel celular
1	0,9415	Domicílios particulares com recepção de sinal digital de televisão aberta ou televisão por assinatura
4	0,9406	Domicílios particulares com tablet ou microcomputador
6	0,9192	Moradores em domicílios particulares com banda larga fixa e móvel

Fonte: Os Autores (2021)

Vale ressaltar que todas as variáveis que possuem maior impacto são relacionadas diretamente a aparelhos (como telefone celular, *tablets* e microcomputadores) e a forma de conexão com algum meio de informação. Portanto, o plano de ação deve levar em consideração o fornecimento dos itens relacionados a estas variáveis.

Considerando o resultado dos escores fatoriais finais ponderados, concluiu-se que os três estados com os menores escores permaneceram os mesmos, sendo eles: Maranhão, Piauí e Alagoas, nesta ordem de classificação. Portanto, estes devem ser os estados com maior ação de políticas de melhoria de acessibilidade à informação. Da mesma maneira, os estados mais bem classificados continuaram os mesmos nas duas análises.

Para facilitar a classificação dos estados e a priorização do direcionamento do plano de ação, a classificação final foi dividida em três categorias. Desta forma, podem-se criar planos de ação diferentes para cada categoria, considerando características diferentes para cada um, como por exemplo, tamanho de equipe, velocidade do resultado, definição de cronograma, entre outros. Vale ressaltar que, apesar da classificação de alguns estados mudarem de uma análise para a outra, eles mantiveram nas mesmas categorias, com exceção dos estados de Paraíba, Mato Grosso e Mato Grosso do Sul.

Referências

DAMÁSIO, Bruno Figueiredo. **Uso da análise fatorial exploratória em psicologia**. 2012. Disponível em: <<http://pepsic.bvsalud.org/pdf/avp/v11n2/v11n2a07.pdf>>. Acesso em: 22 ago. 2021.

DINI, Ariane Polidoro et al. Validity and reliability of a pediatric patient classification instrument. **Revista Latino-americana de Enfermagem**, [s.l.], v. 22, n. 4, p.598-603, ago. 2014. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0104-1169.3575.2457>. Disponível em: <http://www.scielo.br/pdf/rlae/v22n4/pt_0104-1169-rlae-22-04-00598.pdf>. Acesso em: 07 ago. 2021.

FACHEL, Jandyra Maria Guimarães. **Análise Fatorial**. 1976. 86 f. Dissertação (Mestrado) - Curso de Matemática, Universidade de São Paulo, São Paulo, 1976. Disponível em: <<https://lume.ufrgs.br/handle/10183/118553>> Acesso em: 20 set. 2021.

HAIR, Joseph F. et al. **Análise multivariada de dados**. 6. ed. São Paulo: Artmed Editora Sa, 2009. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=oFQs_zJI2GwC&oi=fnd&pg=PA7&dq=an%C3%A1lise+multivariada&ots=KI4GQi69ps&sig=2-g27CIsKODOw0Klw5csUnckvOQ#v=onepage&q&f=false> Acesso em: 20 out. 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. PNDA - Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal 2015. **IBGE, Coordenação de Trabalho e Rendimento**. Rio de Janeiro, 2016. 87 p.

KAISER, Henry F. **The application of electronic computers to factor analysis**. Educational and Psychological Measurement, v.20.p.111-117.1960

MARQUES, Marcos Augusto Mendes. **Aplicação da análise multivariada no estudo da infra-estrutura dos serviços de saúde dos municípios paranaenses.2005**. 133 f. Dissertação (Mestrado) - Curso de Engenharia, Setor de Tecnologia e Setor de Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2005. Disponível em: <<https://acervodigital.ufpr.br/bitstream/handle/1884/6199/Servi%C3%A7os%20de%20Sa%C3%BAde.pdf?sequence=2&isAllowed=y>>. Acesso em: 05 set. 2021.

NETO, J.F.B. MENEZES, A.M.F. O acesso à informação e a difusão do conhecimento no fortalecimento da política de educação profissional da Bahia. **Artigos BA&D**, Salvador, v. 28, n. 1, p.92-109, jan.-jun. 2018

PANORAMA SETORIAL DA INTERNET. Acesso à Internet no Brasil: Desafios para conectar toda a população. Universalização do acesso. Ano 8 – Número 1. UNESCO, 2016.

REGIS, Marcela Neves Oliveira. **Impacto do planejamento estratégico sobre o desempenho de bancos no Brasil**. 2012. Dissertação - Curso de Administração, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: <https://www.maxwell.vrac.puc-rio.br/20504/20504_6.PDF>. Acesso em: 24 set. 2021.

SILVA, D.M. RIBEIRO, A.C.D. FILHO, E.A.S. As Redes sociais como ferramenta para o acesso à informação na administração pública. **Perspectivas em Políticas Públicas**, Belo Horizonte, v. XI , n. 21, p. 267-294. jan/jun 2018

SISTEMA DE APOIO A APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS: **Escore fatorial**. Disponível em: <<http://www5.eesc.usp.br/saate/index.php/saate/Indicar-a-T%C3%A9cnica/Descrever/2.-%C3%81rvore-de-Decis%C3%A3o/Gloss%C3%A1rio/Escore-fatorial>>. Acesso em: 05 set. 2021.

SUPORTE AO MINITAB[®] 18. **Interpretar todas as estatísticas e gráficos para Análise fatorial**. Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/modeling-statistics/multivariate/how-to/factor-analysis/interpret-the-results/all-statistics-and-graphs/>>. Acesso em: 06 set. 2021.

HENRIQUES, V. R. **Panorama setorial da Internet**: Acesso à Internet no Brasil: Desafios para conectar toda a população. São Paulo, 2016. 12 p. Disponível em: <https://www.cetic.br/media/docs/publicacoes/6/Panorama_Setorial_11.pdf>. Acesso em: 03 jul. 2021