

ANÁLISE DE MÚLTIPLOS FATORES EM DADOS QUALITATIVOS ORDINAIS: ESTUDO DE CASO DO COMPORTAMENTO DOS CONSUMIDORES EM RELAÇÃO À QUALIDADE DE CAFÉS

Paulo César Ossani¹, Flávio Meira Borém², Gislene Araújo Pereira³, Marcelo Ângelo Cirillo⁴

Resumo: O café tem posição de destaque no cenário socioeconômico brasileiro, pois, além de ser relevante na geração empregos e em toda economia, direta ou indiretamente, também é importante nas relações sociais. O estudo do comportamento dos consumidores é uma ferramenta importante para a identificação das motivações nos processos de compra do café nos diferentes segmentos desse mercado. Entretanto, ao considerar grupos de variáveis qualitativas em uma abordagem multivariada, as técnicas estatísticas usuais são limitadas ou inviáveis para descrever adequadamente o perfil sensorial dos consumidores. Este trabalho foi realizado com o objetivo de utilizar a técnica de Análise Múltiplos Fatores, no estudo do comportamento dos consumidores de café em uma pesquisa de opinião pública aplicada a um grupo de consumidores localizados na cidade de Varginha, MG. Foram usados seis assertivas (*Likert* de cinco pontos) relacionadas às percepções dos consumidores em relação à qualidade do café. Os resultados demonstram similaridade moderada nas respostas dadas pelos indivíduos nos subconjuntos das variáveis (marca e qualidade) e (preferência e preço), e baixa similaridade nos subconjuntos das variáveis (quantidade de xícaras de café consumidas), (torrefadora e produtor) e (região e produção). Concluiu-se que o uso desta técnica é viável, por permitir discriminar os comportamentos dos consumidores de cafés.

Palavras-Chave: perfil sensorial, pesquisa de opinião, técnica MFA.

Abstract: Coffee has a prominent position in the Brazilian socioeconomic scenario, for, in addition to being relevant in the generation of jobs and in all economics, whether directly or indirectly, it is also important in social relationships. The study of the consumers' behavior is an important tool to the identification of the motivations in the coffee purchasing processes in the different segments of this market. However, when considering groups of qualitative variables in a multivariate approach, the usual statistical techniques are limited or infeasible to describe adequately the sensory profile of consumers. This work was carried out with the objective of using the Multiple Factor Analysis technique in the study of the behavior of coffee consumers in a public opinion poll applied to a group of consumers located in the city of Varginha, MG. Six assertions (five-point Likert) were used related to consumers' perceptions regarding coffee quality. The results show moderate similarity in the responses given by individuals in the subsets of the variables (brand and quality) and (preference and price), and low similarity in the subsets of the variables (quantity of cups of coffee consumed), (roaster and producer) and (region and production). It follows that the use of this technique is viable, since it allows discriminating the behaviors of the coffee consumers.

Keywords: sensory profile, opinion poll, MFA technique.

¹ DES-UEM. e-mail: ossanipc@hotmail.com

² DEA-UFLA. e-mail: flavioborem@deg.ufla.br

³ ICSA-UNIFAL. e-mail: gislene.araujo.p@gmail.com

⁴ DEX-UFLA. e-mail: macufla@des.ufla.br

1 Introdução

Segundo dados da Associação Brasileira da Indústria de Café (ABIC), os brasileiros consomem 4,82 kg de café torrado e moído por habitante/ano (ABIC, 2018). Os efeitos benéficos, sejam no âmbito social, do agronegócio ou da saúde, são destacados na mídia, e são vistos como fatores motivadores para o consumo e a produção de cafés no Brasil.

Wanderley *et al* (2017) apontam que o consumo de café contribui para a redução do desenvolvimento de doenças degenerativas de *Alzheimer* e *Parkinson*, diabetes tipo II, câncer, redução de peso, asma, tratamento adjuvante de hipertensão, doenças cardiovasculares e cirrose, também contribui para o desempenho de exercícios físicos, e melhora o humor.

Pesquisa realizada pela ABIC, em 2014, demonstrou que o café é consumido por hábito, prazer e tradição, e que o sabor e o aroma são marcantes em sua qualidade. Esses fatores, naturalmente, sugerem foco no estudo do comportamento dos consumidores, que é uma importante ferramenta, no tocante à compreensão dos aspectos envolvidos no processo de compra e na identificação dos diferentes segmentos do mercado de café. Segundo Noronha (2005), existem diferenças nas percepções individuais, resultantes de fatores genéticos ou culturais, que influenciam os hábitos alimentares.

Convém ressaltar que a importância do café, refletida na interação social, é mencionada por Arruda *et al* (2009), ao relatarem que os consumidores associam o consumo do café a hábito, prazer, família, amizade e trabalho. Isso indica que as motivações para o seu consumo têm um significado social associado aos hábitos e aos costumes brasileiros.

De acordo com Fehr *et al* (2012), o agronegócio café tem relevante importância social devido à sua capacidade de gerar empregos, além de atuar no desenvolvimento socioeconômico brasileiro. Os autores também ressaltam um aumento cada vez maior no consumo interno.

Segundo Alan, Kabadayi e Yilmaz (2016) a preferência pelo consumo de café de uma determinada marca envolve a decisão do consumidor, relacionada a reações cognitivas, afetivas, sociais e físicas ao processo de consumo. Para exemplificar melhor essa relação, observa-se que, na maioria das situações, os consumidores, normalmente, escolhem características simples nas quais baseiam suas percepções de marca, e adotam as marcas com personalidades que refletem a sua autoimagem. Os consumidores avaliam melhor os produtos oriundos de seus países de origem, e também os produtos e as marcas provenientes de países mais desenvolvidos em relação aos menos desenvolvidos (BOAS *et al*, 2009).

Saes e Spers (2006) notaram que os produtos diferenciados pela marca da torrefadora são mais atrativos, quando comparados à região de origem ou ao processo de produção agrícola.

Segundo EMBRAPA (2019) o consumidor de café no Brasil modificou seus hábitos de consumo e adquiriu novas percepções a respeito da bebida. Com isso, novas estratégias buscam a valorização do produto com atributos diferenciados, com aspectos tangíveis ou intangíveis e, assim, a concorrência não se dá apenas via preços, mas também por meio de produtos com características inovadoras.

Para Spers, Saes e Souza (2004), o segmento de cafés diferenciados é o que mais cresce e a compra do café está ligada aos atributos marca e sabor, que se relacionam às experiências passadas inerentes à memória sensorial que caracterizam cada consumidor, que associa a fidelidade à marca ao sabor preferido. Portanto, há evidências de que os consumidores brasileiros estão preparados para cafés de qualidade, uma vez que existem diferenças entre as preferências dos consumidores e os segmentos de cafés deveriam ser atendidos por meio de estratégias de marketing que envolvam padrões de diferenciação, o que elevaria a qualidade, de forma a agregar valor à satisfação do consumidor.

Arruda *et al* (2009) afirmam que marca, qualidade e preço são os atributos mais observados na escolha do café e a presença do selo de pureza, além de aumentar a credibilidade, valida a qualidade do produto.

Moori, Bido e Oliveira (2011) realizaram um estudo para detectar as variáveis relevantes do consumidor do café solúvel por meio da modelagem de equações estruturais. Nesse estudo, as variáveis que caracterizam a classe social apresentam um pequeno efeito do consumo de café, entretanto, a idade do respondente apresentou efeito significativo. Contudo, o estudo evidenciou que, no mercado interno brasileiro, o principal motivo que leva as pessoas a consumirem o café solúvel está associado ao hábito.

Em relação às metodologias estatísticas empregadas nos setores de qualidade, nota-se uma carência quando os dados são de natureza qualitativa. Mormente, utilizam-se testes de hipóteses não paramétricos, sendo esses limitados a conclusões que tratam simultaneamente as relações entre as variáveis envolvidas na mensuração do processo de qualidade. Nesse contexto, em virtude do que foi mencionado, este trabalho foi realizado com o objetivo de apresentar uma análise estatística utilizando a técnica multivariada de Análise de Múltiplos Fatores (MFA, do inglês *Multiple Factor Analysis*) em um estudo de caso envolvendo pesquisa de comportamento de uma amostra de consumidores em relação à qualidade de cafés, a fim de encontrar as similaridades dos grupos de variáveis qualitativas ordinais que compõem esta amostra.

A vantagem proporcionada por essa técnica consiste em possibilitar trabalhar com variáveis categóricas com escalas diferentes, gerando resultados com análise similar ao PCA (*Principal Components Analysis*), de tal forma que a visualização dos grupos de variáveis similares é dada em um espaço de duas ou três dimensões.

2 Referencial Teórico

2.1 Análise de múltiplos fatores

Dentre as inúmeras técnicas propostas para analisar dados, a Análise de Múltiplos Fatores se caracteriza por permitir analisar grupos de variáveis (colunas) com tamanhos diferentes e de natureza distintas, que podem ser quantitativas, categóricas, de frequência ou mistas, definidos no mesmo conjunto de observações (linhas) (ESCOFIER; PAGÈS, 2008).

Este método pode ser muito útil para a análise de estudos em que se podem identificar vários grupos de variáveis, ou para os estudos em que as mesmas perguntas são feitas em intervalos de tempo diversos (ABDI; WILLIAMS; VALENTIN, 2013)

O cerne da técnica MFA é uma análise de fator aplicada a um conjunto de variáveis balanceado, conduzindo a uma representação das observações e das variáveis, como em qualquer análise de fator. Devido ao balanceamento, esta análise de fator pode ser interpretada como uma análise canônica (ESCOFIER; PAGÈS, 2008).

Convém ressaltar que, na Análise de Múltiplos Fatores, o número de variáveis em cada grupo pode diferir e a natureza das variáveis (quantitativa, categórica ou frequência) pode variar de um grupo para o outro, mas as variáveis devem ser de mesma natureza no grupo dado.

No PCA o objetivo principal é explicar a estrutura da variância e covariância de um vetor aleatório, composto de p -variáveis aleatórias, por meio da construção de combinações lineares das variáveis originais. Estas combinações lineares são chamadas de componentes principais e não são correlacionadas entre si. A técnica MFA opera no mesmo princípio, mas, em vez de estabelecer as correlações entre as variáveis originais como no PCA, ela fornece a similaridade entre os grupos de variáveis. A técnica MFA utiliza a estrutura da variância e covariância de um vetor aleatório, composto de p -variáveis aleatórias por meio da construção

de combinações lineares das variáveis originais pertencentes a cada grupo de variáveis, as quais são balanceadas conforme a natureza das variáveis em cada grupo, com isso obtendo-se resultados utilizados na avaliação de similaridade entre os grupos. Assim, pode-se dizer que a técnica MFA é uma generalização do PCA para grupos de variáveis.

2.1.1 Notação

Para cada grupo a notação é representada por $X_1, \dots, X_j, \dots, X_J$, compreendendo um único conjunto de dados, conforme ilustrado na Figura 1.

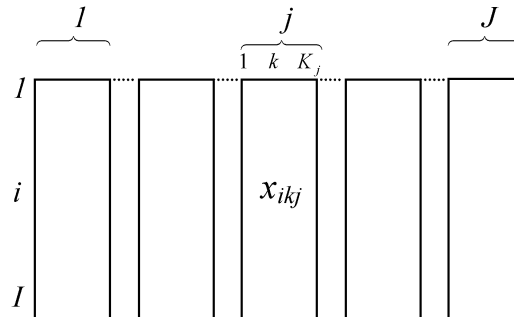


Figura 1 - Layout dos grupos de variáveis justapostas, representando um único grupo.

As I observações foram representadas por J conjuntos de variáveis, o índice j referiu-se a um conjunto, o índice k referiu-se a uma coluna e, por fim, K_j correspondeu ao número de colunas no conjunto j , de modo que $K = \sum_{j \in J} K_j$ indica o número de colunas em todos os conjuntos.

Com essas especificações, uma tabela $I \times K_j$ está associada com cada conjunto j . As tabelas J juntas compõem uma tabela múltipla ou global $I \times K$. Para um j -ésimo conjunto quantitativo, K_j é tanto o número de colunas como o número de variáveis. Entretanto, para um conjunto categórico, K_j indica o número de colunas e o número de categorias. Este tipo de conjunto é representado por uma tabela de variáveis indicadoras, na qual a coluna k está associada com a categoria k (BÉCUE-BERTAUT; PAGÈS, 2008). Logo, no cruzamento da linha i e coluna k (que pertencem a j -ésima tabela), tem-se:

- se j é um conjunto quantitativo, o valor x_{ikj} da variável k para a observação i ;
- se j é um conjunto categórico, $x_{ikj} = 1$ se i pertence à categoria k e 0, caso contrário.

Dados perdidos são permitidos no caso de variáveis categóricas, assim as observações em que não há nenhuma categoria k , terão 0 (zero) para toda as variáveis indicadoras associadas à categoria k .

Dados de frequência e mistos não serão abordados nesse texto, embora a diferença seja apenas quanto ao seu balanceamento, o que pode ser visto, em Bécue-Bertaut e Pagès (2008).

2.1.2 Balanceamento dos dados

Consideram-se os dados centralizados por coluna, isto é, subtrai-se de cada elemento da coluna a sua respectiva média, ou seja,

$$C_{ik} = x_{ik} - \bar{x}_K. \quad (1)$$

Em seguida, normalizam-se as colunas, divide-se cada elemento da coluna pela raiz quadrada da soma do quadrado da respectiva coluna obtida de (1), conforme a expressão

$$x_{ik}^* = \frac{C_{ik}}{\sqrt{\sum_{i=1}^I (C_{ik})^2}} \quad (2)$$

Desse modo, cada coluna irá se comportar como um vetor de módulo 1. Agora, se os dados forem quantitativos, aplica-se um PCA, ou Análise de Correspondência Múltipla se forem de variáveis categóricas, em cada grupo X_j de variáveis, formado por x_{ik}^* valores, com o propósito de determinar os primeiros autovalores λ_{j1} . Assim, torna-se viável calcular $\varphi_k = a_k / \sqrt{\lambda_{j1}}$, em que a_k é o peso da variável k . Para variáveis quantitativas, $a_k = 1$ para todo o k . Para variáveis categóricas, a_k é a proporção das observações que não têm categoria k . Assim segue que

$$Z_j = X_j \times \varphi_k. \quad (3)$$

A matriz global será uma matriz $I \times K$, denotada por

$$Z = [Z_1, \dots, Z_j]. \quad (4)$$

2.1.3 MFA como um PCA

A técnica MFA consiste em um PCA na matriz global Z , cujo termo geral é dado por (4). Do PCA, também chamado de análise global, os componentes principais são obtidos de modo usual por linhas e colunas de F , escores dos fatores globais dados pela equação (6) (ABDI; WILLIAMS, 2010).

Estimada a matriz global Z , aplica-se a *decomposição de valor singular*, dada pela expressão

$$Z = \hat{U}\Lambda\hat{V}^T, \quad (5)$$

em que $\hat{U}^T\hat{U} = \hat{V}^T\hat{V} = I_{id}$ é a matriz de autovetores e $\Lambda = \text{diag}(\sqrt{\lambda_i})$, sendo $\lambda_i > 0$ os autovalores e I_{id} é a matriz identidade.

2.1.4 Os escores dos fatores globais

Os escores dos fatores globais são estimados em por

$$F = M^{\frac{1}{2}}\hat{U}\Lambda, \quad (6)$$

sendo M a matriz diagonal das massas das observações, com $m_i = I$, tendo I como o número de linhas da matriz Z . Em F , cada linha representa as observações e cada coluna, as variáveis.

2.1.5 Análise parcial

A análise global revela a estrutura comum das observações. Para ver como cada variável interpreta este espaço, projeta-se o conjunto de dados de cada variável sobre a análise global. Isto é alcançado obtendo-se a matriz de projeção reescrevendo-se a equação (6), podendo os escores dos fatores globais serem computados como

$$F = M^{\frac{1}{2}} \hat{U} \Lambda = ZZ^T \times M^{\frac{1}{2}} \hat{U} \Lambda^{-1}. \quad (7)$$

Isso mostra que

$$P = M^{\frac{1}{2}} \hat{U} \Lambda^{-1}, \quad (8)$$

é a matriz de projeção que transforma a matriz ZZ^T em escores dos fatores.

A matriz de projeção é utilizada para projetar os grupos sobre o espaço global, fazendo

$$F_j = J \times Z_j Z_j^T \times P, \quad (9)$$

sendo J o número de conjuntos de variáveis.

2.1.6 Inércia parcial entre os grupos de variáveis

As relações entre as variáveis e a solução global são analisadas calculando-se a inércia parcial de cada variável por cada dimensão da análise global. Isto é calculado, para cada variável, como a soma das projeções quadradas das variáveis do autovetor \hat{V} de Z multiplicado pelo autovalor correspondente. Como o autovetor \hat{V} é ortonormal, a soma das inércias parciais para todos os grupos para uma determinada dimensão é igual ao seu autovalor, ou seja (ABDI; WILLIAMS, 2010),

$$R_{ji} = \lambda_p \times \sum_j^{J_k} \hat{V}_{ji}^2, \quad (10)$$

em que j refere-se ao grupo, p o componente principal e λ_p é o p -ésimo autovalor da decomposição de valores singulares realizada em Z .

3 Procedimentos Metodológicos

3.1 Preparo e realização do experimento

Em concordância com os objetivos propostos, uma pesquisa descritiva foi realizada em julho de 2011, com consumidores abordados de forma aleatória em locais de maior movimentação de pessoas no centro da cidade de Varginha, localizada no sul de Minas

Gerais, região com a maior área plantada de café. Se o entrevistado não consumia café, era dispensado, assim uma amostra de 50 consumidores de café foi obtida, com idade a partir de 18 anos, com diferentes graus de escolaridade e renda, com o propósito de definir o comportamento desses consumidores. Esta pesquisa foi constituída de duas questões relativas a dados demográficos, uma questão com o objetivo de mensurar a importância atribuída pelo consumidor à marca do café e, com base no trabalho de Saes e Spers (2006), seis assertivas (Likert de cinco pontos) relacionadas às percepções dos consumidores em relação à qualidade do café.

Diante do exposto, seguindo as especificações dadas por Saes e Spers (2006), consideraram-se as variáveis qualitativas em uma escala ordinal, de modo que as questões elaboradas foram estruturadas de forma a mensurar a importância atribuída pelo consumidor a diferentes marcas, conforme segue descrito na Tabela 1.

Tabela 1 - Layout do questionário aplicado ao estudo sensorial dos consumidores de café da cidade de Varginha, MG.

Natureza dos Grupos	Ref.	Questões	Descrição das Categorias
Escolaridade e Renda	A1	Escolaridade	Analfabeto ou EF incompleto: 1; EF completo ou EM incompleto: 2; EM completo ou Sincompleto: 3; S completo: 4;
	A2	Renda Familiar	< R\$ 1.000,00: 1; [R\$ 1.000,00; R\$ 2.000,00[: 2; [R\$ 2.000,00; R\$ 4.000,00[: 3; [R\$ 4.000,00; R\$ 6.000,00[: 4; ≥ R\$ 6.000,00: 5;
Quantidade de xícaras de café consumidas	B1	Quantidade de xícaras de café consumida em casa.	1; 2; 3; 4; mais de 4;
	B2	Quantidade de xícaras de café consumida no trabalho.	1; 2; 3; 4; mais de 4;
Marca e Qualidade	C1	Qual a importância você atribui a marca ao escolher o café?	Nenhuma: 1; Pouca: 2; Moderada: 3; Acentuada: 4; Muito Acentuada: 5;
	C2	A marca do café interfere na qualidade do café.	1; 2; 3; 4; 5;
Torrefadora e Produtor	D1	A qualidade do café depende mais da firma torrefadora e não do produtor rural.	1; 2; 3; 4; 5;
	D2	O produtor de café interfere diretamente na qualidade do café.	1; 2; 3; 4; 5;
Região e Produção	E1	A região onde é produzido o café interfere na qualidade.	1; 2; 3; 4; 5;
	E2	Quando compro café, procuro informações sobre a região de produção.	1; 2; 3; 4; 5;
Preferência e Preço	F1	Prefiro cafés com a identificação da origem, mesmo que sejam mais caros.	1; 2; 3; 4; 5;
	F2	Prefiro cafés de qualidade superior, mesmo que sejam mais caros.	1; 2; 3; 4; 5;
	F3	Preço maior significa que a qualidade do café é superior.	1; 2; 3; 4; mais de 4;

Escolaridade - EF: Ensino Fundamental (1º grau); EM: Ensino Médio (2º grau); S: Superior
Fonte: Pereira (2014)

Com os dados obtidos, procedeu-se à construção das tabelas para cada grupo, as quais foram organizadas em uma estrutura conveniente de serem extraídos os múltiplos fatores, denominada de tabelas justapostas (Tabela 2).

Tabela 2 - Layout da tabela justaposta para análise de cada grupo de variáveis.

Indivíduos	Escolaridade e renda		Quantidade de xícaras de café consumidas		Marca e qualidade		Torrefadora e produtor		Região e produção		Preferência e preço		
	A1	A2	B1	B2	C1	C2	D1	D2	E1	E2	F1	F2	F3
1	x_{ik1}		x_{ik2}		x_{ik3}		x_{ik4}		x_{ik5}		x_{ik6}		
2													
:													
...													
50													

x_{ikG} = observação i na variável k do grupo G .

4 Resultados

Com base nas especificações matriciais estabelecidas no referencial teórico, a obtenção dos resultados foi feita por meio do desenvolvimento de scripts no software R (R CORE TEAM 2020) para uso público, por meio do pacote MVar versão 2.1.3 (OSSANI; CIRILLO, 2020).

Os resultados descritos na Tabela 3 evidenciam que o estudo das associações entre as variáveis representadas em cada subtabela, analisadas simultaneamente, poderá ser explicada em um gráfico bidimensional, uma vez que λ_i a proporção da variação explicada nos dois primeiros componentes é igual a 17,14%. Entretanto, convém ressaltar que, embora esse percentual seja um número pouco expressivo, a qualidade de ajuste da técnica MFA prioriza a similaridade entre os grupos.

Tabela 3 – Proporção da variação amostral explicada pelos dois primeiros componentes.

Componentes	λ_i	% da variação amostral	% da variação amostral acumulada
1	2,6854	8,82	8,82
2	2,5327	8,32	17,14

As coordenadas obtidas que correspondem à posição global, isto é, considerando as respostas dos indivíduos a todos os subconjuntos de variáveis, estão representadas no mapa perceptual ilustrado na Figura 2, gerado pelos escores obtidos da equação (6), a qual permite interpretar, em termos práticos, que, em relação às questões abordadas em diferentes subconjuntos, não é possível descrever um perfil que caracteriza os consumidores. Tal afirmação é perceptível devido à distribuição dos escores, que estão localizados próximos ao centroide.

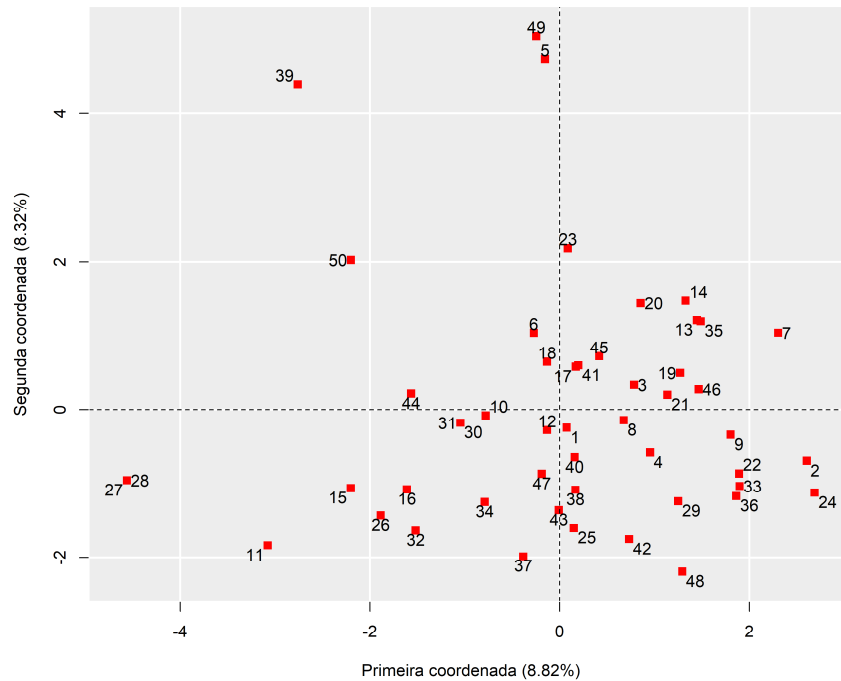


Figura 2 - Gráfico correspondente às respostas dos consumidores a todos os subconjuntos de variáveis na análise global.

Isso leva a crer que, dentre os fatores apontados por Malhotra (2001) que explicam o comportamento do consumidor, reportando aos resultados obtidos, há evidências para afirmar que a distribuição dos escores em torno da média dar-se-á pelo fato do grupo de consumidores analisado ser homogêneo, o que não justifica uma análise por renda ou escolaridade como efeito que caracterize um perfil diferenciado em relação ao comportamento do consumidor, mediante a aquisição do produto.

Em se tratando da influência que cada grupo de variáveis exerceu em relação às respostas dos consumidores (Figura 2), os resultados ilustrados na Figura 3, gerada pelos escores obtidos da equação (9), indicam que as linhas que unem os grupos com os escores provenientes da análise global foram inconclusivas ao destacarem algum conjunto de variáveis relevantes a respostas dos consumidores, uma vez que eles se encontram aglomerados próximo ao centroide, não revelando nenhum grupo de consumidor distinto, em relação à amostra.

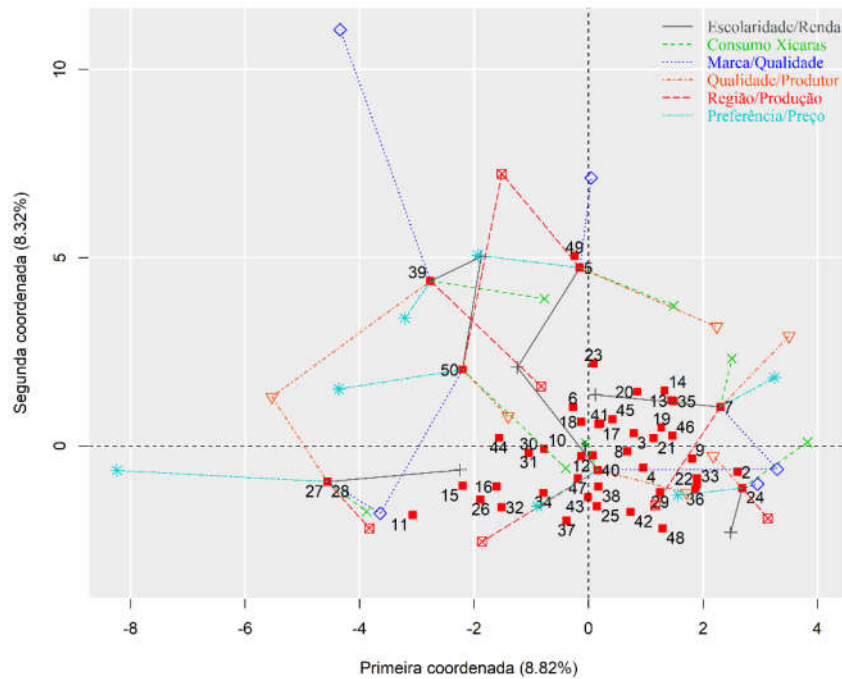


Figura 3 – Gráfico correspondente à influência que cada grupo de variáveis exerceu nas respostas dos consumidores na análise global.

Nota-se, pela magnitude dos coeficientes descritos na Tabela 4, gerada pelos escores obtidos da equação (10) e que descrevem a similaridade entre os grupos, os subconjuntos de variáveis (marca e qualidade) e (preferência e preço) apresentaram maior contribuição para a formação dos componentes.

Tabela 4 - Valores referentes às inércias computadas nos dois primeiros componentes para cada grupo de variáveis.

Comp.	Grupo de variáveis						Inércia total
	Escolaridade e renda	Quantidade de xícaras de café consumidas	Marca e qualidade	Torrefadora e produtor	Região e produção	Preferência e preço	
1	0,2493	0,4917	0,6076	0,3681	0,3119	0,6569	2,6854
2	0,4560	0,4133	0,4972	0,1617	0,4367	0,5677	2,5327

A partir das inércias obtidas em cada grupo na Tabela 4, visando uma melhor interpretação, o gráfico das inércias (Figura 4) é gerado.

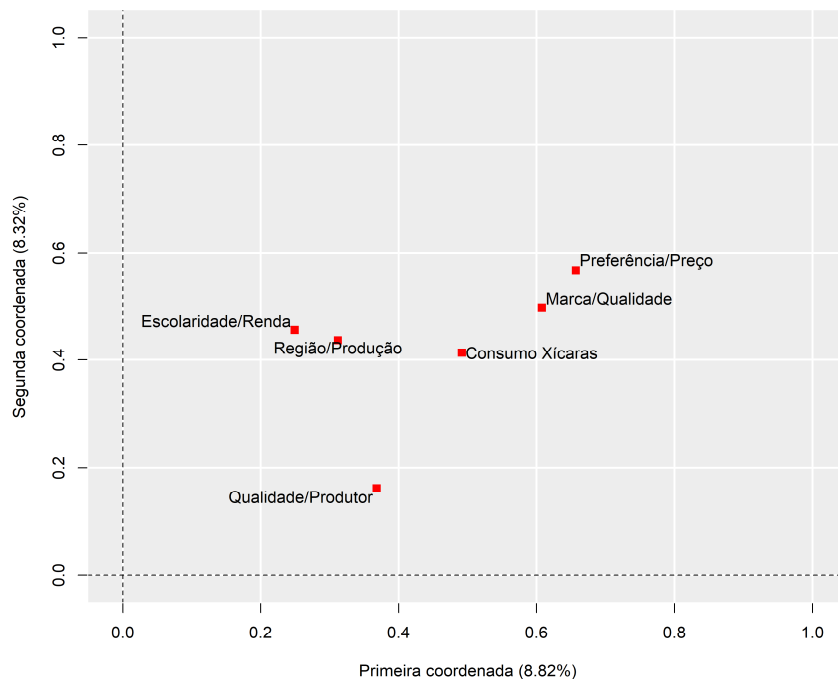


Figura 4 – Inércias dos grupos.

Convém ressaltar que, embora essas dimensões tenham apresentado baixas porcentagens de explicação da variabilidade amostral, os subconjuntos de variáveis (marca e qualidade) e (preferência e preço) foram mais marcantes, apresentando uma similaridade moderada em relação à primeira componente principal, podendo ser visualizado nas projeções dos eixos na Figura 4.

Devido à magnitude dos coeficientes observa-se que os subconjuntos (quantidade de xícaras de café consumidas), (torrefadora e produtor) e (região e produção) apresentaram baixa similaridade entre si.

O subconjunto (escolaridade e renda) apresentou pouquíssima similaridade com os outros subconjuntos. Como a técnica MFA trabalha com a similaridade entre os subconjuntos, verifica-se que esse último subconjunto sofre pouca influência das variáveis de outros subconjuntos na formação da análise final e, assim, o subconjunto (escolaridade e renda), em termos de grupo de variáveis, não se assemelha aos outros subconjuntos.

Em relação à segunda componente principal, pode-se dizer que todos os subconjuntos apresentaram baixa similaridade entre si, exceto o subconjunto (torrefadora e produtor), que apresentou similaridade baixíssima. Mas, como esses subconjuntos já foram explicados na primeira componente de maior explicação, como no PCA, pode-se considerar apenas a primeira componente para explicar a similaridade dos dados.

Porém, é sabido que esses grupos de variáveis são voláteis no tempo e dependentes da conjectura econômica e ou regional. Além desses fatores, as conclusões são limitadas ao tamanho amostral. Entretanto, por se tratar de uma pesquisa empírica, os resultados obtidos são apreciados no contexto de direcionar estratégias metodológicas de marketing e segmentação de mercado.

5 Conclusões

Estando de acordo com os objetivos propostos, a técnica MFA mostrou-se promissora ao analisar as similaridades de grupos de variáveis qualitativas ordinais.

Em relação aos resultados da aplicação, conclui-se que:

- i) as respostas dadas pelos indivíduos nos subconjuntos de variáveis (marca e qualidade) e (preferência e preço) caracterizam-se por uma similaridade moderada;
- ii) os subconjuntos (quantidade de xícaras de café consumidas), (torrefadora e produtor) e (região e produção) apresentaram baixa similaridade entre si.

Esses resultados mostram a viabilidade da técnica MFA na identificação do comportamento de subconjuntos de consumidores com diferentes características qualitativas ordinais, o que favorece a sua utilização em pesquisas que envolvem questionários com variáveis categóricas.

A técnica MFA ao estabelecer as similaridades entre os grupos de variáveis possibilitou que pesquisas futuras sejam simplificadas, assim ao direcionar grupos específicos de variáveis, já se sabe qual a relação com os outros grupos, deste modo ao pesquisar marca e qualidade de café, já se sabe que há uma similaridade moderada com preferência e preço por parte dos consumidores, isto é informação importante para aqueles que desejam direcionar o marketing ou a venda ao público consumidor de cafés.

Referências

- [1] ABDI, H.; WILLIAMS, L. Principal component analysis. **WIREs Computational Statistics**, New York, v. 2, n. 4, p. 433-459, July/Aug. 2010.
- [2] ABDI, H.; WILLIAMS, L.; VALENTIN, D. Multiple factor analysis: principal component analysis for multitable and multiblock data sets. **WIREs Computational Statistics**, New York, v. 5, n. 2, p. 149-179, Feb. 2013.
- [3] ARRUDA, A. C. *et al.* Justificativas e motivações do consumo e não consumo de café. **Ciência e Tecnologia de Alimentos**, Campinas, v. 29, n.4, p. 754-763, out.-dez. 2009.
- [4] ALAN, A. K., KABADAYI, E. T.; YILMAZ, C. Cognitive and affective constituents of the consumption experience in retail service settings: effects on store loyalty. **Serv Bus**. Vol. 10. p. 715-735. 2016. <https://doi.org/10.1007/s11628-015-0288-8>
- [5] ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE CAFÉ - ABIC. **Pesquisa tendências de consumo**. Tendências no consumo do café - 2014. Disponível em: <<https://www.abic.com.br/estatisticas/pesquisas-2/pesquisa-tendencias-de-consumo/>> Acesso em: dezembro 2020.
- [6] ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE CAFÉ - ABIC. **Estatísticas – indicadores da indústria de café - 2020**. Disponível em: <<https://www.abic.com.br/estatisticas/indicadores-da-industria/indicadores-da-industria-de-cafe-2018-2/>>. Acesso em: dezembro 2020.

- [7] ASSOCIAÇÃO BRASILEIRA DA INDÚSTRIA DE CAFÉ - ABIC. **Pesquisa Nacional Fiesp/IBOPE sobre o Perfil do Consumo de Alimentos no Brasil**. Disponível em: <http://www.abic.com.br/media/EST_PESQFoodTrendsI.pdf>. Acesso em: dezembro 2014.
- [8] BÉCUE-BERTAUT, M.; PAGÈS, J. Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. **Computational Statistics & Data Analysis**, New York, v. 52, n. 6, p. 3255-3268, Feb. 2008.
- [9] BOAS, L. H. *et al.* Marcas regionais de café em Uberlândia: a percepção de valor dos consumidores e a relação com o “ser regional”. in: 47º Congresso Sociedade Brasileira de Economia, Administração e Sociologia Rural (SOBER), 47., 2009, Porto Alegre. **Anais...** Porto Alegre: UFRGS, 2009.
- [10] EMBRAPA, Empresa brasileira de pesquisa agropecuária – **Ministério da agricultura, pecuária e abastecimento**. Consumo interno dos Cafés do Brasil representa 13% da demanda mundial. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/41277124/consumo-interno-dos-cafes-do-brasil-representa-13-da-demanda-mundial>>. Acessado: dezembro 2020.
- [11] ESCOPIER, B.; PAGÈS, J. **Analyses factorielles simples et multiples: objectifs, méthodes et interprétation**. 4th. ed. Paris: Dunod, 2008. 318 p.
- [12] FEHR, L. C. F. de A.; DUARTE, S. L.; TAVARES, M.; REIS, E. A. dos. Análise temporal das variáveis de custos da cultura do café arábica nas principais regiões produtoras do Brasil. **Custos e Agronegócio Online**, v. 8, n. 1 - Jan/Mar - 2012.
- [13] MALHOTRA, N. K. **Pesquisa de Marketing: uma orientação aplicada**. 3. ed. Porto Alegre: Bookman, 2001. p. 720.
- [14] MOORI, R. G.; BIDO, D. S.; OLIVEIRA, L. H. de. **Variáveis Relevantes do Consumidor do Café Solúvel sob o Enfoque da Diferenciação. Organizações Rurais e Agroindustriais (UFLA)**, v. 13, p. 124-138, 2011.
- [15] <https://www.embrapa.br/busca-de-noticias/-/noticia/41277124/consumo-interno-dos-cafes-do-brasil-representa-13-da-demanda-mundial>. Acessado: dezembro 2020.
- [16] NORONHA, R. L. F.; DELIZA, R.; SILVA, M. A. A. P. A expectativa do consumidor e seus feitos na avaliação sensorial e aceitação de produtos alimentícios. **Alimentos e Nutrição**, Araraquara, v.16, n.3, p. 299-308, jul./set. 2005.
- [17] OSSANI, P. C.; CIRILLO, M. A.; **MVar: Multivariate Analysis**. 2020. URL <<https://cran.r-project.org/web/packages/MVar/index.html>>. R package version 2.1.3.
- [18] PEREIRA, G. A. **Estimadores de Ridge Generalizados adaptados em modelos de equações estruturais: Estudo de simulação e aplicação no perfil de consumidores de café**. 2014. 80 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2014.

- [19] R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: Vienna University of Economics and Business, 2020. Available at: <[http:// www.R-project.org/](http://www.R-project.org/)>. Accessed at: december 2020.
- [20] SAES, M. S. M.; SPERS, E. E. Percepção do consumidor sobre os atributos de diferenciação no segmento rural: café no mercado interno. **Organizações Rurais & Agroindustriais**. Lavras, v.8, n.3, p.354-367, nov. 2006.
- [21] SPERS, E. E.; SAES, M. S. M.; SOUZA, M. C. M. Análise das preferências do consumidor brasileiro de café: um estudo exploratório dos mercados de São Paulo e Belo Horizonte. **RAUSP - Revista de Administração da Universidade de São Paulo**, São Paulo, v. 39, n. 1, p. 53-61, 2004
- [22] WANDERLEY, A. B. *et al.* Functional Benefits of Green Coffee in Metabolic Syndrome Prevention: A Review Study. **Journal of Chemical and Pharmaceutical Research**, v. 9, n.7, p. 5-12, 2017.