



De 11 a 14 - NOVEMBRO DE 2024

**Inteligência Artificial
na Gestão de Operações:**
Limitações e possibilidades



APLICAÇÃO DA ANÁLISE DE CONGLOMERADOS PARA AVALIAÇÃO DE PREÇOS DE COMBUSTÍVEIS AUTOMOTIVOS COMERCIALIZADOS EM GOIÁS

AGENOR SOUSA SANTOS NETO – agenor.neto@ueap.edu.br
UNIVERSIDADE DO ESTADO DO AMAPÁ - UEAP

JÚLIO CESAR VALANDRO SOARES – julio_cesar_valandro@ufg.br
UNIVERSIDADE FEDERAL DE GOIÁS - UFG

ÁREA: 7. ENGENHARIA ECONÔMICA

SUBÁREA: 7.2 – GESTÃO DE CUSTOS

RESUMO: A DEMANDA DO CONSUMIDOR POR PRODUTOS DE BAIXO CUSTO TORNOU-SE UMA CONSTANTE NO MERCADO BRASILEIRO, SOBRETUDO PARA COMBUSTÍVEIS AUTOMOTIVOS QUE TÊM UM IMPACTO FINANCEIRO RELEVANTE NA VIDA DAS PESSOAS E NA ECONOMIA DO PAÍS. LOGO, O OBJETIVO DESTES ESTUDO É APRESENTAR INDICADORES ECONÔMICOS, CONSTRUÍDOS A PARTIR DE ANÁLISE DE CONGLOMERADOS, E QUE PERMITAM IDENTIFICAR QUAIS BANDEIRAS DE POSTOS DE COMBUSTÍVEIS ESTÃO PRATICANDO O MENOR PREÇO GLOBAL. OS DADOS PARA A PESQUISA FORAM OBTIDOS DO PORTAL DE DADOS ABERTOS POR MEIO DA SÉRIE HISTÓRICA DE PREÇOS DE COMBUSTÍVEIS E DE GLP DA AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS (ANP). A ANÁLISE DE PREÇOS RECAIU SOBRE O ESTADO DE GOIÁS, TENDO COMO FOCO TRÊS PRODUTOS: GASOLINA COMUM, GASOLINA ADITIVADA E ETANOL. OITO BANDEIRAS DE POSTOS DE COMBUSTÍVEIS FORAM OBSERVADAS, AS QUAIS FORAM AGRUPADAS EM TRÊS CLUSTERS. RODOIL É A BANDEIRA QUE PRÁTICA OS MENORES PREÇOS GLOBAIS DE COMBUSTÍVEIS, ALESAT, CIAPETRO E RAIZEN SÃO AS BANDEIRAS QUE FORNECEM COMBUSTÍVEIS COM OS MAIORES PREÇOS GLOBAIS. O AUMENTO DOS PREÇOS DA GASOLINA E DO ETANOL, ALÉM DE OUTROS PRODUTOS, COMO GÁS DE COZINHA E DIESEL, GERA DEBATES ENTRE A POPULAÇÃO. EM RAZÃO DISSO, É IRREFUTÁVEL A NECESSIDADE DE O PAÍS CONSIDERAR UMA NOVA POLÍTICA DE PREÇOS PARA SEUS COMBUSTÍVEIS.

PALAVRAS-CHAVES: ANÁLISE DE CONGLOMERADOS; PREÇOS; COMBUSTÍVEIS; GOIÁS.

1. INTRODUÇÃO

De acordo com Mingoti e Silva (1997) a demanda do consumidor por produtos de baixo preço tornou-se uma constante no mercado brasileiro. O custo de bens e serviços, como alimentação, combustível, calçados, vestuário, saúde e educação tem grande impacto nos orçamentos familiares, necessitando de estudos de preços pelo consumidor em diversas instituições. Um dos problemas que os consumidores enfrentam é que é difícil saber onde comprar pelo menor preço global uma vez que existem diversas marcas ou especificações para um produto.

No mundo dos derivados de petróleo, o governo brasileiro, por meio do seu controle acionário na Petrobras, influenciou preços que eram livres desde o início dos anos 2000. Como a Petrobras detém quase a totalidade dos derivados de refino e importação, sua estratégia acaba determinando os preços em todo o mercado nacional. Além disso, o governo federal também reduziu os impostos sobre a gasolina e o diesel para efetuar reajustes de preços nas refinarias sem repassá-los aos consumidores finais (ALMEIDA; OLIVEIRA; LOSEKANN, 2015).

Ainda segundo Almeida, Oliveira e Losekann (2015) as políticas de preços de combustíveis orientadas para o controle da inflação tiveram um impacto econômico significativo na Petrobras. Esses efeitos econômicos negativos, além de reduzirem a atratividade do investimento privado no setor downstream da indústria petrolífera brasileira, também prejudicam a capacidade de investimento da empresa. Nesse sentido, fica muito clara a necessidade de o país considerar uma nova política de preços para seus combustíveis.

O aumento dos preços do gás de cozinha, da gasolina, do etanol e do diesel é um tema muito debatido entre a população. Estes produtos têm um impacto relevante na vida das pessoas e na economia. Desde outubro de 2016, os brasileiros convivem com instabilidade e escalada nos preços desses produtos devido à decisão da administração da Petrobras de monitorar as alterações no preço internacional do barril de petróleo, bem como as alterações cambiais nos preços do petróleo. Essa realidade é vivenciada apenas por países que não possuem petróleo e refinarias capazes de abastecer sua população e por isso têm que importar esses produtos, mas este não é o caso do Brasil (DIEESE, 2021).

O DIEESE (2021) destaca também que de 2016 a 2020, a produção nacional de petróleo aumentou 18%, atingindo 3,7 milhões de barris equivalentes/dia (BOE/dia). O rápido crescimento da produção dos campos do pré-sal descobertos em 2006 e desenvolvidos nos

anos subsequentes foi o grande responsável por isso, com o país exportando cerca de 1 milhão de barris por dia em 2020. Na estrutura mundial, o Brasil ocupa uma posição dominante. Em 2019, foi o décimo maior produtor mundial, com o nono maior complexo refinador com capacidade de produção de 2,4 milhões de barris por dia, e o nono maior consumidor de petróleo do mundo.

Nestas condições, como explicar a contínua subida dos preços dos derivados nas refinarias ou nos postos de gasolina? Por que os brasileiros gastam tanto dinheiro com gasolina, etanol e diesel? Logo, este estudo tem como objetivo apresentar indicadores econômicos, construídos a partir de Análise de Conglomerados, e que permitam identificar quais bandeiras de postos de combustíveis estão praticando o menor preço global em uma determinada época de interesse.

2. REFERENCIAL TEÓRICO

2.1 Mercado de gasolina e etanol no Brasil

O mercado de gasolina no Brasil está ligado ao mercado de etanol. Desde a década de 1970, com a primeira crise do petróleo, o Proálcool estimulou a produção de etanol com o objetivo de substituir a gasolina e produzir automóveis movidos a etanol (o chamado etanol hidratado). Além disso, a gasolina consumida pelos utilizadores de veículos movidos a combustíveis fósseis é misturada com etanol, denominado etanol anidro, não só para estabilizar os preços nacionais da gasolina, mas também para reduzir as emissões de poluição (MELO; SAMPAIO, 2014).

Portanto, até 2003, a elasticidade de substituição entre a demanda de etanol e os preços da gasolina era baixa, e os consumidores que trocaram de combustível tiveram que trocar também de veículo para consumir um combustível mais barato. Porém, a partir de maio de 2003, uma nova tecnologia foi implementada em veículos que permitia aos consumidores abastecer seus veículos com qualquer mistura de etanol hidratado e gasolina. Os veículos equipados com essa tecnologia são chamados de *flex-fuel*. Isso permite que os consumidores escolham o combustível mais barato (TOKGOZ; ELOBEID, 2006).

A inter-relação entre os mercados de etanol e gasolina no Brasil é única em comparação com outros países, onde o etanol e a gasolina são bens complementares porque os consumidores compram combustível em que o etanol anidro e gasolina já se encontram misturados. No Brasil, de acordo com Balcombe e Rapsomanikis (2008), a nova composição

da frota resultou em uma relação complexa entre gasolina e etanol, uma vez que podem ser combustíveis complementares ou substitutos. Segundo os autores, para os consumidores que possuem veículos movidos inteiramente por uma mistura de etanol anidro e gasolina, um aumento nos preços da gasolina levaria a uma diminuição na demanda por etanol anidro consumido junto a gasolina. Para os consumidores que possuem veículos *flex*, um aumento no preço da gasolina faz com que a curva de demanda mais elástica do etanol se desloque para a direita. Logo, os preços da gasolina afetam a demanda por etanol anidro e a demanda por etanol hidratado.

Contudo, os consumidores de veículos *flex-fuel* não olham apenas para os preços da gasolina e do etanol quando trocam de combustível, mas sim para os preços relativos entre os dois combustíveis. Esse comportamento é razoável porque o desempenho dos carros movidos a etanol hidratado é 30% inferior ao dos carros movidos a gasolina. Ou seja, se a relação preço do álcool em relação à gasolina for superior a 0,7, os consumidores optarão por adicionar gasolina aos seus carros. Se a relação for inferior a 0,7, os consumidores optarão por adicionar etanol aos seus carros (LOSEKANN; CASTRO, 2011).

Ferreira, Prado e Silveira (2009) constataram que o preço dos combustíveis dependerá da proporção de álcool, gasolina e veículos *flex* no estoque total. A demanda por determinado combustível dependerá dos preços esperados do álcool e da gasolina. No longo prazo, existe uma forte relação causal entre os preços da gasolina e do etanol. Losekann e Castro (2011) por sua vez apontam que o preço relativo dos combustíveis é o principal critério de escolha do consumidor.

2.2 Análise de conglomerados

De acordo com Guimarães, Asmus e Burdorf (2013) a análise de conglomerados ou de *cluster*, visa resolver o seguinte problema: dada uma amostra de n indivíduos, cada um caracterizado por p variáveis, deve-se criar um critério para classificar os indivíduos em classes ou *clusters* de modo que indivíduos com características semelhantes pertencem à mesma classe. Portanto, o objetivo da análise de conglomerados é atribuir indivíduos a grupos mutuamente exclusivos, ou seja, os indivíduos pertencem apenas a um grupo, de modo que os elementos de um mesmo grupo sejam tão semelhantes quanto possível, enquanto os indivíduos de grupos diferentes sejam heterogêneos.

A análise de conglomerados é uma técnica de avaliação ou análise de interdependência que visa agrupar elementos de acordo com sua estrutura natural (HAIR et al., 2005). Em

termos gerais, Abram e Treinish (1999) afirmam que a análise de *cluster* é bem-sucedida se fornecer informações que melhorem agrupamentos previamente identificados em um conjunto de dados ou ajude a formalizar sua estrutura hierárquica. Johnson e Wichern (2007) acrescentam que esta é uma técnica exploratória importante porque pode reduzir a dimensionalidade dos dados e identificar valores atípicos (*outliers*) e hipóteses relevantes para os indivíduos ao estudar a estrutura natural da associação das populações.

A análise de conglomerados visa classificar indivíduos (ou variáveis) em grupos internamente homogêneos, mutuamente heterogêneos e mutuamente exclusivos (não mais do que um grupo com a mesma particularidade) com base em alguma característica de similaridade, que mede o grau de semelhança entre dois indivíduos ou dissimilaridade que mede o quão diferentes esses dois indivíduos são (HAIR et al., 2005).

De modo geral, as semelhanças e diferenças entre os elementos da amostra são expressas pela sua distância no espaço variável medido. Uma distância comumente usada é a distância euclidiana. Resumidamente, o método de análise de *cluster* (ANDERBERG, 1973) pode ser descrito da seguinte forma: Seja $X = (X_1, X_2, \dots, X_p)$ o conjunto contendo as p variáveis medidas em cada um dos n elementos amostrais, aqui representados como: $(E_1, E_2, \dots, E_n)$. Com base no conjunto X os n elementos amostrais são agrupados em grupos g_i de modo que:

- a) se E_l e $E_k \in g_i \rightarrow E_l$ e E_k são semelhantes;
- b) se E_l e $E_k \notin g_i \rightarrow E_l$ e E_k são distintos.

3. METODOLOGIA

3.1 Coleta dos dados

Os dados para a pesquisa foram obtidos do Portal de Dados Abertos por meio da Série Histórica de Preços de Combustíveis e de GLP da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis – ANP (<https://dados.gov.br/dados/conjuntos-dados/serie-historica-de-precos-de-combustiveis-e-de-glp>).

Devido a diversidade de bandeiras de postos de combustíveis, dez, o estado eleito para análise de preços foi Goiás. As variáveis utilizadas para construção do modelo se referem aos três produtos observados: Gasolina Comum, Gasolina Aditivada e Etanol.

Os dados foram coletados pela ANP no período de 02 a 31 de janeiro de 2024, totalizando 2202 amostras.

3.2 Etapas da análise de conglomerados

Este estudo tem como objetivo fornecer indicadores econômicos construídos a partir de análises de conglomerados, a fim de identificar quais bandeiras de postos de gasolina oferecem os preços mais baixos globalmente durante um período específico de interesse. Para tanto, Hair et al. (2006) declara que existem cinco etapas necessárias para usar análise de conglomerados em um determinado projeto de pesquisa.

O primeiro passo é definir a amostra. Para Hair et al. (2006), o tamanho da amostra na análise de *cluster* é irrelevante para questões de inferência estatística na análise de regressão. Em outras palavras, nenhuma tentativa é feita para estimar até que ponto os resultados encontrados na amostra podem ser generalizados para a população. Na verdade, o tamanho da amostra deve garantir que pequenos grupos da população estejam adequadamente representados. Além disso, diferentemente de outras técnicas multivariadas, não existe uma regra geral que especifique um tamanho mínimo de amostra (DOLNICAR, 2002).

Após a seleção da amostra (primeira etapa), o pesquisador deve decidir quais variáveis utilizar para estimar a distância/semelhança entre os casos (segunda etapa). Como a análise de *cluster* não distingue entre variáveis relevantes e irrelevantes, esta inclusão deve ser orientada pela teoria. Hair et al. (2006) afirmam que devem ser incluídas apenas variáveis que caracterizem os objetos a serem agrupados e que estejam especificamente relacionadas aos objetivos da análise de conglomerados.

Após selecionar as variáveis utilizadas para estimar a similaridade entre os casos (etapa 2), o pesquisador deve definir a medida de similaridade a ser utilizada (etapa 3). Pohlmann (2007) declara que similaridade entre objetos é uma medida da correspondência ou semelhança entre objetos a serem agrupados. Existem diferentes maneiras de calcular essas medidas, e diferentes medidas tendem a produzir soluções diferentes.

Uma vez calculadas as similaridades, o próximo passo é decidir sobre o método de agrupamento (algoritmo matemático). Ou seja, o pesquisador deve definir como as distâncias são calculadas e quantos conglomerados (grupos) devem ser criados. O *Statistical Package for Social Sciences* (SPSS) fornece três métodos gerais para a criação de *clusters*: a) agrupamento hierárquico (*hierarchical clustering*); b) agrupamento *k-means* (*k-means clustering*) e c) agrupamento em duas etapas (*two-step clustering*). O método de agrupamento hierárquico é mais adequado para amostras pequenas, geralmente, $n < 250$. A opção de agrupamento *k-means* é mais adequada para amostras maiores ($n > 1.000$) porque não calcula uma matriz de proximidade de distância/similaridade entre todos os casos observados. Os

métodos de agrupamento em duas etapas são considerados ideais para grandes bancos de dados porque tanto o agrupamento hierárquico quanto o agrupamento *k-means* sofrem de problemas de escala quando a amostra é muito grande (GARSON, 2009).

Após selecionar a medida de similaridade e o método de agrupamento (algoritmo matemático) (etapa 3), o pesquisador deve determinar o número de grupos (*k*) que serão formados (etapa 4). Finalmente, na quinta etapa, os investigadores devem validar os resultados. Hair et al. (2005) alertam que os pesquisadores devem ter muito cuidado para verificar e garantir o significado prático da solução final. A validação envolve garantir que a solução encontrada é representativa da população, descrevendo um padrão relativamente estável em outras amostras.

4. RESULTADOS E DISCUSSÕES

Inicialmente tinha-se 2202 amostras referentes às dez bandeiras de postos de combustíveis de Goiás, sendo elas: i) Alesat, ii) Branca, iii) Ciapetro, iv) Ipiranga, v) Masut Distribuidora, vi) Raizen, vii) Rio Branco, viii) Rodoil, ix) Totalenergies e x) Vibra. Como critério de exclusão decidiu-se apenas pelas bandeiras que comercializassem, simultaneamente, os três produtos: gasolina comum, gasolina aditiva e etanol. Rio Branco e Totalenergies não ofertam gasolina aditivada, logo não participaram deste modelo. Sendo assim, após a exclusão, restaram para a análise 2196 amostras. Na Tabela 1 é mostrado o preço médio dos combustíveis para cada bandeira.

TABELA 1 – Preço médio dos combustíveis de cada bandeira.

Bandeira	Preço Médio Gasolina Comum	Preço Médio Gasolina Aditivada	Preço Médio Etanol
ALESAT	R\$ 5,76	R\$ 5,93	R\$ 3,75
BRANCA	R\$ 5,59	R\$ 5,70	R\$ 3,54
CIAPETRO	R\$ 5,75	R\$ 5,77	R\$ 3,75
IPIRANGA	R\$ 5,61	R\$ 5,77	R\$ 3,56
MASUT DISTRIBUIDORA	R\$ 5,59	R\$ 5,62	R\$ 3,54
RAIZEN	R\$ 5,70	R\$ 5,98	R\$ 3,66
RODOIL	R\$ 5,39	R\$ 5,39	R\$ 3,19
VIBRA	R\$ 5,64	R\$ 5,80	R\$ 3,60

Pela Tabela 1 é percebido que os menores preços globais são praticados pela Rodoil e os maiores preços globais são praticados pela Alesat. Após isso, é aplicada a análise de conglomerados a fim de agrupar as bandeiras dos postos por preço global. No primeiro

momento, o pesquisador pode optar por não estabelecer um número determinado de agrupamentos ou intervalo de soluções. Entretanto, para este modelo vamos supor que o ideal de agrupamentos estaria entre 2 e 4 grupos. Na Tabela 2 é mostrada a matriz de proximidade entre as variáveis.

TABELA 2 – Matriz de proximidade pela distância quadrática euclidiana.

Caso	Distância Euclidiana Quadrática							
	1:ALESAT	2:BRANCA	3:CIAPETRO	4:IPIRANGA	5:MASUT DISTRIBUIDORA	6:RAIZEN	7:RODOIL	8:VIBRA
1:ALESAT	0	5,04	0,76	3,52	6,31	0,59	28,41	2,25
2:BRANCA	5,04	0	3,39	0,19	0,19	3,64	9,6	0,59
3:CIAPETRO	0,76	3,39	0	2,56	3,9	1,74	23,54	1,61
4:IPIRANGA	3,52	0,19	2,56	0	0,7	2,2	12,09	0,14
5:MASUT DISTRIBUIDORA	6,31	0,19	3,9	0,7	0	5,15	8,33	1,25
6:RAIZEN	0,59	3,64	1,74	2,2	5,15	0	24,19	1,33
7:RODOIL	28,41	9,6	23,54	12,09	8,33	24,19	0	14,79
8:VIBRA	2,25	0,59	1,61	0,14	1,25	1,33	14,79	0

Numa matriz de proximidade, os valores representam a semelhança ou dissimilaridade entre cada par de observações. Neste caso, a matriz de proximidade é a matriz de dissimilaridade porque a medida de distância (distância quadrática euclidiana) é utilizada como medida de similaridade. Quanto menor a distância, mais próximos os elementos estão, ou seja, mais semelhantes eles são. De acordo com a Tabela 2, verifica-se que os elementos agrupados primeiro pela medida quadrática de distância serão os elementos com menores coeficientes. Neste caso específico, trata-se das bandeiras Ipiranga (4) e Vibra (8), com coeficiente de 0,14. Na Tabela 3 é mostrado o esquema de aglomeração que indica a ordem de agregação (estágios) das bandeiras nos respectivos *clusters*.

TABELA 3 – Esquema de aglomeração.

Estágio	Cluster combinado		Coeficientes	O cluster de estágio é exibido primeiro		Próximo estágio
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	4	8	0,14	0	0	4
2	2	5	0,19	0	0	4
3	1	6	0,59	0	0	5
4	2	4	0,68	2	1	6
5	1	3	1,25	3	0	6
6	1	2	3,41	5	4	7
7	1	7	17,28	6	0	0

Na Tabela 3, a coluna “Estágio” mostra em que etapa cada elemento foi agrupado. As duas próximas colunas “Cluster combinado” indicam quais observações ou grupos foram unidos em cada estágio. A coluna “Coeficientes” mostra a distância entre dois grupos unidos em cada estágio. Neste caso, utilizamos a distância quadrática euclidiana para agrupar os elementos de acordo com a menor distância média entre todos os pares de indivíduos dos dois grupos. No estágio 1, agrupou-se o registro 4 e 8, que são, respectivamente, as bandeiras Ipiranga e Vibra, que apresentaram a menor distância média na matriz de proximidade, sendo apresentada na coluna “Coeficientes” o valor de 0,14. Nesse momento, tem-se 7 *clusters*. No estágio 2, a bandeira 2 (Branca) foi agrupada à bandeira 5 (Masut Distribuidora). O total de *clusters* no estágio 2 é 6. No estágio 3, a bandeira 1 (Alesat) foi agrupada à bandeira 6 (Raizen), e assim, sucessivamente. As colunas “O cluster de estágio é exibido primeiro” indicam em qual estágio anterior cada elemento foi associado a outro. A coluna “Próximo estágio” indica quando o *cluster* formado naquele estágio vai se juntar com outro. Na Tabela 4 é apresentado o número de *clusters* que cada empresa pertence de acordo com os grupos.

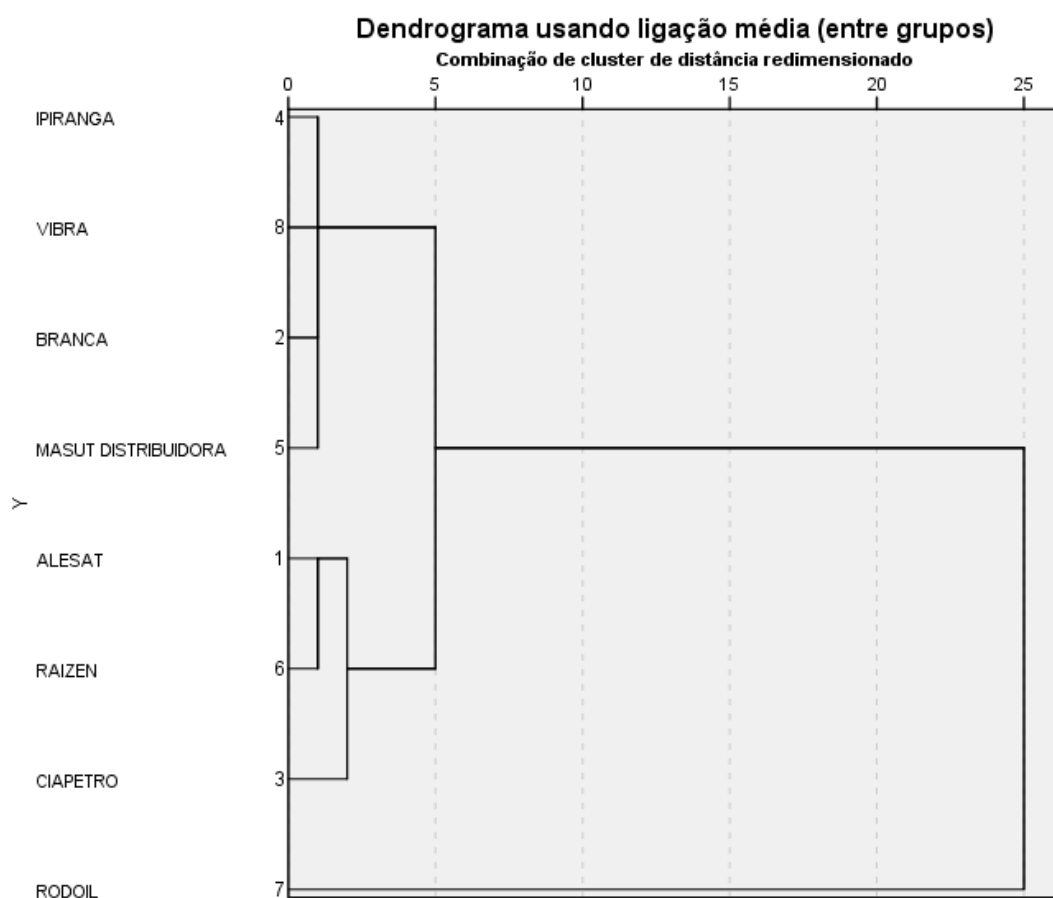
TABELA 4 – Associação do cluster.

Caso	4 Clusters	3 Clusters	2 Clusters
1:ALESAT	1	1	1
2:BRANCA	2	2	1
3:CIAPETRO	3	1	1
4:IPIRANGA	2	2	1
5:MASUT DISTRIBUIDORA	2	2	1
6:RAIZEN	1	1	1
7:RODOIL	4	3	2
8:VIBRA	2	2	1

Neste caso, como escolhemos uma solução entre 2 e 4 grupos, cada coluna indica o número de grupos formados e as bandeiras que pertencem a cada grupo. Se a solução com 2 *clusters* for selecionada, ela será composta das seguintes bandeiras: *Cluster 1* (1, 2, 3, 4, 5, 6 e 8) e *Cluster 2* (7), colocando a Rodoil em um grupo isolado. Por outro lado, se a solução com 3 *clusters* for selecionada, ela será composta das seguintes bandeiras: *Cluster 1* (1, 3 e 6), *Cluster 2* (2, 4, 5 e 8) e *Cluster 3* (7), colocando a Rodoil em um grupo isolado novamente. Por fim, caso a solução com 4 *clusters* for eleita, ela será formada pelas seguintes bandeiras: *Cluster 1* (1 e 6), *Cluster 2* (2, 4, 5 e 8), *Cluster 3* (3) e *Cluster 4* (7), colocando agora, a Ciapetro e Rodoil em grupos isolados.

Analisando os grupos, tem-se que a solução com 2 *clusters* traz um agrupamento (*Cluster 1*) composto por bandeiras (1, 2, 3, 4, 5, 6 e 8) com preços mais elevados de combustíveis e outro agrupamento (*Cluster 2*) composto por uma única bandeira (7), Rodoil, com o menor preço de combustíveis se comparado às demais bandeiras. Já a solução com 3 *clusters* traz um agrupamento (*Cluster 1*) composto por bandeiras (1, 3 e 6) com os preços mais elevados de combustíveis, outro agrupamento (*Cluster 2*) composto por bandeiras (2, 4, 5 e 8) com uma faixa intermediária de preços de combustíveis, e o último agrupamento, *Cluster 3*, representa a bandeira com o menor preço de combustíveis, Rodoil. Por fim, a solução com 4 *clusters* traz um agrupamento (*Cluster 1*) composto por bandeiras (1 e 6) com os preços mais elevados de combustíveis, outro agrupamento (*Cluster 2*) composto por bandeiras (2, 4, 5 e 8) com uma faixa intermediária de preços de combustíveis, tem-se o agrupamento (*Cluster 3*) incluindo uma única bandeira, Ciapetro, com uma faixa de preço moderadamente inferior às bandeiras do *Cluster 1*, e o último agrupamento, *Cluster 4*, representa a bandeira com o menor preço de combustíveis, Rodoil. Fazendo uma avaliação geral, constata-se que o agrupamento com três *clusters* é o que melhor atende a realidade do modelo, pois temos clareza entre as bandeiras que praticam o menor preço, preço intermediário e maior preço. A Figura 1 contém o dendograma que representa graficamente o esquema de aglomeração.

FIGURA 1 – Dendrograma do esquema de aglomeração.



Pela Figura 1 é percebido que o coeficiente (distância) é expresso no intervalo de 0 a 25. Além disso, por meio do dendrograma podemos visualizar os elementos de cada *cluster* com base nos cortes que se deseja fazer, bastando imaginar uma linha vertical no dendrograma. Neste caso, fica perceptível a existência de três grupos, o que é corroborado pela análise anterior, onde também foi percebida a necessidade de três grupos para alocação das oito bandeiras de postos de combustíveis analisadas em Goiás. Sendo assim, Rodoil é a bandeira que pratica os menores preços globais de combustíveis, Alesat, Ciapetro e Raizen são as bandeiras que fornecem combustíveis com os maiores preços globais, e na faixa intermediária de preços tem-se Branca, Ipiranga, Masut Distribuidora e Vibra.

5. CONSIDERAÇÕES FINAIS

Os resultados apresentados ilustram a importância prática da utilização de técnicas estatísticas multivariadas. Muitas vezes, são realizadas e publicadas pesquisas de preços em postos de combustíveis com o objetivo de apresentar à população uma escolha no momento

de abastecimento de seus veículos entre os diversos estabelecimentos. Entretanto, em geral, estas pesquisas não levam em consideração a utilização de técnicas estatísticas que permitem identificar mais globalmente a política de preços praticada pelos postos de combustíveis. A técnica de Análise de Conglomerados é, relativamente, de fácil utilização e pode levar a comparações com resultados mais conclusivos e precisos.

Logo, este estudo que tinha como objetivo apresentar indicadores econômicos, construídos a partir de Análise de Conglomerados para identificar quais bandeiras de postos de combustíveis estão praticando o menor preço global em uma determinada época de interesse alcançou seu êxito, demonstrando que a Rodoil é a bandeira que fornece aos seus consumidores de Goiás os menores preços globais para gasolina comum, gasolina aditivada e etanol. Três bandeiras (Alesat, Ciapetro e Raizen) demonstram ofertar produtos a um preço global mais elevado em relação aos seus concorrentes, enquanto quatro bandeiras (Branca, Ipiranga, Masut Distribuidora e Vibra) praticam um preço global intermediário, ou seja, situado entre a faixa superior e inferior de valores analisados.

É importante ressaltar que a análise de conglomerados permite aos pesquisadores agrupar casos e até mesmo variáveis em grupos homogêneos com base em semelhanças entre observações baseadas em variáveis previamente selecionadas. O objetivo desta técnica é buscar a estrutura “natural” dos dados agrupando objetos com base em suas próprias características. A escolha das variáveis define a estratificação obtida e, portanto, qualquer afirmação de que duas observações são semelhantes para um conjunto de variáveis não consideradas na análise de conglomerados é puramente especulativa. Como a análise de conglomerados é uma técnica exploratória, o simples fato de identificarmos que determinados objetos pertencem ao mesmo cluster não significa que sejam semelhantes em todos os aspectos. Os agrupamentos formados refletem a estrutura inerente às variáveis selecionadas.

Como sugestão para trabalhos futuros recomenda-se a comparar os resultados promovidos pela Análise de Conglomerados com outras técnicas de estatística multivariada, como a Análise de Componentes Principais, a fim de avaliar se ambas técnicas conduzem a conclusões semelhantes quanto a política de preços praticada pelos postos de combustíveis de Goiás ou de outro estado de interesse.

REFERÊNCIAS

ABRAM, G.; TREINISH, L. **An Extended Data-Flow Architecture for Data Analysis and Visualization**. New York: Computer Graphics, 1 ed., 1999.

ALMEIDA, E. L. F.; OLIVEIRA, P. V.; LOSEKANN, L. Impactos da contenção dos preços de combustíveis no Brasil e opções de mecanismos de precificação. **Revista de Economia Política**, v. 35, n. 3, p. 531-556, 2015.

ANDERBERG, M. R **Cluster analysis for applications**. Academic Press, 1973.

BALCOMBE, K.; RAPSOMANIKIS, G. Bayesian estimation and selection of nonlinear vector correction models: the case of the sugar-ethanol-oil nexus in Brazil. **American Journal of Agricultural Economics**, v. 90, n. 3, p. 658-668, 2008.

DIEESE – DEPARTAMENTO INTERSINDICAL DE ESTATÍSTICA E ESTUDOS SOCIOECONÔMICOS. **O necessário debate sobre os preços dos combustíveis no Brasil**. Nota Técnica, n. 251. 2021. Disponível em: <<https://www.dieese.org.br/notatecnica/2021/notaTec251PrecosCombustiveis.pdf>>. Acesso em: jun. 2024.

DOLNICAR, S. **A review of unquestioned standards in used cluster analysis for data-driven market segmentation**. Faculty of Commerce – Papers, 2002. Disponível em: <<https://ro.uow.edu.au/commpapers/273/>>. Acesso em: jun. 2024.

FERREIRA, A. L.; PRADO, F. P. A.; SILVEIRA, J. J. Flex cars and the alcohol price. **Energy Economics**, v. 31, n. 3, p. 382-394, 2009.

GARSON, G. D. **Statnotes: Topics in Multivariate Analysis**. 2009. Disponível em: <<http://faculty.chass.ncsu.edu/garson/PA765/statnote.htm>>. Acesso em: jun. 2024.

GUIMARÃES, R. M.; ASMUS, C. I. R. F.; BURDORF, A. Caracterização da exposição de população a organoclorados: uma aplicação da análise de cluster. **Revista Brasileira de Epidemiologia**, v. 16, n. 2, p. 231-239, 2013.

HAIR, J. F.; BLACK, W. C; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise Multivariada de Dados**. Porto Alegre: Bookman, 2005.

_____. **Multivariate Data Analysis**. 6 ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

JOHNSON, R.; WICHERN, D. **Applied multivariate statistical analysis**. Upper Saddle River: Pearson education, 6 ed., 2007.

LOSEKANN, L.; CASTRO, G. R. **Automóveis flex fuel: entendendo a escolha de combustível**. Infopetro, Etanol, 21 mar. 2011. Disponível em: <<http://infopetro.wordpress.com/2011/03/21/automoveis-flexfuel-entendendo-a-escolha-de-combustivel/>>. Acesso em: jun. 2024.

MELO, A. S.; SAMPAIO, Y. S. B. Impactos dos preços da gasolina e do etanol sobre a demanda de etanol no Brasil. **Revista de Economia Contemporânea**, v. 18, n. 1, p. 56-83, 2014.

MINGOTI, S. A.; SILVA, A. F. Um exemplo de aplicação de técnicas de estatística multivariada na construção de índices de preços. **Nova Economia**, Belo Horizonte, v. 7, n. 2, p. 203-212, 1997.

TOKGOZ, S.; ELOBEID, A. **An analysis of the link between ethanol, energy, and crop markets**. Working Paper, Iowa State University, n. 06-WP 435, 2006.