

Análise de Agrupamentos como Ferramenta de Inovação em Gestão da Saúde: Consumo de Álcool dos Discentes do IFBA/Campus Salvador

Jowâner de Oliveira Araújo
Maria Teresinha Tamanini

José Lamartine de Andrade Lima Neto
Antônio Clodoaldo Almeida Neto

Maria Aparecida da Silva Modesto
Rivailda Silveira Nunes de Argollo

RESUMO

O consumo de álcool é um problema complexo em vários países, sendo uma grave e crescente temática de saúde pública; no Brasil, esse aumento tem sido considerável, inclusive no ambiente educacional. Pesquisas apontam que, em 2006, 13% da população com menos de 15 anos tinham experimentado bebidas alcoólicas; em 2012, esse número passou para 22%. Gerenciar essa preocupante e ascendente tendência, no âmbito institucional, deve ser uma das prioridades dos gestores de saúde, públicos e privados, fazendo uso de informações relevantes. Resultados iniciais sobre o consumo de álcool no Instituto Federal da Bahia foram apresentados em eventos de pesquisa em 2018, com dados obtidos a partir da aplicação do questionário *Alcohol Use Disorders Identification Test* (AUDIT), utilizado como referência para identificação de problemas relacionados ao uso do álcool, ao evidenciar o correspondente grau de consumo pelo indivíduo, bem como desordens associadas e possível dependência. O presente trabalho utilizou a Análise de Agrupamentos para tratamento da informação e geração de conhecimento visando a assertividade da gestão, a partir de um extrato do banco de dados original, composto por uma amostra de 256 estudantes do IFBA/Campus Salvador. Objetivou verificar a associação entre o consumo de álcool e nove variáveis sociais demográficas, para identificar o segmento de consumo mais crítico, nesse universo, possibilitando ações gerenciais, inovadoras e mais efetivas, no tratamento desta complexa e preocupante questão. A análise de agrupamentos em conjunto com o teste de independência do consumo de álcool no ambiente educacional são ferramentas disponíveis a serem utilizadas como inovações na gestão da saúde.

Palavras-chaves: Análise de Agrupamentos; Inovação em Gestão da Saúde; Consumo de Álcool.

1 Introdução

Em pleno contexto de uma epidemia mundial, resultante de um inédito vírus, o tema da saúde ganha enorme relevância, nas suas perspectivas pública e privada. A gestão da saúde, em todas as suas nuances, ganha correspondente importância, requerendo abordagens disruptivas. Se por um lado a já tradicional Medicina Baseada em Evidências, com toda sua consistente metodologia estatística, brilha no cenário epidêmico mundial, com sua vertente de maior acuidade no trato com os medicamentos, por outro, ganha evidência novos paradigmas de como encarar a saúde e a medicina nestes tempos de um “novo normal”, com uma abordagem focada muito maior na prevenção e, secundariamente, a remediação. Dentre estes novos enfoques, vale a pena observar um recente ramo destas novas abordagens, que é a Medicina do Estilo de Vida (MEV), baseada no trabalho da Dra. Beth Pegg Frates, que criou um estilo de vida em forma de medicina, na *Harvard Extension School* em 2014, cujo sucesso a inspirou a fazer uma parceria com o *American College of Lifestyle Medicine* para construir um programa de estudos que pudesse ser amplamente distribuído entre profissionais e alunos de diversas áreas da saúde.

A MEV é uma abordagem baseada em evidências para prevenir, tratar e reverter doenças, substituindo comportamentos não saudáveis por positivos, como comer de forma saudável, ser fisicamente ativo, controlar o estresse, evitar abuso de substâncias tóxicas, sono adequado e um forte sistema de apoio através de estímulo da sociabilidade (FRATTES, B. *et*

al., DATA). Ressaltamos aqui, a importância nesta nova abordagem, do não uso de substâncias tóxicas e que causam dependência, como cigarro e álcool, decorrente de estilo de vida inadequado.

Segundo informe da Organização das Nações Unidas (ONU), muitos dos custos diretos e indiretos do problema das drogas (álcool inclusive) podem ser quantificados em termos monetários. Estudos econômicos mostraram que os custos flutuaram entre 0,07% e 1,7% do PIB dos países examinados. Mais difíceis de contabilizar são os custos intangíveis, como perda de vida ou deterioração da qualidade de vida, bem como os anos de vida degradados com uma deficiência adquirida (UNODC 2016).

Como se depreende, facilmente, a gestão da saúde tem que, também, sofrer *upgrades* em várias dimensões, inovando seus procedimentos, protocolos e padrões, bem como metodologias e ferramentas, técnicas e gerenciais, dentre essas as estatísticas. O adequado tratamento de dados propicia a geração de informações, constituindo-se de suma importância no mundo atual e um elemento básico para decisão e ações gerenciais. Sendo devidamente tratada contribui, significativamente, para a geração de conhecimentos, desenho de estratégias, resolução de problemas e busca de melhorias e inovações, com aumento da efetividade e agregação de valor, cada mais consequente, para todos os *stakeholders* integrantes dos diversos sistemas de saúde, privado, municipal, estadual e federal. Esse é o enfoque desse artigo: aplicar uma técnica estatística para aumentar a acuidade da gestão com futuras ações mais efetivas para a mitigação do uso de álcool e, consequente, favorecimento à saúde da população envolvida.

Vale enfatizar que o consumo de substâncias psicoativas é caracterizando como problema abrangente para vários países. No Brasil, o consumo de álcool tem aumentado, inclusive entre o público adolescente. No levantamento realizado por Laranjeira (2014), entre 2006 e 2012 aumentou a população que experimentou álcool com menos de 15 anos saindo de 13% para 22%.

Em pesquisa¹ realizada no Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA), em 2017, sobre consumo de álcool e drogas, relacionamentos, comportamentos de risco e estilo de vida, abrangendo um universo de 19.750 estudantes, obteve-se os primeiros resultados os quais foram divulgados em 2018, no Congresso do IFBA/*Campus* Salvador, XII Congresso Norte e Nordeste de Pesquisa e Inovação – XII CONNEPI e I Simpósio Brasileiro de Teoria e Ciência das Redes – SimBraRedes. (LIMA-NETO et al, 2018a-c).

Tendo em vista as múltiplas possibilidades de correlações e análise que podem ser extraídas da pesquisa supracitada, neste artigo, propõem-se a utilização inovadora da técnica de agrupamento da análise multivariada de dados com o teste de independência para verificar a associação entre o consumo de álcool e as variáveis sócias demográficas: maior grau de instrução da família, remuneração, nota mais frequente, idade estratificada, prática religiosa, raça, sexo, modalidade de curso e classificação sócio demográfica do domicílio.

Para este trabalho fez-se um recorte dos dados da pesquisa destacando o *Campus* Salvador², pela exclusão dos demais *Campi*; a amostra correspondente é de 256 estudantes numa população de 5.833 (IFBA, 2017), com erro amostral de 6,11%.

Assim, o objetivo deste estudo é testar a “Análise de Agrupamentos”, integrante da metodologia de “Análise Multivariada de Dados”, como ferramenta estatística que evidenciasse o grupo de risco dentre os dados obtidos. Como é uma aplicação inovadora, vamos detalhar neste artigo o uso dessa abordagem, de forma a permitir sua replicação e melhorias.

Para análise, dentre vários instrumentos, utilizou-se o *Alcohol Use Disorders Identification Test* (AUDIT), um teste para identificação de problemas relacionados ao uso do

¹ Submetida a Comitê de Ética em Pesquisa sob o código CAAE: 73745317.7.0000.5031, aprovada conforme o Parecer Circunstanciado nº 2.307.870, em 01 de outubro de 2017.

²Incluindo a unidades de Salinas das Margaridas.

álcool, elaborado pela Organização Mundial de Saúde (OMS). Esse teste rápido mostra o quão arriscado está o consumo de álcool pelo indivíduo, as desordens associadas ao uso de álcool e drogas e, possíveis dependentes de álcool, validado no Brasil por Lima *et al.*, (2005)³.

Neste artigo, o pressuposto é que a aplicação de métodos estatísticos mais refinados no tratamento de dados e análise de informações podem ser úteis ao permitir a identificação do público alvo mais crítico, sobre o qual deverão advir, prioritariamente, ações futuras mais consequentes e efetivas na gestão da saúde da população considerada. Espera-se que a aplicação inovadora da metodologia no diagnóstico do consumo de álcool em instituições de ensino, como ferramenta auxiliar seja significativa para a tomada de decisão e estratégias organizacionais, oriente caminhos mais efetivos para uma possível mitigação do consumo de álcool no *Campus* Salvador do IFBA e, por extensão, que auxilie no tratamento dessa significativa questão de saúde nas organizações e na sociedade como um todo.

2 Contextualização

O consumo de álcool, fortemente influenciado pelo estímulo da propaganda, tem sido reconhecido como um elemento de socialização que contribui nos processos de autoafirmação e de inclusão, tanto na esfera juvenil quanto adulta, cujo exagero traz enormes prejuízos pessoais e coletivos.

Estudos apontam que o consumo de drogas vem sendo um importante fator de risco para comportamentos violentos como homicídios, suicídios, violência doméstica e acidentes de trânsito. (LARANJEIRA; DUAILEBI; PINSKY, 2005; MORAES, 2001).

Segundo Kerr-Corrêa *et al.* (2008), os danos provocados pelo álcool sobre os neurônios do cérebro interrompem processos-chave do desenvolvimento trazendo como consequência uma perda cognitiva que pode levar, simultaneamente, há um aumento do descontrole na quantidade de bebida alcoólica ingerida. Ressalta o autor que, os danos causados são inversamente proporcionais à idade de início do consumo de álcool:

Tais danos estão associados ao cérebro, prejudicando a memória, comprometendo diretamente a aprendizagem. Além disso, o consumo de álcool pode favorecer o surgimento de problemas familiares, incluindo violência, vida sexual promíscua, caracterizado como comportamento de alto risco por acarretar risco de contrair AIDS e gravidezes precoces. (KERR-CORRÊA *et al.*, 2008, p. 3).

A NIAAA⁴ (2003) e Slawewski (2002), ao tratar da estrutura cerebral, alertam que o hipocampo é o responsável pela aprendizagem e memória, sendo a região mais prejudicada pelo uso do álcool. Kerr-Corrêa *et al.* (2008, p. 12) afirmam que:

[...] as alterações no seu volume estariam relacionadas à idade de início do uso da bebida, sendo maior quanto mais precoce for o início do consumo de álcool. Mesmo com o declínio da ingestão alcoólica, o dano cerebral permanece durante a fase adulta.

Merece destaque que as alterações na capacidade de compreensão da emoção ligada a linguagens provocadas pelo álcool, resultam em erros de julgamento. Se associado à capacidade de agressividade causada pela ação psicoativa do álcool, seu uso pode desencadear comportamentos violentos. (DE ALMEIDA; PASA; SCHEFFER, 2009).

Considerando que a juventude é a fase de construção da identidade, não é desejável que nessa idade seja criado o hábito de experimentar situações sobre o efeito do álcool, como festas, namoros, relações sexuais ou afetivas, criando, assim, um pareamento emocional entre o uso de bebidas e as decorrentes sensações de prazer, de forma que o consumo alcoólico, cada vez mais frequente, poderá abrir as portas da dependência.

³O AUDIT é composto de 10 questões que se referem aos últimos 12 meses sendo: três itens sobre frequência e quantidade do uso de álcool, quatro itens sobre dependência e três itens sobre problemas decorrentes do consumo. De um total de 47 possibilidades de resposta, como algumas delas têm peso zero, o valor final do escore que varia de 0 a no máximo 40. Uma pontuação superior a oito (8) indica a necessidade de um diagnóstico mais específico.

⁴National Institute of Alcohol Abuse and Alcoholism

As alterações na capacidade de compreensão da emoção ligada à linguagem, provocadas pelo álcool, resultando em erros de julgamento que podem apresentar comportamentos de irresponsabilidade e negligência com as tarefas ou consigo mesmo, conflitos interpessoais ou familiares além de estar presente em grande parte dos casos de comportamentos violentos, como homicídios, suicídios, violência doméstica e acidentes de trânsito. Ademais tem um problema por vezes negligenciado pelos jovens, famílias e escolas que é o comprometimento cerebral, especialmente nessa fase da adolescência, quando ocorre a formação de novas conexões sinápticas em um cérebro que pode estar inundado de substância psicoativas exógenas.

Os dados da pesquisa sobre o consumo de álcool pelos discentes no *Campus Salvador/IFBA*, utilizando o instrumento AUDIT, estratificada do Banco de Dados original, evidenciaram o Perfil Sócio Demográfico da amostra, sintetizado na Tabela 1, conforme o recorte deste estudo. Apresenta-se, para cada variável que caracteriza o Perfil, a categoria que mais prevaleceu, em negrito.

Tabela 1 – Perfil Sócio Demográfico da amostra - IFBA/Campus Salvador

Nota mais frequente	Absoluta	Percentual	Prática Religiosa		
A (>8-10)	39	15,23%	Pr0- Não tenho religião	84	32,81%
B (>6-8)	166	64,84%	Pr1- Só oro/rezo ou acredito	80	31,25%
C (>4-6)	49	19,14%	Pr2- menos que 1x / mês	23	8,98%
D (>2-4)	1	0,39%	Pr3- pelo menos 2x / mês	8	3,13%
E (0-2)	1	0,39%	Pr4- 1x / semana	31	12,11%
Idade Estratificada			Pr5- 2x / semana ou mais	30	11,72%
F1-Idade invalida	9	3,52%	Raça (Cor)		
F3-Idade>=16 e <=17	18	7,03%	Amarelo (asiático, japonês)	2	0,78%
F4-Idade>=18 e < 22	92	35,94%	Branco	54	21,09%
F5-Idade>=22 e < 29	69	26,95%	Miscigenado	1	0,39%
F6-Idade>=29	68	26,56%	Pardo	116	45,31%
Remuneração			Preto	83	32,42%
R0-Não trabalhei	105	41,02%	Class. Sócio Demográfica do Domicílio		
R1- Outras Rendas	43	16,80%	A	12	4,69%
R2- Esporádico	10	3,91%	B	98	38,28%
R3- Período parcial	38	14,84%	C	103	40,23%
R4- Período integral	60	23,44%	D-E	43	16,80%
Sexo			Modalidade		
Feminino	114	44,53%	M1- Não Informado	1	0,39%
Masculino	134	52,34%	M2- Integrado	84	32,81%
Prefiro não dizer	8	3,13%	M3- Integrado - Proeja	1	0,39%
Maior Grau Instrução Família			M4- Subsequente	24	9,38%
4ª Série Fundamental	18	7,03%	M5- EAD	2	0,78%
Analf/até 3ª Série EF	10	3,91%	M6- Superior	137	53,52%
EF completo	22	8,59%	M7- Pós-graduação	7	2,73%
Médio completo	109	42,58%	Total geral	256	100,00%
Superior completo	97	37,89%			

Fonte: LIMA-NETO *et al.*, 2018d.

3 Desenvolvimento Conceitual: Abordagem Estatística

Este item descreve o suporte teórico, fundamentos e a base científica para o alcance do objetivo inicialmente proposto cujos conceitos centrais fornecem a base teórica que foi utilizada para a análise, discussão e conclusões. Por ser uma aplicação inovadora, mas significativamente técnico, cabe maior aprofundamento e detalhamento da metodologia utilizada, que embasará a análise, conclusão e possíveis replicações.

3.1 Análise de Agrupamentos (AG)

São variadas as definições acerca do significado e aplicações da Análise de Agrupamentos (AG). De modo geral, a AG classifica os objetos, agrupando-os. Os grupos formados devem ter uma baixa heterogeneidade interna, dentro de cada grupo e elevada heterogeneidade externa, entre grupos. Em uma AG bem realizada, os objetos dentro do agrupamento estarão mais próximos graficamente e distantes em diferentes agrupamentos; portanto, a AG avalia a estrutura, agregando objetos baseada na distância.

A AG é conceituada por Hair *et al.* (2009), como um conjunto de técnicas com objetivo de agregar objetos, com base em suas características; por exemplo, respondentes, produtos. A AG tem sido empregada em diversas áreas do conhecimento, entre elas: psicologia, biologia, sociologia. Segundo os autores, o papel mais comum da AG é a redução de dados através da diminuição da informação ou geração de hipóteses, desenvolvendo-as ou examinando-as.

Provost e Fawcett (2016) definem AG como técnica de segmentação não supervisionada. Isto significa não ter uma característica predeterminada, uma variável alvo. Eles focam, principalmente, na noção de distância dos objetos. Por outro lado, Silva *et al.* (2016) opinam sobre a AG como sendo um processo que permite descobrir relações existentes entre exemplares de um conjunto de dados descritos por uma série de características.

As definições supracitadas convergem em semântica. A variável estatística de Agrupamentos de Hair *et al.* (2009, p. 430), é “o conjunto de variáveis que representam as características que servem como base para comparar objetos”. Já Provost e Fawcett (2016, p. 46, 142) consideram “cada objeto com um vetor de características” por meio de um conjunto de atributos e Silva *et al.* (2016, p. 145) empregam o termo “atributos descritivos”.

Logo, a AG depende do conceito de similaridade. Segundo Hair *et al.* (2009, p. 429), “similaridade corresponde a associação em dois objetos, baseada nas variáveis estatísticas”. Foster, Provost e Fawcett (2016) entendem similaridade como “medida de distância entre objetos”, coincidindo com a visão de Silva *et al.* (2016). Hair *et al.* (2009), definem a Distância Euclidiana como uma medida comumente mais usada de similaridade entre dois objetos; é uma medida de comprimento de um segmento de reta desenhado entre eles.

A AG pode ser classificada em duas grandes categorias: Análise Hierárquica de Agrupamento (AHG) e Análise Não-Hierárquica de Agrupamento (ANHG), conhecida por Foster e Provost e Fawcett (2016), como Agrupamento em Torno de Centroide (Centro de Agrupamento) e por Silva *et al.* (2016), Agrupamento por Partição.

A AHG envolve uma série de “n-1” decisões de agrupamento das “n” observações, que se combinam duas a duas numa estrutura hierárquica ou do tipo árvore. Os procedimentos básicos são o aglomerativo e o divisivo. No procedimento aglomerativo, cada grupo é formado a partir do agrupamento de uma observação ou grupo de observações com outra observação ou grupo com menor distância até restar um único grupo. Já o procedimento divisivo é o procedimento inverso ao aglomerativo, separando primeiro o grupo e/ou observações de maior distância até cada grupo possuir observações com mesmas características.

ANHG utiliza o conceito de centroide, valor médio dos objetos contidos em cada agrupamento. As observações ocupam uma posição e os centroides se deslocam, remodelando

grupos com observações mais próximas. Este processo continuará até que não haja mudanças nos agrupamentos. O Algoritmo de Agrupamento desta categoria mais conhecido é o K-means⁵.

Hair *et al.* (2009) sugerem o uso da AHG para gerar um conjunto completo de soluções aplicáveis que será entrada para o refinamento da solução por meio do emprego da ANHG. Durante a AG, a multicolinearidade e as observações atípicas devem ser consideradas. A multicolinearidade corresponde ao grau em que uma variável é explicada por outras variáveis. Seu alto valor aumenta a complexidade na interpretação das variáveis estatísticas de agrupamento, dificultando a identificação das inter-relações das variáveis. Emprega-se a Análise Fatorial como técnica de detecção da multicolinearidade. Já as observações atípicas são aquelas com substanciais diferenças em relação às outras, por exemplo, valor muito extremo em uma ou mais características. Cabe ao pesquisador decidir a manutenção, ou não, da análise.

3.2 Análise Fatorial (AF)

Segundo Hair *et al.* (2009, p. 102), a Análise Fatorial (AF) é uma “técnica de interdependência cujo propósito é a definição inerente entre as variáveis”. Assim, a AF avalia a estrutura com base nos padrões de variações (correlação) dos agrupamentos das variáveis, conhecidos como fatores.

Esta técnica resume as dimensões e/ou reduz dados. No resumo das dimensões, a variável estatística da AF é a composição linear das variáveis, maximizando a explicação do conjunto inteiro de variáveis. Desta forma, mitiga a complexidade da estrutura das variáveis. Na redução de dados, a AF identifica as variáveis mais representativas a partir de um conjunto muito maior de variáveis para uso posterior de técnicas multivariadas ou pela criação de um conjunto novo de variáveis, muito menor, para substituir parcial ou completamente o conjunto original de variáveis. Por tanto, pode ser utilizada numa perspectiva exploratória, conhecida como AF do tipo R ou a confirmatória, conhecida como AF do tipo Q.

As variáveis da AF devem ser numéricas. A existência de variáveis categóricas necessita ser transformada em variáveis dicotômicas, sendo um (1) representando pela existência da categoria e zero (0), caso contrário. Silva *et al.* (2016) sugerem um processo de transformação das variáveis categóricas para as dicotômicas. Inicialmente, cria-se uma matriz zerada, trocando por um somente as variáveis na qual a categoria esteja presente.

3.3 Metodologia de Análise de Agrupamentos

Esse estudo adota o tratamento estruturado para construir Modelos Multivariados em seis estágios, sugerido por Hair *et al.* (2009). Os três primeiros estágios lidam com questões relativas à preparação para a estimação do modelo real conforme Figura 1, em anexo.

O Estágio 1 corresponde a Definição do Problema da Pesquisa, dos Objetivos e da Técnica Multivariada; o Estágio 2, denominado de Desenvolvimento do Plano de Análise, define o local e período da realização da pesquisa. Sendo uma Análise de *Agrupamentos*, torna-se necessário definir: tamanho amostral, padronização, detecção de observações atípicas e medida de similaridade; o Estágio 3 ocorre com a Avaliação das Suposições Inerentes à Técnica Multivariada e Suposições da Análise de *Agrupamentos*, verifica o impacto de multicolinearidade das variáveis, empregando a Análise Fatorial.

Os últimos três estágios tratam dos aspectos pertinentes à estimação do modelo, interpretação e validação (Figura 2, em anexo).

O Estágio 4, denominado de Estimação do Modelo Multivariado e Avaliação do Ajuste Geral do Modelo, utiliza o resultado da Análise Fatorial, sendo composto de dois passos: Análise Hierárquica de Agrupamento (AHG) e Análise Não-Hierárquica de Agrupamento (ANHG).

⁵Foster e Provost, 2016, p. 170.

O Estágio 5, consiste na Interpretação das Variáveis Estatísticas, consiste em interpretar os agrupamentos encontrados, analisando as características marcantes de cada grupo. A partir daí, pode-se rotular os grupos.

O Estágio 6, denominado de Validação e Perfil dos Agrupamentos, compreende adotar procedimento de validação, selecionando os pontos sementes por meio da amostragem estratificada por grupo que sirva de entrada para a Análise Não-Hierárquica de Agrupamento. Os centroides encontrados neste estágio são comparados com os centroides do quarto estágio.

4 Aplicação da Metodologia: Tratamento Estatístico dos Dados

Inicialmente, apresenta-se uma síntese da metodologia utilizada e em seguida seu detalhamento passo a passo.

Para o recorte da pesquisa original, foram utilizadas inicialmente todas as possibilidades de respostas do questionário AUDIT, no total de 47, que se referem apenas ao “Consumo de Álcool” pelo corpo discente do *Campus* Salvador/IFBA e de Salinas das Margaridas.

A metodologia utilizada exigia que essas 47 questões do AUDIT fossem transformadas em variáveis dicotômicas. O algoritmo utilizado na metodologia selecionou apenas 14 delas, das quais apenas sete foram utilizadas para diferenciar os grupos. Essas sete originaram quatro agrupamentos que foram correlacionados com as nove informações sócio demográficas dos respondentes, seguintes: (i) Idade do entrevistado; (ii) Modalidade do curso; (iii) Classificação econômica do domicílio (IBGE Critério Brasil 2016); (iv) Maior Grau de Instrução da família; (v) Cor ou raça (segundo o IBGE); (vi) Notas (conceitos) com maior frequência; (vii) Remuneração econômica regular, periódica; (viii) Sexo do entrevistado; (ix) Prática Religiosa.

Esse recorte de informações foi, então, tratado pela Análise de *Agrupamentos* da Análise Multivariada de Dados, obtendo-se quatro grupos com baixa heterogeneidade interna e elevada heterogeneidade externa, rotulados como: (i) Gr1 – Sem consumo de álcool; (ii) Gr2 – Baixo consumo de álcool; (iii) Gr3 – Médio consumo de álcool; (iv) Gr4 – Alto consumo de álcool.

Em seguida, verificou-se a associação entre estes grupos e as informações sócio demográficas, calculando o teste de independência entre eles, utilizando a tabela do Qui-Quadrado, obtendo a probabilidade da significância de cada associação. Assim, calculou-se o Coeficiente de Contingência Moderado. A partir daí, realizou-se a análise vertical e horizontal dos dados das associações com maior probabilidade de rejeição da hipótese “Não existência de associação entre Consumo de Álcool” e cada informação sócio demográfica.

Para cada um dos quatro grupos evidenciou-se os dados absolutos e percentuais correspondente a cada variável sócio demográfico (Tabela 9). Na sequência, verificou-se a associação entre estes grupos e as informações sócio demográficas, calculando-se o teste de independência entre eles, utilizando a tabela do Qui-Quadrado, obtendo a probabilidade da significância de cada associação. A partir daí, definiu-se o nível de significância a ser aceito a partir da significância prática, selecionando as associações que atendessem este critério, assim calculando o Coeficiente de Contingência Moderado (Tabela 10). Em seguida, foram feitas as análises de dados a partir das Tabelas 9 e 10 obtidas, com as informações mais significativas e conclusivas, a partir das associações entre variáveis de maior probabilidade estatística.

Conforme a metodologia apresentada, este estudo obteve os resultados a seguir, descritos em cada estágio.

4.1 Definição do Problema: Pesquisa, Objetivos e Técnica Multivariada (Estágio 1)

Foi definido como objetivo o desenvolvimento de uma taxonomia que particione objetos em grupos, conforme o seu consumo de álcool, ou não. Para tanto, utilizou-se as 10 questões do AUDIT. Pelo fato destas questões serem categóricas, estas foram transformadas em variáveis dicotômicas no estágio adequado com objetivo de serem usadas na Análise de *Agrupamentos*.

4.2 Desenvolvimento do Plano de Análise (Estágio 2)

O tamanho amostral foi de 256 indivíduos da população de 5.833 alunos do *Campus Salvador* e também Salinas da Margarida, correspondendo cerca de 4,39% da população, com erro amostral de 6,11%.

A base de dados disponível possui 153 perguntas, tratadas como variáveis categóricas, dentre elas, as dez do AUDIT. Somente estas foram consideradas variáveis estatísticas, sendo suas respostas transformadas em 47 variáveis dicotômicas, segundo Silva *at al.* (2016), adequando-as às premissas de submissão para a AF e AG, utilizando a Linguagem R. Doravante, o campo “Var” da Tabela 2 representa a respectiva pergunta e resposta. As respostas marcadas pelos respondentes terão valor um (1) ou zero (0), caso contrário. Na Tabela 2 encontram-se as Variáveis Dicotômicas conforme o questionário AUDIT⁶

Os valores mínimos e máximos encontrados de cada variável foram 0 (zero) e 1 (um) respectivamente, caracterizando a inexistência de valores fora deste limite, devido a representação de variáveis categóricas. Os desvios padrões das variáveis diferente de zero refletem a não ocorrência de constantes nestas variáveis. Os desvios padrões das variáveis V16, V26, V29 e V31 foram zerados, logo foram excluídas da análise. A medida de similaridade utilizada foi calculada através do coeficiente de comunidade de Jaccard⁷. As observações atípicas não foram detectadas e nem padronização dos valores foram necessários por se tratar de dados dicotômicos. Por fim, os pesquisadores resolveram utilizar toda amostra por considerar relevante a identificação de possíveis pequenos grupos. Os valores mínimos, máximos, moda e desvio padrão das variáveis dicotômicas, são apresentados na Tabela 2.

Tabela 2 – Valores Mínimo, Máximo, Moda e Desvio Padrão das Variáveis Dicotômicas

Var	Mínimo	Maximo	Moda	Desvio Padrão	Var	Mínimo	Maximo	Moda	Desvio Padrão
V1	0	1	0	0,454	V23	0	1	0	0,235
V2	0	1	0	0,486	V24	0	1	0	0,063
V3	0	1	0	0,44	V25	0	1	0	0,063
V4	0	1	0	0,203	V27	0	1	1	0,174
V5	0	1	0	0,163	V28	0	1	0	0,163
V6	0	1	0	0,476	V30	0	1	0	0,063
V7	0	1	0	0,406	V32	0	1	1	0,385
V8	0	1	0	0,419	V33	0	1	0	0,356
V9	0	1	0	0,34	V34	0	1	0	0,139
V10	0	1	0	0,22	V35	0	1	0	0,063
V11	0	1	0	0,194	V36	0	1	0	0,088
V12	0	1	0	0,499	V37	0	1	1	0,391
V13	0	1	0	0,475	V38	0	1	0	0,364
V14	0	1	0	0,356	V39	0	1	0	0,108
V15	0	1	0	0,228	V40	0	1	0	0,063
V17	0	1	1	0,303	V41	0	1	0	0,124
V18	0	1	0	0,263	V42	0	1	1	0,275
V19	0	1	0	0,139	V43	0	1	0	0,203
V20	0	1	0	0,063	V44	0	1	0	0,194
V21	0	1	0	0,063	V45	0	1	1	0,281
V22	0	1	1	0,249	V46	0	1	0	0,163
					V47	0	1	0	0,235

Fonte: Autores.

⁶ LIMA-NETO, J. L. A. *at al*, 2018d, p.54

⁷ <https://www.rdocumentation.org/packages/vegan/versions/1.6-2/topics/vegdist>.

4.3 Avaliação das Suposições Inerentes à Técnica Multivariada e Suposições da Análise de Agrupamentos (Estágio 3)

Averiguou-se a multicolinearidade⁸ entre variáveis de agrupamento, utilizando a Análise Fatorial do tipo R nas variáveis do AUDIT no intuito de ponderar igualmente as variáveis ao empregar AG, conforme fluxo de decisão proposto por Hair *et al.* (2009, p. 101).

Durante a Análise Fatorial, foi constatada uma relação “ruim” entre número de observação e variáveis, com valor de 5,38, ou seja, inferior a oito (8), com 288 das 1.081 correlações com significância de 0,01 conforme tabela abaixo. Constatou-se a impossibilidade de obtenção da matriz de correlação, por não admitir matriz inversa. Logo, a Análise Fatorial é uma técnica inadequada para este estudo conforme resultados (Tabela 3).

Tabela3 – Resultados do Processo da Análise Fatorial

Critério	Amostra Estudada
Total de registros	256
Relação Número de Observação/Variável *	5,38
Número de Correlações Significantes	288
Número de Total de Correlações	1081
% Correlações Significantes	26,64%
Matriz Correlação Parcial	Impossibilidade de cálculo

Fonte: Autores.

Nota: * Relação Número de Observação / Variável ideal ≥ 8 .

4.4 Estimação do Modelo Multivariado e Avaliação do Ajuste Geral do Modelo (Estágio 4)

4.4.1 Análise Hierárquica de Agrupamento (AHG) - Passo 1

Foi utilizado o procedimento hierárquico para a formação de grupos com até 10 soluções, através do cálculo da matriz de similaridade pelo coeficiente de Jaccard da função *Vegdist* da Linguagem de Programação “R”, conforme a Tabela 4.

Tabela 4 – Quantidade de grupo por Solução

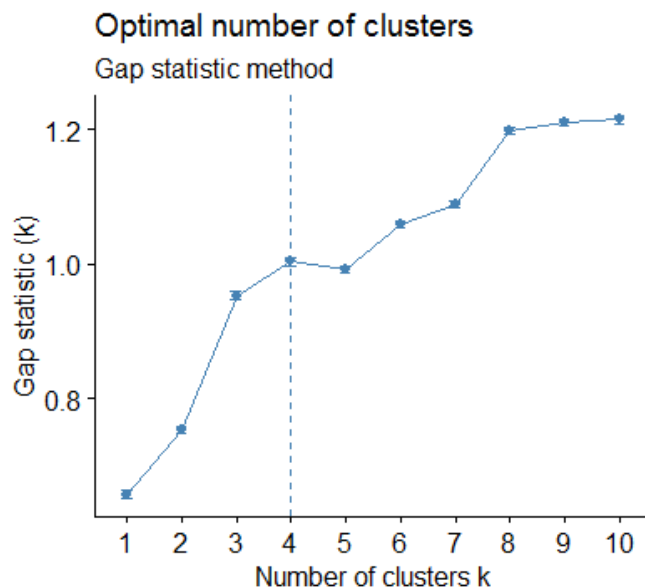
Soluções / Grupos	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
G1	256	223	111	111	111	80	80	80	80	80
G2	0	33	112	112	65	65	65	65	65	65
G3	0	0	33	25	25	25	7	7	7	7
G4	0	0	0	8	47	31	18	18	18	18
G5	0	0	0	0	8	47	31	31	31	31
G6	0	0	0	0	0	8	47	47	43	43
G7	0	0	0	0	0	0	8	3	4	4
G8	0	0	0	0	0	0	0	5	3	2
G9	0	0	0	0	0	0	0	0	5	5
G10	0	0	0	0	0	0	0	0	0	1
Total	256	256	256	256	256	256	256	256	256	256

Fonte: Autores.

A Tabela 4, a partir do método *Gap Statistic*, dá origem ao Gráfico 1, que sugeriu “quatro”, como quantidade adequada de grupos.

⁸Hair *et al.* (2019, p. 22) “Extensão em que uma variável pode ser explicada pelas outras variáveis na análise”.

Gráfico 1 - Número ótimo de agrupamentos

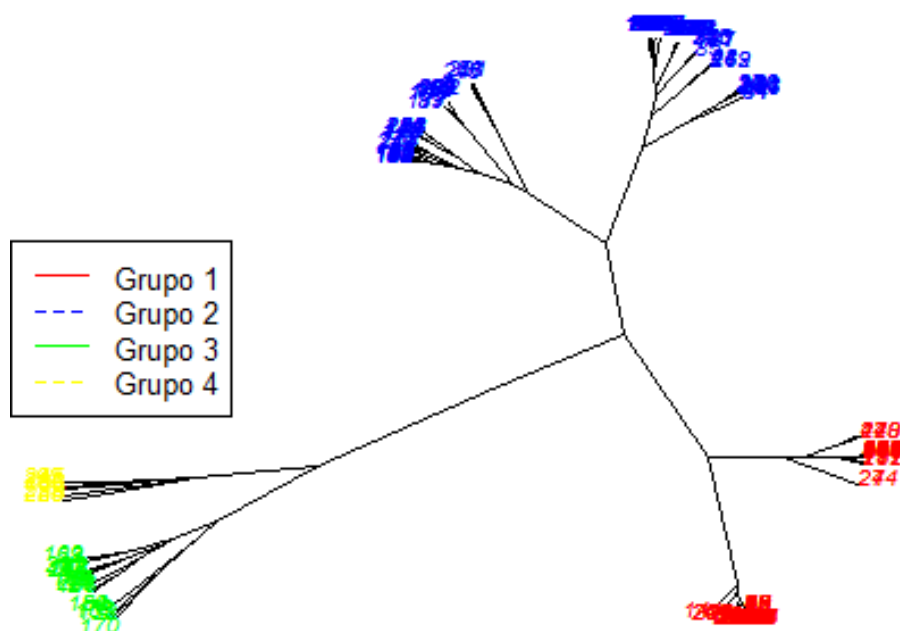


Fonte: Autores.

Assim, o percentual do tamanho de cada grupo para quantidade de quatro soluções está apresentado na Tabela 6. Como definição, esta pesquisa não pretendeu excluir nenhum grupo, logo não houve necessidade de re-especificação dos resultados. Os grupos estão representados pelo filograma, retratando cada agrupamento, obtido através da utilização da Linguagem “R” aplicada ao agrupamento S4.

Cabe destacar que a variável “consumo de álcool” foi a única utilizada para categorizar os grupos pelo fato dela expressar a diferenciação ente os quatro agrupamentos obtidos, a partir do cálculo da matriz de similaridade pelo coeficiente de Jaccard, conforme Figura 3.

Figura 3 - Filograma do agrupamento S4



Fonte: Autores.

Uma vez definida a quantidade de soluções, os pontos sementes de cada grupo das 43 variáveis estão explicitados conforme Tabela 5, que mostra os pontos sementes.

Tabela 5 - Pontos Sementes do Agrupamento S4

Variável/Grupo	G1	G2	G3	G4	Variável/Grupo	G1	G2	G3	G4
V1	1	0	0	0	V24	0	0	0	0
V2	0	1	0	1	V25	0	0	0	0
V3	0	0	1	0	V27	1	1	1	1
V4	0	0	0	0	V28	0	0	0	0
V5	0	0	0	0	V30	0	0	0	0
V6	1	0	0	0	V32	1	1	1	0
V7	0	0	0	1	V33	0	0	0	1
V8	0	1	0	0	V34	0	0	0	0
V9	0	0	1	0	V35	0	0	0	0
V10	0	0	0	0	V36	0	0	0	0
V11	0	0	0	0	V37	1	1	0	0
V12	1	0	0	1	V38	0	0	1	1
V13	0	1	0	0	V39	0	0	0	0
V14	0	0	1	0	V40	0	0	0	0
V15	0	0	0	0	V41	0	0	0	0
V17	1	1	0	1	V42	1	1	0	0
V18	0	0	1	0	V43	0	0	1	0
V19	0	0	0	0	V44	0	0	0	1
V20	0	0	0	0	V45	1	1	1	1
V21	0	0	0	0	V46	0	0	0	0
V22	1	1	1	1	V47	0	0	0	0
V23	0	0	0	0					

Fonte: Autores.

4.4.2 Análise Não-Hierárquica de Agrupamento (ANHG) - Passo 2

Foi utilizado o Agrupamento Não-Hierárquico K-modes, por tratar de dados categóricos, tendo os quatro pontos sementes iniciais gerados pelo agrupamento hierárquico. A tabela abaixo mostra os centroides encontrados. As variáveis V4, V5, V10, V11, V15, V18, V19, V20, V21, V23, V24, V25, V28, V30, V34, V35, V36, V39, V40, V41, V43, V44, V46, V47 estão com todos os mesmos valores zerados. Já as variáveis V17, V22, V27, V42, V45 dos centroides estão com valor 1, nos quatros grupos respectivamente, sendo todas elas suprimidas da Tabela 6. Consequentemente, todas variáveis desta tabela foram suprimidas, já que a igualdade dos valores não diferencia os grupos.

Tabela 6 - Centroides obtidos do Agrupamento S4 pelo K-modes

Variável / Grupo	G1	G2	G3	G4	Variável / Grupo	G1	G2	G3	G4
V1	1	0	0	0	V13	0	1	0	0
V2	0	1	0	1	V14	0	0	1	0
V3	0	0	1	0	V32	1	1	0	1
V6	1	0	0	0	V33	0	0	1	0
V7	0	0	0	1	V37	1	1	0	1
V8	0	1	0	0	V38	0	0	1	0
V9	0	0	1	0	Total	95	92	37	32
V12	1	0	0	1	Perc (%)	37,11	35,94	14,45	12,5

Fonte: Autores.

4.5 Interpretação das Variáveis Estatísticas (Estágio 5)

Foram selecionadas características definidoras de cada grupo, que vão permitir a consolidação dos dados das tabelas subsequentes:

- ✓ O Grupo1, doravante **Gr1-SEM Consumo de Alcool**: não consumiu bebida alcoólica, correspondendo o grupo de maior tamanho;
- ✓ O Grupo 4, doravante **Gr2 - BAIXO Consumo de Alcool**: consumiu 1 a 2 doses, no máximo, uma vez ao mês, pouca ingestão comparado aos grupos que consumiram álcool;
- ✓ O Grupo 2, doravante **Gr3 - MÉDIO Consumo de Alcool**: consumiu de 3 e 4 doses no máximo 1 vez ao mês;
- ✓ O Grupo 3, doravante **Gr4- ALTO Consumo de Alcool**: consumiu de 5 a 6 doses, de 2 a 4 vezes ao mês, com muita ingestão de álcool comparado aos demais grupos.

Do banco de dados original, obteve-se a tabela abaixo sobre a frequência dos eventos, por grupo, do agrupamento S4 utilizado.

Tabela 7 – Detalhamento do número de respostas do Agrupamento S4

Quantidade	Respondentes do Consumo de Álcool por Agrupamento				
	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
Total geral	111	8	112	25	256
Percentuais	43,36%	3,13%	43,75%	9,77%	100,00%

Fonte: Autores

4.6 Validação e Perfil dos Agrupamentos (Estágio 6)

Validou-se a Análise Não-Hierárquica de Agrupamento, tomando uma amostra estratificada de cada grupo como pontos sementes iniciais. Após aplicação do K-modes, os centroides encontrados de cada grupo neste estágio coincidiram com os centroides dos grupos da análise anterior. A partir do banco de dados original, para facilitar um estudo de correlação entre as muitas variáveis da pesquisa, decidiu-se por correlacionar estes agrupamentos de consumo de álcool com apenas **nove variáveis**, mostradas no Quadro 1.

Quadro 1 – Variáveis do estudo

VARIÁVEIS ESTUDADAS		
1. Idade do entrevistado	4. Grau de instrução do chefe da família	7. Se trabalhou com remuneração ou recebeu bolsa de estudo nos últimos 6 meses
2. Modalidade do curso	5. Cor ou raça (segundo o IBGE)	8. Sexo do entrevistado
3. Classificação econômica do domicílio (IBGE Critério Brasil 2016)	6. Notas (conceitos) com maior frequência	9. Prática religiosa

Fonte: Autores.

Estas nove variáveis foram escolhidas no estudo para serem objetos de análise de associação com o consumo de álcool pelos estudantes, em função da necessidade de síntese analítica e por terem importante significado social econômico, obtendo-se a Tabela 8, cujas análises dos dados e de possíveis associações serão feitas no próximo item.

5 Análise dos Dados

A partir das informações sobre do número de respostas do Agrupamento S4 (Tabela 5), que foram detalhadas no Perfil da Solução S4 sobre as Variáveis Sócio Demográficas escolhidas (Tabela 8), observou-se que dos quatro agrupamentos considerados, as maiores frequências dos respondentes foram relativas ao em Gr1 e Gr3, Sem Consumo e Médio Consumo de Álcool, respectivamente, com 43,36% e 43,75% de participação. No entanto, por apresentar maior preocupação social, uma análise mais detalhada ocorrerá nos **Gr3 e Gr4, Médio e Alto Consumo** a partir das ocorrências entre estes e as variáveis sócio demográficas. Logo ressalta-se as seguintes características:

- ✓ Gr3-Médio e Gr4-Alto Consumo: Alunos do sexo masculino da modalidade superior cujas notas mais frequentes estão entre 6 e 8, idade entre 18 e 22, cor parda, da classe sócio demográfica do domicílio B ou C;
- ✓ GR3-Médio Consumo apenas: Alunos quem não trabalham, cuja maior instrução do chefe de família é Ensino Médio, apenas ora ou reza;
- ✓ GR4-alto Consumo apenas: Alunos que trabalham parcialmente, cuja maior instrução do chefe de família é superior e não tem religião.

Tabela 8 – Perfil da Solução S4 sobre as Variáveis Sócio Demográficas escolhidas

Variáveis escolhidas	Consumo de Alcool									
Classe de análise	Números Absolutos					Percentuais				
Nota mais Frequente	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
A (>8-10)	16	2	16	5	39	6,25%	0,78%	6,25%	1,95%	15,23%
B (>6-8)	77	2	72	15	166	30,08%	0,78%	28,13%	5,86%	64,84%
C (>4-6)	18	4	22	5	49	7,03%	1,56%	8,59%	1,95%	19,14%
D (>2-4)			1		1	0,00%	0,00%	0,39%	0,00%	0,39%
E (0-2)			1		1	0,00%	0,00%	0,39%	0,00%	0,39%
Idade Estratificada	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
F1-Idade Invalida	3		5	1	9	1,17%	0,00%	1,95%	0,39%	3,52%
F3-Idade>=16 e <=17	12	1	5		18	4,69%	0,39%	1,95%	0,00%	7,03%
F4-Idade>=18 e < 22	33	1	47	11	92	12,89%	0,39%	18,36%	4,30%	35,94%
F5-Idade>=22 e < 29	32	4	26	7	69	12,50%	1,56%	10,16%	2,73%	26,95%
F6-Idade>=29	31	2	29	6	68	12,11%	0,78%	11,33%	2,34%	26,56%
Remuneração	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
R0-Não trabalhei	48	2	48	7	105	18,75%	0,78%	18,75%	2,73%	41,02%
R1- Outras Rendas	19	2	19	3	43	7,42%	0,78%	7,42%	1,17%	16,80%
R2- Esporádico (bicos)	2		6	2	10	0,78%	0,00%	2,34%	0,78%	3,91%
R3- Período parcial	18		13	7	38	7,03%	0,00%	5,08%	2,73%	14,84%
R4- Período integral	24	4	26	6	60	9,38%	1,56%	10,16%	2,34%	23,44%
Sexo	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
Feminino	55	2	48	9	114	21,48%	0,78%	18,75%	3,52%	44,53%
Masculino	54	6	58	16	134	21,09%	2,34%	22,66%	6,25%	52,34%
Prefiro não dizer	2		6		8	0,78%	0,00%	2,34%	0,00%	3,13%
Maior Grau Instrução	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
4ª Série Fundamental	10		7	1	18	3,91%	0,00%	2,73%	0,39%	7,03%
Analfabeto/ até 3ª Série Fundamental	6		4		10	2,34%	0,00%	1,56%	0,00%	3,91%
Fundamental completo	9	1	12		22	3,52%	0,39%	4,69%	0,00%	8,59%
Médio completo	52	4	45	8	109	20,31%	1,56%	17,58%	3,13%	42,58%
Superior completo	34	3	44	16	97	13,28%	1,17%	17,19%	6,25%	37,89%
Prática Religiosa	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
Pr0- Não tenho religião	30	4	39	11	84	11,72%	1,56%	15,23%	4,30%	32,81%
Pr1- Só oro/rezo ou acredito	30	2	42	6	80	11,72%	0,78%	16,41%	2,34%	31,25%
Pr2- menos que 1x / mês	9	2	10	2	23	3,52%	0,78%	3,91%	0,78%	8,98%
Pr3- pelo menos 2x / mês	4		3	1	8	1,56%	0,00%	1,17%	0,39%	3,13%
Pr4- 1x / semana	18		10	3	31	7,03%	0,00%	3,91%	1,17%	12,11%
Pr5- 2x / semana ou mais	20		8	2	30	7,81%	0,00%	3,13%	0,78%	11,72%
Raça (Cor)	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
Amarelo (asiático, japoneses)	1			1	2	0,39%	0,00%	0,00%	0,39%	0,78%
Branco	22		23	9	54	8,59%	0,00%	8,98%	3,52%	21,09%
Miscigenado	1				1	0,39%	0,00%	0,00%	0,00%	0,39%
Pardo	44	4	59	9	116	17,19%	1,56%	23,05%	3,52%	45,31%
Preto	43	4	30	6	83	16,80%	1,56%	11,72%	2,34%	32,42%
Class. SocioEconômica	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
A	2		7	3	12	0,78%	0,00%	2,73%	1,17%	4,69%
B	39	6	43	10	98	15,23%	2,34%	16,80%	3,91%	38,28%
C	48		45	10	103	18,75%	0,00%	17,58%	3,91%	40,23%
D-E	22	2	17	2	43	8,59%	0,78%	6,64%	0,78%	16,80%
Modalidade	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral	Gr1-Sem	Gr2-Baixo	Gr3-Médio	Gr4-Alto	Total geral
M1- Não Informado	1				1	0,39%	0,00%	0,00%	0,00%	0,39%
M2- Integrado	42	2	35	5	84	16,41%	0,78%	13,67%	1,95%	32,81%
M3- Integrado - Proeja (formado)				1	1	0,00%	0,00%	0,00%	0,39%	0,39%
M4- Subsequente	15		8	1	24	5,86%	0,00%	3,13%	0,39%	9,38%
M5- EAD			2		2	0,00%	0,00%	0,78%	0,00%	0,78%
M6- Superior	52	5	62	18	137	20,31%	1,95%	24,22%	7,03%	53,52%
M7- Pós-graduação	1	1	5		7	0,39%	0,39%	1,95%	0,00%	2,73%
Total geral	111	8	112	25	256	43,36%	3,13%	43,75%	9,77%	100,00%

Fonte: Autores.

Em seguida, o teste de independência entre os grupos foi realizado, conforme Tabela 10, considerando as seguintes hipóteses:

- ✓ **H0**: inexistência de associação entre Consumo de Álcool e cada variável sócio demográfica;
- ✓ **H1**: o caso contrário, ou seja, a existência de associação entre Consumo de Álcool e cada variável sócio demográfica.

Tabela 9 -Testes de Independência e cálculo do Coeficiente de Contingência Moderado

Associação entre Consumo de Álcool e Variáveis Sócio Demográficas	Qui-quadrado	G.L.*	p-valor	P.S.**	C.C.M.***	Associação
Nota Mais Frequente	10,47	12	0,674	57,52%	0,677	Moderada
Idade Estratificada	12,67	12	0,561	39,33%	0,719	Forte
Remuneração	13,28	12	0,463	34,88%	0,729	Forte
Sexo	6,96	6	0,431	32,45%	0,742	Forte
Maior grau Família	13,96	12	0,62	30,34%	0,740	Forte
Prática Religiosa	18,81	15	0,34	22,24%	0,765	Forte
Raça (Cor)	16,42	12	0,353	17,27%	0,775	Forte
Classe Sócio econômica	14,63	9	0,273	10,17%	0,798	Forte
Modalidade	28,37	18	0,144	5,66%	0,819	Forte

Fonte: Autores.

Nota: * GL- Grau de Liberdade

**P.S. Probabilidade de Significância da Rejeição de H0

***CCM Coeficiente de Contingência Moderado

O primeiro critério adotado para escolha das variáveis sócio demográficas mais significativas no desdobramento das associações entre as variáveis foi a pouca ou nenhuma evidência real contra a hipótese H0. Todas as variáveis atenderam a este critério cujo valor de p-value é superior a 0,10. O próximo critério de escolha foi das variáveis com Probabilidade de Significância de Rejeição de H0 (P.S.) superior a 39%, sendo escolhidas **Nota mais Frequente** e **Idade Estratificada**, evidenciando-se estas variáveis sócio demográficas como as mais significativas dentre as demais, na definição do agrupamento prioritário, a ser objeto de futuras ações mitigadoras do consumo de álcool para a amostra considerada.

6 Considerações Finais

Este estudo inicial, significativamente técnico e inovador, permite algumas considerações a título de conclusão e recomendações. A Técnica de Agrupamentos da Análise Multivariada de Dados enquanto um recurso empregado na análise do consumo de álcool do *Campus* Salvador do IFBA, foi validada, pelo Teste de Independência, nas associações verificadas entre consumo de álcool e variáveis sócio demográficas, evidenciando o caráter inovador da aplicação na gestão da saúde.

Do ponto de vista gerencial, principalmente, a aplicação da Técnica de Agrupamentos permitiu identificar o grupo prioritário a ser objeto de ações efetivas a serem adotadas para mitigar o consumo de álcool, correspondendo aos Grupos **Gr3-Médio** e **Gr4 Alto Consumo**, caracterizados por: **alunos cujas notas mais frequentes estão entre 6 e 8, idade entre 18 e 22** e, complementarmente, do sexo masculino, da modalidade superior, cor parda, da classe sócio demográfica do domicílio B ou C, que não trabalha ou trabalha parcialmente, cuja maior instrução do chefe de família é no mínimo médio, apenas ora ou reza ou sem religião. Esse foi o extrato considerado o mais crítico, a ser objeto de ações gerenciais subsequentes mais significativas.

Vale destacar, que a metodologia empregada neste estudo é diferente daquela constituinte originalmente como suporte ao AUDIT, não sendo possível, portanto, uma comparação analítica simples entre os resultados das duas pesquisas, a original e está aqui apresentada, por distinção metodológica entre ambas. Além de ser um recorte da pesquisa original, esta pesquisa categoriza/classifica variáveis e grupos distintamente e é nisso que se configura seu ineditismo, que precisa, naturalmente, de desdobramentos para ampliar, ainda mais, sua validação.

Neste sentido, cabem diversas recomendações adicionais: aplicar a metodologia com outros arranjos de grupos (3 ou 5, por exemplo) e comparar resultados; ampliar o número de variáveis estudadas e não apenas as nove consideradas; estender a metodologia deste trabalho para todo o IFBA utilizando o Banco de Dados completo da pesquisa feita por Lima Neto, com dados de 2017 e, posteriormente, realizar uma pesquisa atual de maior alcance entre o corpo discente do Instituto, verificando alterações de consumo com o passar do tempo.

Assim, no entendimento dos autores, este estudo cumpre um seu papel orientador mais pragmático, ao utilizar uma metodologia estatística refinada no tratamento dados, para identificar amostra mais crítica de consumo de álcool no IFBA, a partir da pesquisa realizada no *Campus* Salvador, em 2017. Ele pode vir a favorecer ações mais efetivas e consequentes para trabalhos mitigantes futuros na área de psicologia, educação, pedagogia e sociologia, de forma a obter melhores resultados a partir de recursos aplicados de forma mais focada na amostra vital identificada, obtendo-se, assim, maior agregação de valor na gestão da saúde nos âmbitos público e privado.

Como possibilidade de posteriores intervenções pontua-se, dentre as muitas opções existentes, aprofundar o estudo e até experimentar a proposta por Dimeff *et al.* (2002) quando desenvolveram o *Brief Alcohol Screning and Intervention for College Students* (BASICS), uma intervenção breve que foi desenvolvida e padronizada especificamente para estudantes universitários, com o objetivo de reduzir o consumo excessivo de álcool por meio de ações educativas que irão ensinar os estudantes a utilizarem o álcool com moderação.

Esse modelo de intervenção não visa o confronto ou o julgamento sobre o comportamento de consumo; pelo contrário, visa a compreensão e a orientação direcionadas para promover a redução do consumo e das consequências negativas, utilizando os princípios da entrevista motivacional e de estratégias cognitivo-comportamentais específicas (SILVA; TUCCI, 2015). Para os mais jovens, por exemplo, é fundamental iniciar um programa de prevenção do uso de álcool e, estender a todos os níveis (palestras, mesas redondas, campanhas de esclarecimentos etc.).

Seguindo recomendações da OMS sobre mitigação do uso de drogas, se evidenciam três áreas de ação que exigem o “treinamento e apoio ao pessoal educacional, para garantir sua capacidade de fornecer e apoiar respostas de prevenção de alta qualidade ao uso de substâncias nas escolas. A formação geral de professores pode ser fornecida como parte da formação de professores de pós-graduação” (UNODC 2017, p. 12):

a) Respostas baseadas em evidências a nível escolar, que compreende: 1) abordagens de prevenção universal destinadas a garantir um ambiente seguro; 2) prevenção seletiva e indicada que visa indivíduos ou grupos considerados com risco particular de iniciar ou que já estão envolvidos em comportamentos de uso de substâncias;

b) Serviços de saúde escolar adequados, que podem desempenhar um papel importante na prevenção, detecção precoce de distúrbios de uso de substâncias bem como no fornecimento de breves intervenções.

c) Gerenciamento eficaz da resposta do setor educacional, que assegura o estabelecimento de recursos e infraestrutura adequados para o planejamento, coordenação, monitoramento e avaliação das respostas do setor educacional ao uso de substâncias por estudantes a longo prazo.

Neste contexto de uso crescente de drogas, em especial do álcool, e em um mundo cada vez mais volátil, incerto, complexo e ambíguo (VICA), com fluxos de informações e relacionamentos entre indivíduos e organizações crescentemente registrados em armazéns de dados (*Data Warehouse*), faz-se ainda mais necessário tratamentos adequados dos mesmos, via métodos, modelos e algoritmos inteligentes, com inovadoras e necessárias sistematizações de informações. Inéditas aplicações metodológicas de abordagem multidisciplinar, conectando e integrando diversos campos do saber, favorecem a geração de conhecimento e a tomada de decisão com menos risco, maior assertividade e geração de valor para todas as partes interessadas, sendo, portanto, cada vez mais, uma questão de sobrevivência organizacional e de efetividade da gestão, em geral, e, também, no contexto da saúde privada e da pública.

REFERÊNCIAS

- DE ALMEIDA, R. M. M.; PASA, G. G.; SCHEFFER, M. **Álcool e violência em homens e mulheres**. Psicologia: Reflexão e Crítica, v. 22, n. 2, p. 252-260, Porto Alegre, 2009. Disponível em: <<http://www.scielo.br/pdf/prc/v22n2/a12v22n2>>. Acesso em: 27 mai. 2016.
- DIMEFF, Linda A.; BAER, John S.; KIVLAHAN, Daniel R.; MARLATT, G. Alan. **Alcoolismo entre estudantes universitários: Uma Abordagem de Redução de Danos**. (Trad. J. M. Bertolote). Botucatu: UNESP, 2002. Não encontrei no texto
- FRATTES, B. *et al.* **Life Style Medicine Handbook: An Introduction to the Power of Health Habits**. American College of Lifestyle Medicine. Healthy Learning, 2019. ISBN 978-1-60679-413-5 2019.
- HAIR, J. F.; BLACK, W. C.; BABIN, B.J., ANDERSON, R.E. TATHAM, R. L. **Análise Multivariada de Dados**. 6a ed. Porto Alegre: Bookman, 2009.
- IFBA. **Relatório de Gestão Institucional 2017 IFBA**. Disponível em: <https://portal.ifba.edu.br/proap/transparencia-arquivos/relatorios-de-gestao-do-ifba/copy_of_RELATRIODEGESTOINSTITUCIONAL2017_Versocomgrficos_08.10.2018_REVISANDO.pdf>. Acesso em: 18 jan. 2020.
- KERR-CORRÊA, F.; LIMA, M. C. P.; BERTOLETE, J. M.; TRINCA, L. A.; SIMÃO, M. O.; OLIVEIRA, J. B.; COLL, G. S.; PEREIRA, P. L.; FONTANELLA, B. C.; COSTA, D. L. C. **Relatório UNESP sobre o levantamento de uso de álcool e drogas por estudantes do ensino fundamental, médio e supletivo do município de Botucatu**. Botucatu: UNESP, 2008.
- LARANJEIRA, R. (supervisão). **II Levantamento Nacional de Álcool e Drogas (LENAD) – 2012**. São Paulo: Instituto Nacional de Ciência e Tecnologia para Políticas Públicas de Álcool e Outras Drogas (INPAD), UNIFESP. 2014. Disponível em: <<https://inpad.org.br/wp-content/uploads/2014/03/Lenad-II-Relatório.pdf>>. Acesso em: 18 nov. 2016.
- LIMA, C. T.; FREIRE, A. C. C.; SILVA, A. P. B.; TEIXEIRA, R. M.; FARRELL, M.; PRÍNCIPE, M. **Concurrent and construct validity of the Audit in na urban Brazilian sample**. Álcool e Alcoolismo, v. 40, n. 6, p. 584–589, 2005. Disponível em: <<https://doi.org/10.1093/alcalc/agh202>>. Acesso em: 6 abr. 2020.
- LIMA-NETO, J. L. A.; MAGRIS, P. N.; RIBEIRO, N. M.; PEREIRA, H. B. B.; CASAS, T. H. P.; SANTOS, C. S. **Saúde Mental de alunos de uma instituição federal de ensino: A análise de dados a partir da aplicação de elementos da Teoria e Ciência das Redes**. UEFS, Feira de Santana, 2018a. Disponível em: <https://www.researchgate.net/publication/328934916_Saude_Mental_de_alunos_de_uma_in

stituicao_federal_de_ensino_A_analise_de_dados_a_partir_da_aplicacao_de_elementos_da_Teoria_e_Ciencia_das_Redes>. Acesso em: 6 mar. 2020.

LIMA-NETO, J. L. A.; SOUZA, C. R. S.; RIBEIRO, N. M.; FREITAS, M. M.; SANTOS, C. S. **Consumo de álcool por alunos do IFBA: uma avaliação local de um problema global**. XII Congresso Norte-Nordeste de Pesquisa e Inovação (Connepi 2018) At: Recife-PE. 2018b. Disponível em: <https://www.researchgate.net/publication/329171142_CONSUMO_DE_ALCOOL_POR_ALUNOS_DO_IFBA_UMA_AVALIACAO_LOCAL_DE_UM_PROBLEMA_GLOBAL>. Acesso em: 6 mar. 2020.

LIMA-NETO, J. L. A.; SOUZA, C. R. S.; RIBEIRO, N. M.; FREITAS, M. M.; SANTOS, C. S. **Consumo de álcool e drogas por alunos do IFBA: o que os dados nos dizem e o que podemos fazer**. (Apresentação) 1º CONGRESSO DO IFBA CAMPUS DE SALVADOR, Junho 2018c. Disponível em: <https://www.researchgate.net/publication/325790744_Presentation_Consumo_de_alcool_e_drogas_por_alunos_do_IFBA_o_que_os_dados_nos_dizem_e_o_que_podemos_fazer>. Acesso em: 6 mar. 2020.

LIMA-NETO, J. L. A.; SOUZA, C. R. S.; RIBEIRO, N. M.; FREITAS, M. M.; SANTOS, C. S. **Relatório sobre o Levantamento de uso de álcool e drogas por estudantes do IFBA**. Salvador-BA, GEPIO-IFBA, Junho 2018d. Disponível em: <https://www.researchgate.net/publication/326679051_RELATORIO_SOBRE_O_LEVANTAMENTO_DE_USO_DE_ALCOOL_E_DROGAS_POR_ESTUDANTES_DO_IFBA>. Acesso em: 6 mar. 2020.

MORAES, R. **Drogas e álcool: Prevenção e tratamento**. São Paulo, SP: Komedi, 2001.

NIAAA (National Institute of Alcohol Abuse and Alcoholism). **Underage drinking: a major public health challenge**. Alcohol Alert. n. 59, p. 1-4. EUA, 2003. Disponível em: <<https://pubs.niaaa.nih.gov/publications/aa59.htm>>. Acesso em: 8 fev. 2017.

PROVOST, F.; FAWCETT, T. **Data Science para Negócio**. Rio de Janeiro: Alta Books Editora, 2016.

SILVA, E. C.; TUCCI, A. M. A. **Intervenção Breve para Redução do Consumo de Álcool e suas Consequências em Estudantes Universitários Brasileiros**. Psicologia: Reflexão e Crítica. 2015, n. 28, v. 4, p. 728-736. Disponível em: <<https://www.redalyc.org/articulo.oa?id=18842573011>>. Acesso em: 6 abr. 2020.

SILVA, L. A.; PERES, S. M.; BOSCARIOLI, C. **Introdução à Mineração de Dados com Aplicações em R**. Rio de Janeiro: Elsevier Editora, 2016.

SLAWECKI, C. J. **Altered EEG responses to ethanol in adult rats exposed to ethanol during adolescence**. Alcoholism: Clinical and Experimental Research. v. 26, p. 246-254, 2002.

UNODC. **Good Policy and Practice in health education - Booklet 10. Education Sector responses to the use of alcohol, tobacco and drugs**. World Health Organization, Switzerland, 2017. Disponível em https://www.unodc.org/documents/drug-prevention-and-treatment/UNODC_UNESCO_WHO_GoodPolicyAndPracticeInHealthEducation.pdf. Acesso em: 19 jul. 2020.

UNODC. **Informe Mundial Sobre Las Drogas - Resumen Ejecutivo**. World Health Organization, 2016. Disponível em <https://www.unodc.org/doc/wdr2016/WDR_2016_ExSum_spanish.pdf>. Acesso em: 19 jul. 2020.

Figura 1 - Estágios 1-3 do diagrama de decisão da análise de agrupamentos

Fonte: Hair *et al.* (2009, p. 437)

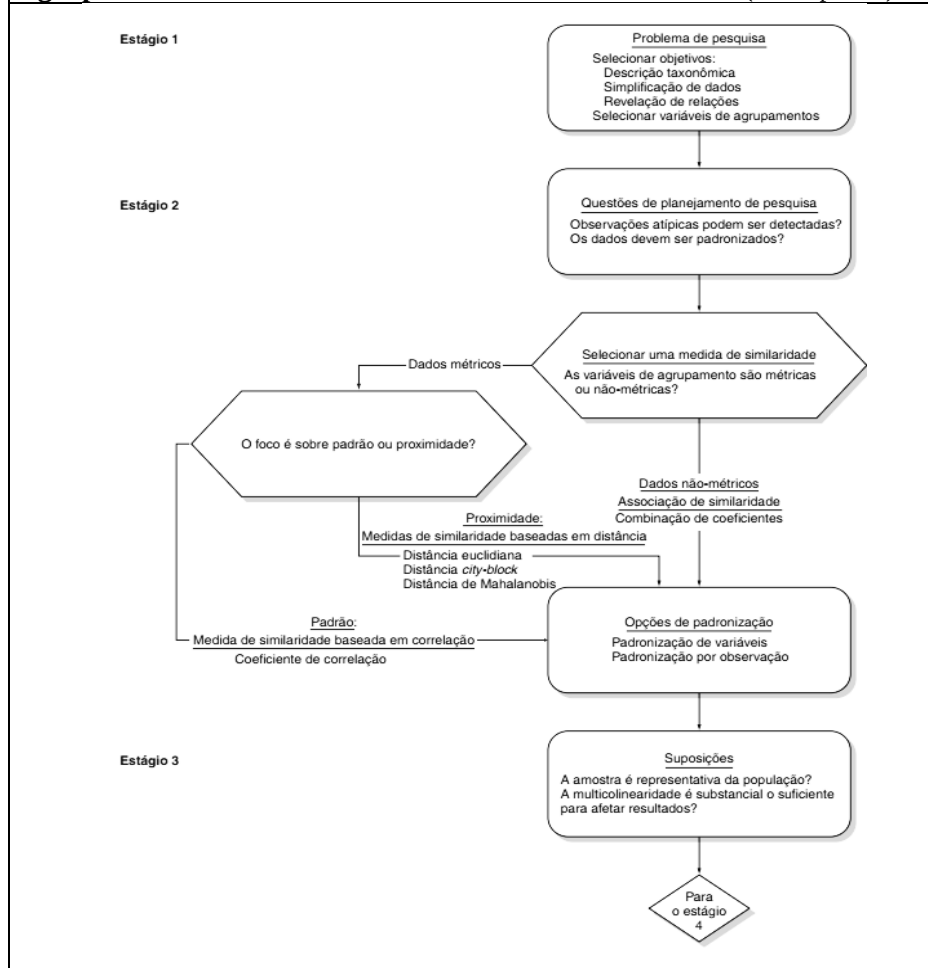


Figura 2 - Estágios 4-6 do diagrama de decisão da análise de agrupamentos

Fonte: Hair *et al.* (2009, p. 448)

