



**Universidade Federal de Uberlândia
Faculdade de Matemática**

Bacharelado em Estatística

**ANÁLISE DE CLUSTER PARA O
QUANTITATIVO, PER CAPITA, DE
RESÍDUOS SÓLIDOS RESIDENCIAIS
RECICLÁVEIS, NA CIDADE DE
UBERLÂNDIA-MG**

Paloma Larissa Viana Silva

Uberlândia-MG

2021

Paloma Larissa Viana Silva

**ANÁLISE DE CLUSTER PARA O
QUANTITATIVO, PER CAPITA, DE
RESÍDUOS SÓLIDOS RESIDENCIAIS
RECICLÁVEIS, NA CIDADE DE
UBERLÂNDIA-MG**

Trabalho de conclusão de curso apresentado à Co-ordenação do Curso de Bacharelado em Estatística como requisito parcial para obtenção do grau de Bacharel em Estatística.

Orientador: Prof. Dr. Ednaldo Carvalho Guimarães

Uberlândia-MG

2021



**Universidade Federal de Uberlândia
Faculdade de Matemática**

Coordenação do Curso de Bacharelado em Estatística

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Estatística.

Uberlândia, _____ de _____ de 20_____

BANCA EXAMINADORA

Prof. Dr. Ednaldo Carvalho Guimarães

Profa Dra. Aurélia Aparecida de Araújo Rodrigues

Profa Dra. Camilla Soueneta Nascimento Nganga

Uberlândia-MG

2021

AGRADECIMENTOS

Agradeço a Deus, por ter permitido que eu tivesse saúde e determinação para não desanimar durante a realização deste trabalho. Agradeço a toda minha família e a todos os meus amigos pelo apoio nessa importante fase do meu processo de formação profissional. Aos meus pais e irmãos, que me incentivaram nos momentos difíceis e compreenderam a minha ausência enquanto eu me dedicava à realização deste trabalho. Ao professor Ednaldo Carvalho Guimarães, por ter sido meu orientador e ter desempenhado tal função com dedicação e amizade. Aos meus colegas de curso, com quem convivi intensamente durante os últimos anos, pelo companheirismo e pela troca de experiências que me permitiram crescer não só como pessoa, mas também como formando.

RESUMO

A crescente produção de resíduos sólidos nos municípios se configura um problema ambiental, social e econômico para as gestões das cidades. A reciclagem de resíduos sólidos residências é uma importante forma de redução dos resíduos que são destinados aos aterros sanitários, além de ser uma fonte de renda para os recicladores e contribuir para a saúde do meio ambiente. Objetivo desse estudo foi analisar o comportamento do quantitativo per capita (kg) da coleta oficial de resíduos sólidos recicláveis, por setor em Uberlândia-MG no período de janeiro a dezembro do ano de 2020. Foi utilizada a metodologia de análise multivariada (análise de cluster hierárquico) nas informações sobre o quantitativo per capita (kg) de resíduos sólidos recicláveis da coleta seletiva realizada pelo DMAE-Uberlândia. Os resultados permitiram identificar setores que se agruparam com quantitativos semelhantes entre si mas cujos quantitativos se diferenciaram entre os grupos e esse mesmo comportamento pôde ser verificado para os meses. A análise de cluster mostrou-se eficaz na identificação de agrupamentos das variáveis em estudo, permitindo fazer a distinção das regiões da cidade e dos meses do ano com os maiores quantitativos da coleta seletiva; Políticas de planejamento e expansão da coleta seletiva; ações de incentivo e de conscientização à destinação correta dos resíduos sólidos recicláveis podem se embasar em estudos como este na definição de suas diretrizes e procedimentos.

Palavras-chave: Análise Multivariada, Cluster, Coleta Seletiva .

Abstract

The growing production of solid waste in municipalities is an environmental, social and economic problem for city management. Recycling solid residential waste is an important way of reducing waste that goes to landfills, in addition to being a source of income for recyclers and contributing to the health of the environment. The objective of this study was to analyze the behavior of the quantitative per capita (kg) of official collection of recyclable solid waste, by neighborhood in Uberlândia-MG, from January to December of 2020. The methodology of multivariate analysis was used (cluster analysis) hierarchical) in the information on the per capita quantity (kg) of recyclable solid waste from the selective collection carried out by DMAE-Uberlândia. The results allowed the identification of neighborhoods that were grouped with similar amounts but whose amounts differed between the groups and this same behavior could be verified for the months. Cluster analysis proved to be effective in identifying groupings of the variables under study, allowing for a distinction to be made between city regions and months of the year with the highest amounts of selective collection; Policies for planning and expanding selective collection; actions to encourage and raise awareness of the correct disposal of recyclable solid waste can be based on studies such as this one in defining its guidelines and procedures.

Keywords: Multivariate Analysis, Cluster, Selective Collection.

SUMÁRIO

Lista de Figuras	I
Lista de Tabelas	III
1 Introdução	1
1.1 Objetivo.....	3
2 Fundamentação Teórica	5
2.1 Dissimilaridade baseada na correlação de Pearson	5
2.2 Distância Euclidiana	6
2.3 Distância Manhattan	6
2.4 Distância de Mahalanobis.....	7
2.5 Método de Agrupamento Hierárquico.....	7
2.6 Definição de Quantidade de Grupos	8
3 Metodologia	9
4 Resultados	11
4.1 Análise Descritiva.....	11
4.2 Método aglomerativo	12
4.3 Análise de agrupamento dos Setores	13
4.4 Análise de agrupamento dos meses.....	16
5 Conclusões	19
Referências Bibliográficas	21

LISTA DE FIGURAS

2.1	Representação Ilustrativa de um Dendograma	7
4.1	Dendograma com 2 grupos referente aos setores atendidos pela coleta seletiva de lixo, na cidade de Uberlândia, no ano de 2020.	13
4.2	Visualização espacial da distribuição dos setores atendidos pela coleta seletiva considerando os grupos extraídos do dendograma, na cidade de Uberlândia- MG, no ano de 2020.	15
4.3	Dendograma com 2 grupos referente aos meses coleta seletiva de lixo, na cidade de Uberlândia, no ano de 2020.	16

LISTA DE TABELAS

4.1	Estatísticas descritivas do quantitativo per capita (kg/mês) da coleta seletiva em Uberlândia-MG, por setor 2020.	11
4.2	Estatísticas descritivas gerais, por mês, do quantitativo per capita (kg/setor) da coleta seletiva nos setores de Uberlândia-MG, no ano 2020.	12
4.3	Grupos formados na análise de agrupamentos em relação aos setores da coleta seletiva de lixo, Uberlândia, 2020.	14
4.4	Estatísticas descritivas e teste de normalidade de Anderson-Darling para o quantitativo per capita de resíduos da coleta seletiva, em Uberlândia, no ano de 2020.	15
4.5	Grupos formados na análise de agrupamentos dos meses da coleta de lixo seletiva, Uberlândia, 2020.	17
4.6	Estatísticas descritivas e teste de normalidade de Anderson-Darling para o quantitativo per capita de resíduos da coleta seletiva, em Uberlândia, no ano de 2020.	17

1. INTRODUÇÃO

A sustentabilidade urbana é um dos maiores desafios enfrentados pelos municípios brasileiros, principalmente, quando se trata da gestão dos resíduos sólidos.

O lixo urbano, atualmente, é um grande desafio para os governantes dos municípios, tendo em vista as leis ambientais que versam sobre o assunto tanto a nível municipal quanto no nível estadual ou federal. Observa-se ainda aumento crescente da produção de resíduos sólidos nas cidades em função do crescimento populacional nas áreas urbanas e ao uso, por parte da população, de produtos industrializados que tendem a gerar maior volume desse tipo resíduo [21].

A questão dos resíduos sólidos tem se mostrado uma preocupação evidente, principalmente a partir da década de 1990 em que os temas ambientais tomaram impulso no Brasil e no mundo. Em 2012, o lixo foi visto como um problema ambiental por 7 em cada 25 brasileiros [6].

Os perigos sanitários somados a preocupação com a saúde pública, em consequência do acúmulo de lixo, tem incrementado a preocupação com a proteção do meio ambiente e com a reutilização dos resíduos. Devido a isso, as administrações municipais são constantemente pressionadas a dar maior ênfase na atenção ao manejo dos resíduos sólidos.

A Coleta Seletiva é uma etapa fundamental do processo de reciclagem, pois por meio disso os resíduos recicláveis, como plástico, papel, metal, vidro e orgânico são separados dos demais objetos.

Em Uberlândia, o programa oficial de Coleta Seletiva teve início em 2011 com o objetivo de incentivar os catadores, por consequência, incentivar estruturas que integram e aumentam a produtividade dos indivíduos envolvidos.

No final do ano de 2020, o serviço de Coleta Seletiva em Uberlândia estava atendendo 45 setores em dias específicos [8][9][10]. Nessa coleta oficial, deve-se colocar os materiais recicláveis na porta da casa para que a coleta seja feita em um caminhão e os resíduos destinados às cooperativas. Entretanto, esse serviço urbano vem sendo implantado de forma gradativa na cidade e novos setores da cidade tem sido adicionados ao programa ao longo do tempo. Cabe à população da cidade que não é atendida por esse tipo de coleta, direcionar o material diretamente aos Ecopontos. Vale lembrar também que ocorre a coleta seletiva na zona rural desde 2013 em que a prefeitura se encarregou da estrutura de armazenamento e a comunidade ficou responsável pela construção do abrigo.

Em relação as informações da coleta seletiva no Brasil, a Associação Brasileira de Empresas de Limpeza Pública e Resíduos Especiais (ABRELPE) criou o Panorama dos Resíduos Sólidos,

o qual deixa evidente que poucos brasileiros possuem acesso aos serviços regulares de coleta porta a porta, tendo em vista que menos de 0,08% da população do Brasil usufrui desse tipo de coleta. Dessa forma, o Brasil ocupa uma posição muito inferior ao comparar com os países que possuem o mesmo nível de renda, de acordo com o presidente da corporação, Carlos Silva Filho [2].

Ainda de acordo com o presidente da ABRELPE, os fatores que influenciam nessa posição é a falta de consciência em relação a importância do manejo correto de resíduos sólidos para salvaguardar o meio ambiente e para prevenir doenças. A maior parte da sociedade e do Poder Público não possui clareza sobre a necessidade desse serviço, e o fator mais grave é que a coleta de resíduos sólidos, por ser conduzido pela esfera municipal e os municípios não possuem renda suficiente em virtude das dívidas, não possuem recursos para custear o serviço da coleta como um todo.

Sendo assim, a ABRELPE enfatiza que a universalização da coleta seletiva é um serviço a ser solucionado há longo prazo. Porquanto, ao passo que a pauta global é a economia circular e alternativas eficazes de destinação dos resíduos, no Brasil há o registro de lixões em todas as regiões e precisa-se solucionar o comportamento problemático e contrário da sociedade brasileira a efetivação da coleta seletiva dos resíduos sólidos, pois não há separação correta dentro de casa, por exemplo.

Enfim, a coleta seletiva de resíduos sólidos, além de um processo de suma importância para o meio ambiente, é também uma grande contribuição social aos catadores que dependem desse insumo para sua geração de renda e, conseqüentemente, para sua sobrevivência e também deve-se destacar sua influência na saúde coletiva da população.

A ideia da coleta seletiva é fazer com que o lixo devidamente separado deixe de ser lixo, proporcionando a geração de renda para os recicladores, por meio do seu processamento e da comercialização e, por consequência, geram benefícios ambientais, sociais, econômicos e de saúde pública para os municípios.

A cidade de Uberlândia é um município com uma população urbana expressiva. No estado de Minas Gerais ocupa a segunda posição em população, com população estimada de 699.097 habitantes no ano de 2020 [11]. Esses números indicam que a cidade se destaca a nível nacional em relação à população, sendo essa população predominantemente urbana.

Cidades com grandes populações urbanas tendem a produzir um volume considerável de resíduos sólido e, boa parte desse volume poderia ser destinado a reciclagem, reduzindo problemas ambientais e aumentando a vida útil dos aterros sanitários. A princípio, os aterros sanitários deveriam ser utilizados apenas para a destinação dos resíduos não recicláveis, mas, na prática, isso não ocorre.

Portanto, é necessário a implementação de políticas públicas que incentivem a população a destinar de forma correta os resíduos produzidos em suas residências e comércios. Estudos científicos sobre o comportamento da população em relação a destinação de resíduos sólidos recicláveis, são importantes ferramentas para subsidiar essas políticas públicas. As análises estatísticas de dados tem se destacado de forma relevante para proporcionar maior segurança as

conclusões científicas nas mais diversas áreas da ciência e, não obstante, exerce papel relevante também na área ambiental.

1.1 OBJETIVO

Nesse estudo o objetivo geral é analisar o comportamento da coleta oficial per capita de resíduos sólidos recicláveis, em setores de Uberlândia-MG, nos meses de janeiro a dezembro do ano de 2020, utilizando a metodologia da análise estatística univariada (análise descritiva e análise de variância) e da análise multivariada (análise de cluster).

Os objetivos específicos são: verificar se há diferença significativa na coleta per capita entre os setores e também entre os meses do ano; Obter agrupamentos dos setores da cidade de Uberlândia, para classifica-los de acordo com sua semelhança em relação a produção do lixo reciclado.

2. FUNdAMEntação TEÓRICA

A análise de agrupamentos é conhecida por outros nomes, como, análise de cluster ou análise de conglomerados, e dependendo da área de estudo em que é aplicada possui ainda outras denominações.- um método estatístico que permite agrupar elementos, indivíduos, produtos e até mesmo comportamentos de elementos de uma amostra, com base nas similaridades e diferenças das características que estes itens possuem [7].

A análise de cluster é um método usado para classificar objetos ou circunstância em grupos relativamente semelhantes chamados de clusters. Os objetos em cada cluster tendem a ser semelhantes entre si, mas diferentes de objetos em outros clusters, não havendo qualquer informação, à argumento, sobre a estrutura ou comportamento dos dados [14].

A Análise de Cluster Hierárquico interliga os dados por suas associações, produzindo um dendograma onde os dados semelhantes, são agrupadas entre si. Logo temos que a interpretação do dendograma quanto menor a distância entre os pontos, maior a semelhança entre os dados. [16].

Logo temos medidas que podem ser utilizadas como medidas de similaridade ou dissimilaridade, como: distância Euclidiana, distância de Manhattan, Dissimilaridade baseada na correlação de Pearson.

Genericamente, a análise de agrupamento compreende cinco etapas [1]:

1. A seleção de elementos ou de uma amostra de elementos a serem agrupados;
2. A definição de um conjunto de variáveis a partir das quais serão obtidas informações necessárias ao agrupamento dos elementos;
3. A definição de uma medida de semelhança ou distância entre os elementos;
4. A escolha de um algoritmo estatístico de partição/classificação;
5. Por último, a validação dos resultados encontrados.

2.1 DISSIMILARIDaDE BASEADa na CORRELAÇÃO DE PEARSON

Segundo Kauffman [12], dissimilaridade é a medida de diferença entre dois objetos. Existem várias maneiras várias maneiras possíveis de se encontrar essa medida. Entre muitas que são conhecidas, às que mais chamam atenção são a distância Euclidiana e a distância de Manhattan, que podem e são amplamente utilizadas na literatura, mas medidas de dissimilaridade baseadas no Coeficiente de Correlação de Pearson são também muito úteis quando o objetivo

é o agrupamento de dados, pois ele gera o nível de relacionamento entre duas variáveis. Temos que a equação da correlação Pearson dada por:

$$R(x, y) = \frac{\sqrt{\sum_{i=1}^p (x_i - \mu_x)(y_i - \mu_y)}}{(\sum_{i=1}^p (x_i - \mu_x)^2)(\sum_{i=1}^p (y_i - \mu_y)^2)}$$

em que μ_x e μ_y são as médias amostrais de X e Y .

As medidas de Dissimilaridade baseadas no Coeficiente de correlação de Pearson são (Kauffman, 1990):

$$d(x, y) = \frac{(1-R(x,y))}{2}$$

$$d(x, y) = 1 - |R(x, y)|$$

onde $d(x, y)$ é a medida de dissimilaridade entre os dados x e y .

2.2 DISTÂNCIA EUCLIDIANA

A distância euclidiana, com certeza, é a mais utilizada quando se aplica as medidas de distâncias. Segundo Mingoti [15], temos que a distância entre dois indivíduos a partir da raiz quadrada do somatório dos quadrados das diferenças entre os valores de i e j para todas as variáveis ($v=1, 2, \dots, p$), sendo definida pela seguinte fórmula:

$$d_{ij} = \sqrt{\sum_{v=1}^p (x_{iv} - x_{jv})^2}$$

X_{iv} representa a característica do indivíduo i ;

X_{jv} representa a característica do indivíduo j ;

p é o número de parcelas na amostra;

v é o número de indivíduos na amostra;

2.3 DISTÂNCIA MANHATTAN

A distância de Manhattan é uma simplificação da distância Euclidiana, e, por isso, é uma medida mais simples e de fácil implementação. É mais eficiente para aplicações em tempo real devido a sua simplicidade [13]. A Distância Manhattan, temos que entre dois (i e j) pontos num espaço euclidiano com um sistema cartesiano de coordenadas fixo como a soma dos comprimentos da projeção da linha que une os pontos com os eixos das coordenadas.

$$d_{ij} = \sum_{k=1}^p |x_{ik} - x_{jk}| = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|$$

2.4 DISTÂNCIA DE MahalanOBIS

Segundo Mahalanobis, a distância foi baseada nas correlações entre variáveis com os quais distintos padrões podem ser identificados e analisados. É uma estatística útil para determinar a similaridade entre uma amostra desconhecida e uma conhecida. Distingue-se da distância euclidiana já que tem em conta as correlações do conjunto de dados e é invariante à escala, ou seja, não depende da escala das medições. Os dados em ambos os grupos a comparar deverão ter o mesmo número de variáveis (ou seja, o mesmo número de colunas) mas não necessariamente o mesmo número de elementos (o número de linhas pode ser diferente).

A distância de Mahalanobis entre os grupos i e j é usualmente estimada segundo Rao [17] por:

$$D_{ij}^2 = (\bar{X}_i - \bar{X}_j)' \Sigma^{-1} (\bar{X}_i - \bar{X}_j)$$

\bar{X}_i é o vetor de médias do i -ésimo grupo;

\bar{X}_j é o vetor de médias do j -ésimo grupo;

Σ é a estimativa combinada da matriz da covariância/variância dentro dos grupos.

2.5 MÉTODO DE Agrupamento Hierárquico

O método hierárquico de Cluster [5], consiste em uma série de sucessivos agrupamentos ou sucessivas divisões de elementos, onde os elementos são agregados ou desagregados. Os métodos hierárquicos são subdivididos em métodos aglomerativos e divisivos. Os grupos, nos métodos hierárquicos, são geralmente representados por um diagrama bi-dimensional chamado de dendograma ou diagrama de árvore. Neste diagrama, cada ramo representa um elemento, enquanto a raiz representa o agrupamento de todos os elementos. A Figura 1 demonstra um exemplo de dendograma.

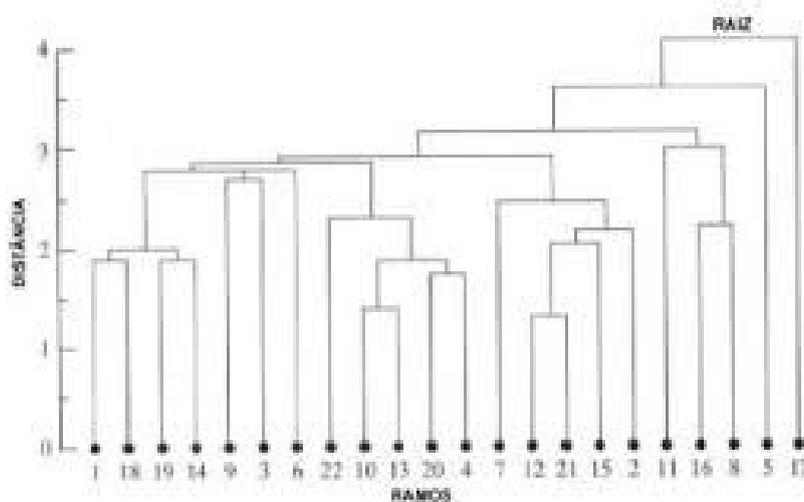


Figura 2.1: Representação Ilustrativa de um Dendograma

De acordo com [15], no método aglomerativo, cada elemento inicia-se representando um grupo, e a cada passo, um grupo ou elemento é ligado a outro de acordo com sua similaridade, até o último passo, onde é formado um grupo único com todos os elementos.

Existe uma variedade de métodos aglomerativos, que são caracterizados de acordo com o critério utilizado para definir as distâncias entre grupos. Entretanto, a maioria dos métodos são basicamente formulações alternativas de três grandes conceitos de agrupamento aglomerativo [3]:

- 1) Métodos de ligação single linkage (ligação simples), complete linkage (ligação completa), average linkage (ligação das médias), median linkage (ligação de medianas);
- 2) Método de centróide;
- 3) Métodos de minimização da soma de erros quadráticos ou variâncias (método de Ward).

2.6 DEFINIÇÃO DE QUANTIDADE DE GRUPOS

Determinar a quantidade de grupos é uma das partes mais importante para o desenvolvimento do processo de agrupamento.

Para Barroso [4], o número de grupos pode ser definido a priori, através de algum conhecimento que se tenha sobre os dados, pela conveniência do pesquisador, por simplicidade, ou ainda pode ser definido a posteriori com base nos resultados da análise.

De acordo com Aaker et al.[1], para determinar o número apropriado de grupos, existem diversas abordagens possíveis: (i), o pesquisador pode especificar antecipadamente o número de grupos (clusters).(ii) o pesquisador pode estimar o número de grupos a partir do uso de um método de agrupamento hierárquico. (iii) outra abordagem é representar, graficamente, a razão entre a variância total interna dos grupos e a variância entre os grupos, em relação ao número de grupos formados.

3. METODOLOGIA

A variável a ser analisada neste estudo é a produção per capita (kg/mês) de resíduos sólidos recicláveis que são coletados e gerenciados pelo Departamento Municipal de Água e Esgoto (DMAE) na cidade de Uberlândia-MG. Os dados foram apresentados pelo DMAE em montante total coletado por mês em cada setor. Para a obtenção da estimativa do montante per capita por mês, os valores informados pela empresa foram divididos pelo total da população estimada em cada setor.

Em uma primeira verificação tinha-se, em 2020, 30 setores atendidos pela coleta seletiva e que apresentam a estimativa populacional necessária para o cálculo da produção per capita. Além desses Bairros existem setores específicos de coleta, como por exemplo o Hospital de Clínicas da Universidade Federal de Uberlândia, que não foram incluídos no estudo, em função de não se ter a possibilidade de se obter a produção per capita.

Foram selecionados, para esta pesquisa, apenas os setores que apresentavam consistência de coleta de resíduos sólidos recicláveis ao longo do ano de 2020, ou seja, aqueles setores em que havia registro de coleta seletiva em todos os meses do ano. Esta organização dos dados gerou uma planilha (Lista 3.1) de informações contendo 19 setores e 12 meses de coleta.

A análise descritiva inicial foi realizada com a finalidade de explorar as informações gerais da produção de resíduos sólidos recicláveis na cidade de Uberlândia. Esta análise foi feita por meio de cálculos de medidas de posição (Média e Mediana) e de dispersão (Desvio Padrão e Coeficiente Variação) tanto para os setores quanto para os meses do ano.

Em uma segunda etapa de análise os dados foram submetidos à análise multivariada por meio da técnica de análise de cluster para a distinção de agrupamentos de setores e ou agrupamentos de meses em função desse procedimento. Nesta etapa, quando o objetivo for distinguir os grupos de setores, estes serão considerados os objetos de agrupamento e as produções mensais serão as variáveis de agrupamentos e, no caso de agrupamentos dos meses estes serão considerados os objetos de agrupamentos e os setores as variáveis de agrupamentos.

A seleção da medida de similaridades assim como a método aglomerativo foi realizado em função das características da variável em análise e baseado em procedimentos descritos em Mingoti [15] e Moita [16].

Todos os procedimentos das análises estatísticas dos dados foram executados no software R. [22], tendo como base os procedimentos descritos em Ritter [18].

Setores participantes da análise estatística

- Brasil/Aparecida
- Custódio Pereira
- Alto Umuarama
- Cazeca
- Saraiva
- Tibery
- Roosevelt
- Centro
- Patrimônio
- Vigilato Pereira
- Jaraguá
- Luizote/Dom Z./Mansour/Jardim Patricia
- Osvaldo R./Martins/Bom Jesus/Daniel F.
- Santa Mônica/Segismundo
- Cidade Jardim
- Fundinho/Tabajaras
- Jardim Karaiba
- Lídice
- Umuarama

4. RESULTADos

4.1 AnÁLISE DESCRITIVA

Na Tabela 4.1 são apresentadas as principais estatística do quantitativo per capita (kg/mês) de resíduos sólidos residenciais coletados pelo programa do DMAE de coleta seletiva, nos setores atendidos pelo programa, no ano de 2020, na cidade de Uberlândia- MG. E na Tabela 4.2 são apresentadas as estatísticas referentes por mês do quantitativo per capita (kg/setor) no ano de 2020.

Tabela 4.1: Estatísticas descritivas do quantitativo per capita (kg/mês) da coleta seletiva em Uberlândia-MG, por setor 2020.

Setor	Estatísticas					
	Média	Mediana	DP	CV(%)	Mín	Máx
Alto Umuarama	0.864	0.895	0.173	20.080	0.556	1.058
Brasil/Aparecida	0.468	0.455	0.092	19.742	0.345	0.600
Cazeca	1.083	1.066	0.162	14.999	0.815	1.334
Centro	2.311	2.370	0.274	11.852	1.719	2.671
Cidade Jardim	1.475	1.483	0.374	25.383	0.838	2.196
Custódio Pereira	0.384	0.384	0.067	17.479	0.276	0.486
Fundinho/Tabajaras	1.282	1.262	0.137	10.721	1.121	1.540
Jaraguá	0.431	0.427	0.130	30.232	0.247	0.742
Jardim Karaiba	2.188	2.211	0.485	22.179	1.549	3.099
Lídice	1.709	1.705	0.223	13.022	1.423	2.237
Luizote/Dom Z./Mansuor/ Jardim P.	0.526	0.539	0.044	8.322	0.426	0.585
Oswaldo R./Martins/Bom J./ Daniel F.	0.512	0.502	0.061	11.895	0.382	0.612
Patrimônio	0.852	0.837	0.121	14.206	0.724	1.041
Santa Mônica/Segismundo	0.456	0.449	0.064	14.007	0.370	0.558
Saraiva	0.556	0.549	0.096	17.359	0.419	0.762
Tibery	0.532	0.526	0.071	13.303	0.430	0.650
Umuarama	1.535	1.552	0.211	13.775	1.106	1.740
Vigilato Pereira	1.207	1.190	0.243	20.168	0.786	1.620
Roosevelt	0.425	0.404	0.075	17.559	0.333	0.552

DP=Desvio Padrão/**CV**= Coeficiente de Variação/ **Mín**= Mínimo/**Máx**= Máximo

Analisando os resultados obtidos, foi observado por meio da Tabela 4.1 , que os valores da média e de mediana para cada setor estão relativamente próximos. Classifica-se as distribuições da coleta seletiva nos setores sendo aproximadamente simétricas. Em relação ao coeficiente

variação (CV), observa-se pela Tabela 4.1, que a variabilidade avaliada pelo CV, pode ser considerada baixa (dados homogêneos), indicando que ocorreu pequena dispersão na quantidade coletada de resíduos sólidos recicláveis entre os setores de Uberlândia. Pode-se observar na Tabela 4.1 que a variável Jaraguá apresentou os maior valor de coeficiente de variação, o que significa que houve uma dispersão alta na quantidade coletada de resíduos sólidos. Já a variável Luizote/Dom Z./Mansour/Jardim Patricia, obteve o menor coeficiente de variação, o que indica uma dispersão baixa na quantidade coletada de resíduos sólidos.

Tabela 4.2: Estatísticas descritivas gerais, por mês, do quantitativo per capita (kg/setor) da coleta seletiva nos setores de Uberlândia-MG, no ano 2020.

Estatísticas						
Meses	Média	Mediana	Desvio Padrão	Coeficiente de Var.(%)	Mínimo	Máximo
Jan	1.204	0.853	0.617	60.275	0.396	2.385
Fev	0.824	0.724	0.512	62.192	0.359	2.074
Mar	0.944	0.627	0.562	59.555	0.302	2.224
Abr	0.898	0.912	0.538	59.886	0.345	2.098
Mai	0.891	0.724	0.515	57.822	0.380	2.052
Jun	0.950	0.650	0.597	62.868	0.276	2.176
Jul	1.052	0.815	0.649	61.626	0.416	2.623
Ago	0.963	0.724	0.658	68.339	0.247	2.486
Set	1.092	0.724	0.725	66.412	0.334	2.671
Out	1.064	0.884	0.702	66.018	0.455	2.881
Nov	0.990	0.837	0.648	65.450	0.304	2.479
Dez	1.177	1.018	0.848	72.055	0.309	3.099

Pode-se observar que o mês de dezembro se destaca com a maior média e desvio padrão de coleta seletiva, também obteve-se o maior coeficiente de variação (CV). Analisando os valores obtidos, nota-se por meio da Tabela 4.2, que a variabilidade avaliada pelo CV, pode ser considerada alta (dados heterogêneos), indicando que ocorreu grande dispersão na quantidade coletada de resíduos sólidos recicláveis entre os setores de Uberlândia em cada mês do ano.

4.2 MÉTODO AGLOMERATIVO

Para os agrupamentos hierárquico pode-se utilizar uma medida bastante comum, que é a correlação cofenética na seleção do método de aglomeração [19]. O coeficiente de correlação cofenética mede o grau de preservação das distancias emparelhadas pelo dendograma resultante do agrupamento em relação às distancias originais [20].

Na análise prévia dos clusters para a seleção do método de aglomeração, verificou-se que o método de Ward apresentou o coeficiente de correlação cofenético de 0,7641, sendo superior aos métodos de ligação simples e de ligação média. Portanto, o agrupamentos foram realizados, neste estudo, por meio da distância euclidiana com o método de aglomeração de Ward.

4.3 ANÁLISE DE AGRUPAMENTO DOS SETORES

Na Figura 4.1 é possível visualizar que na altura de corte de aproximadamente 30 é possível distinguir dois grupos referente aos setores da coleta de lixo seletiva na cidade de Uberlândia, no ano de 2020. Na Tabela 4.3 são apresentados os setores que compõem os dois grupos.

Cabe ressaltar que os setores da coleta seletiva são definidos pelos responsáveis pelo procedimento na cidade de Uberlândia e esses setores, conforme pode ser visualizado, podem ser composto de mais de um setor.

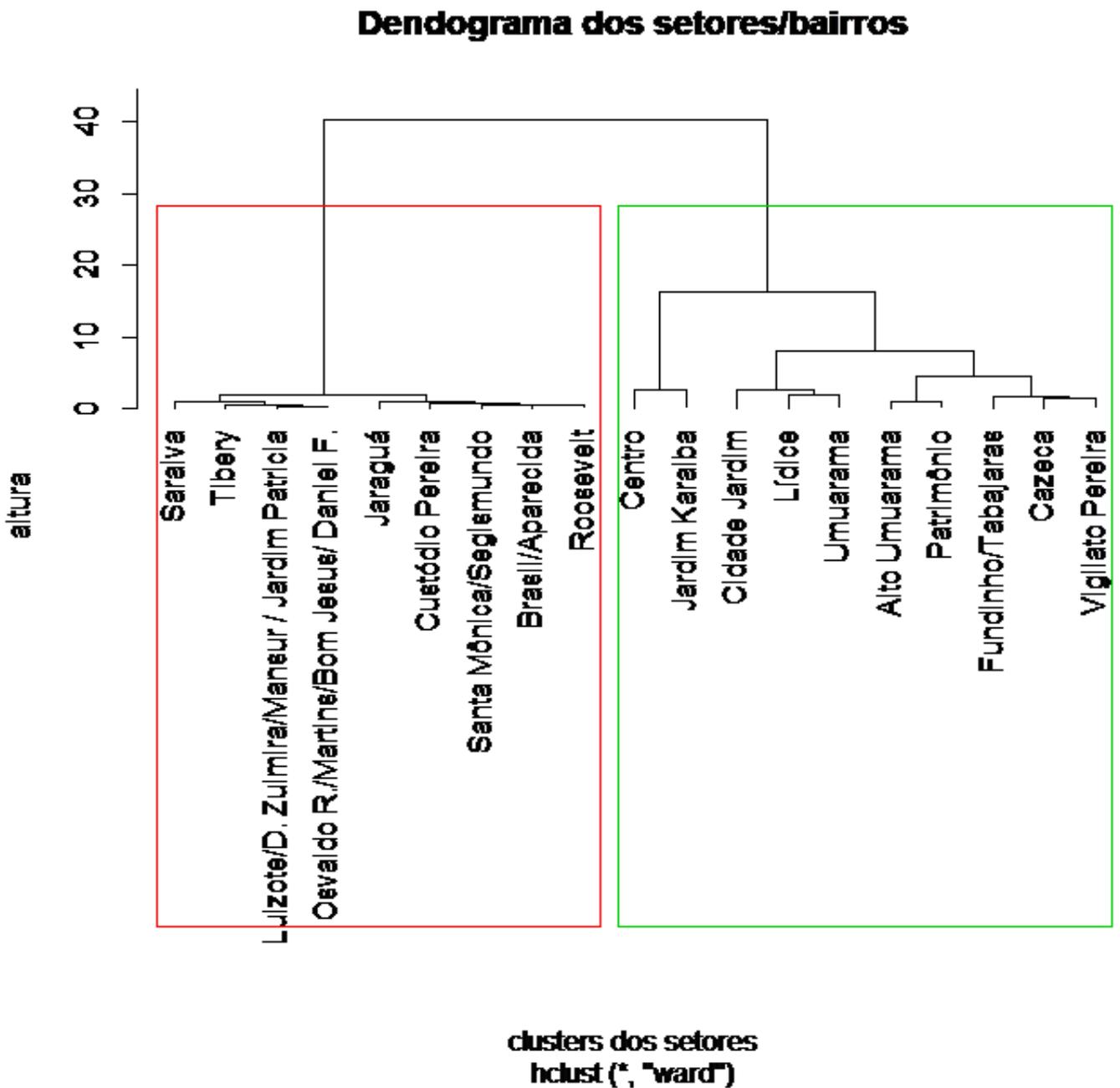


Figura 4.1: Dendograma com 2 grupos referente aos setores atendidos pela coleta seletiva de lixo, na cidade de Uberlândia, no ano de 2020.

Tabela 4.3: Grupos formados na análise de agrupamentos em relação aos setores da coleta seletiva de lixo, Uberlândia, 2020.

Grupo 1	Habitantes
Brasil/Aparecida	24.091
Custódio Pereira	9.551
Jaraguá	8.083
Luizote/ Dona Zulmira / Mansour / Jardim Patricia	38.614
Oswaldo Rezende / Martins / Bom Jesus / Daniel Fonseca	36.625
Santa Mônica / Segismundo	54.274
Saraiva	10.019
Tibery	18.631
Roosevelt	20.724
Total	220.612
Grupo 2	Habitantes
Alto Umuarama	IN
Cazeca	3.202
Centro	7.262
Cidade Jardim	7.378
Fundinho / Tabajaras	9.543
Jardim Karaiba	3.098
Lídice	4.180
Patrimônio	4.420
Umuarama	3.736
Vigilado Pereira	5.000
Total	47.819

O Grupo 1 foi composto por nove setores de coleta. Neste grupo estão presentes os setores mais populosos de Uberlândia, que recebem a coleta seletivo de lixo. Ele é formado pelos setores: Santa Mônica/Segismundo com aproximadamente 54.274 habitantes; Luizote/Dona Zulmira/Mansour/Jardim Patricia com aproximadamente 38.614 habitantes; Oswaldo Resende./Martins/Bom Jesus/ Daniel Fonseca com aproximadamente 36.625 habitantes. Apenas estes setores correspondem a 58,7% dos habitantes desse grupo que é, no total, de 220.612 habitantes. Ainda constam desse grupo os setores : Custódio Pereira, Jaraguá, Brasil/Aparecida , Saraiva , Tibery , Roosevelt.

O Grupo 2 foi composto de dez setores, sendo que nele, estão presentes os setores menos populosos de Uberlândia, que recebem a coleta seletiva de lixo. Neste grupo, os setores mais populosos são: o Fundinho/Tabajaras com aproximadamente 9.543 habitantes, o segundo maior é o Cidade Jardim com aproximadamente 7.378 habitantes, o terceiro maior é o Centro com aproximadamente 7.262 habitantes. Eles correspondem a 50,6% dos dados, dos 47.819 habitantes do grupo 2. Fazem parte ainda deste grupo os setores: Alto Umuarama , Cazeca , Jardim Karaiba , Lídice , Patrimônio , Umuarama , Vigilato Pereira.

Na Tabela 4.4 são apresentadas as estatísticas gerias dos dois grupos de setores que foram selecionados no dendograma.

Os resultados da tabela 4.4 indicam que os dois grupos apresentaram distribuição normal para o quantitativo de recicláveis da coleta seletiva. Na aplicação da estatística F para a

4.4 ANÁLISE DE AGRUPAMENTO DOS MESES

O dendograma dos meses do ano de 2020 para o quantitativo da coleta seletiva em Uberlândia, considerando-se os quantitativos per capita dos setores com sendo as variáveis de agrupamento, é apresentado na Figura 4.3 .

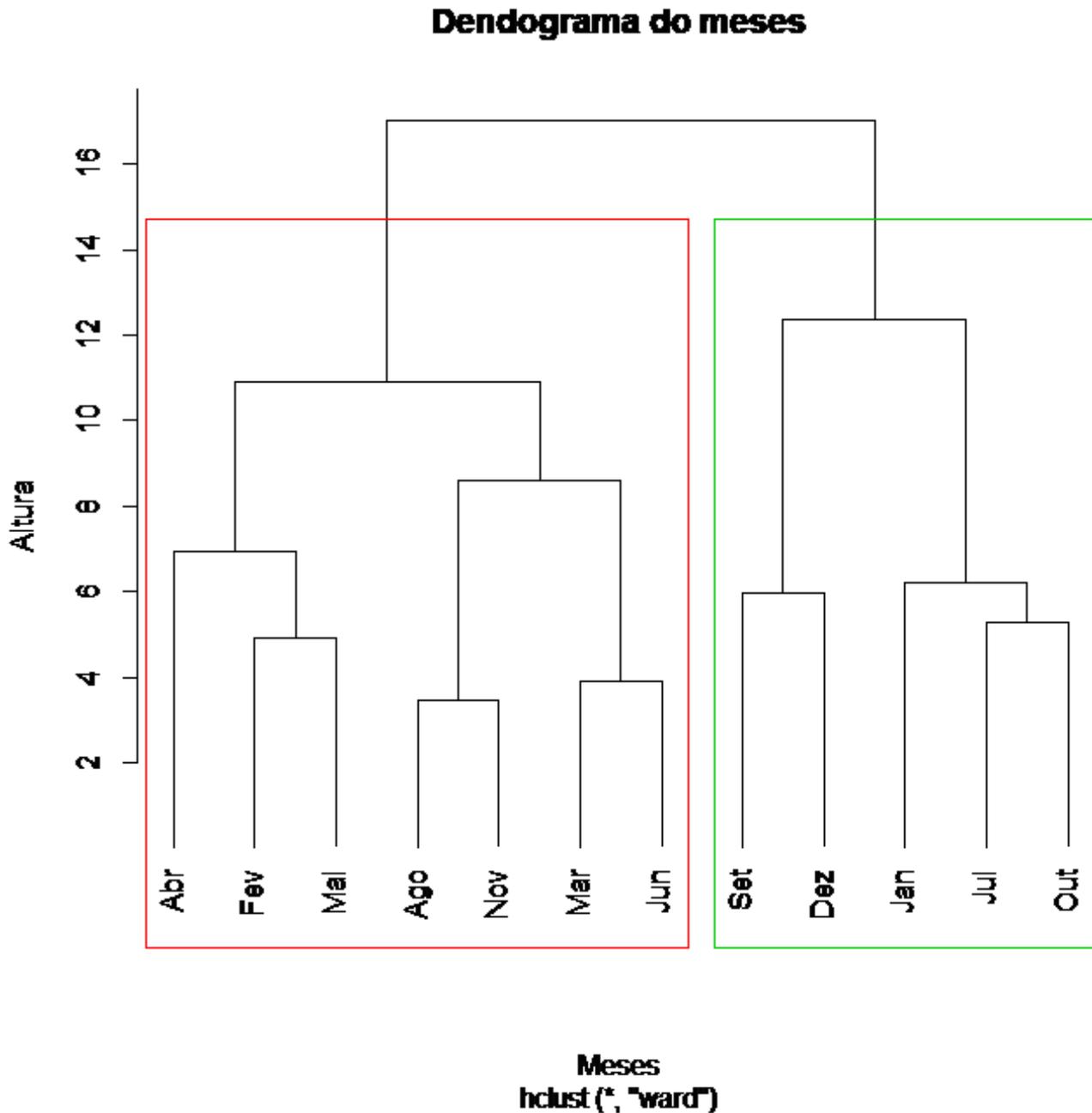


Figura 4.3: Dendograma com 2 grupos referente aos meses coleta seletiva de lixo, na cidade de Uberlândia, no ano de 2020.

Na figura 4.3 é possível ver que, na altura de corte de aproximadamente 15, separa-se a dois grupos referente aos meses da coleta de lixo seletiva, na cidade de Uberlândia, no ano de 2020.

Na Tabela 4.5 são apresentados os meses que compõem cada grupo.

Tabela 4.5: Grupos formados na análise de agrupamentos dos meses da coleta de lixo seletiva, Uberlândia, 2020.

Grupo 1	Grupo 2
Janeiro	Feveiro
Julho	Março
Setembro	Abril
Outubro	Maior
Dezembro	Junho
	Agosto
	Novembro

Nota-se que no grupo 1 tem-se a menor concentração de meses do ano de 2020, entretanto pode-se ver que com cinco meses o grupo 1 tem aproximadamente os mesmos quantitativo de materiais recicláveis coletados pelo DMAE que é aproximadamente 1.145.557 Kg. No grupo 2 é onde se encontra o maior número de meses do ano de 2020, entretanto com sete meses o grupo 2 tem aproximadamente 1.189.738,37 Kg de materiais recicláveis coletados pelo DMAE.

É interessante também observar que meses festivos como dezembro e janeiro (festas de fim de ano) e outubro (mês das crianças) encontram-se no grupo 1, ou seja, espera-se que estes meses tenham grande produção de materiais recicláveis nas residências, impactando no quantitativo da coleta seletiva. Ao se observar as médias da tabela 4.2, nota-se que estes meses apresentaram quantitativo maiores.

Na Tabela 4.6 são apresentadas as estatísticas descritivas para os dois grupos de meses extraídos do dendograma.

Tabela 4.6: Estatísticas descritivas e teste de normalidade de Anderson-Darling para o quantitativo per capita de resíduos da coleta seletiva, em Uberlândia, no ano de 2020.

Grupo	Mínimo	Média	Máximo	Desvio Padrão	p (AD)
1	1.0242	1.0818	1.7770	0.0585	0.4996
2	0.8237	0.9229	0.9905	0.0560	0.4178

Os resultados da Tabela 4.6 indicam normalidade para os dados dos grupos e, portanto, pode-se aplicar o teste t para comparação das médias dos grupos. Neste processo verificou-se, pelo teste F, que as variâncias dos dois grupos são homocedásticas (valor-p = 0,8786) e que os quantitativos médios do grupo 1 (1,0818 kg/mês) foi superior ao do grupo 2 (0,9229 kg) com valor-p = 0,00076. Desta forma, tem-se que em média nos meses do grupo 1, a população apresenta maior quantitativo de recicláveis.

5. Conclusões

O presente estudo permitiu concluir que a Análise de Cluster com método hierárquico mostrou-se eficaz para fazer a distinção entre a coleta per capita de resíduos sólidos recicláveis entre os setores do município. Foram identificados dois grupos em relação as variáveis em estudo, o grupo 1 é composto por nove setores e o grupo 2 com dez setores.

Nota-se que o grupo 1 é composto pelos maiores setores de Uberlândia. Foi possível verificar que dentro desses 2 grupos, o grupo 1 foi o grupo em que apresentou a média de quantitativo per capita da coleta seletiva superior ao grupo 2.

Em relação aos meses foi feita a mesma análise de cluster, onde foram identificados dois grupos sendo o grupo 1 composto por cinco meses e o grupo 2 com sete meses. É interessante observar que os meses festivos como dezembro, janeiro e outubro encontram-se no grupo 1, e que estes são os meses em que a população apresenta maior quantitativo recicláveis.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AAKER, D. A.; KUMAR, V. D. G. S.: *Pesquisa de marketing*. Atlas, 745p ed., 2001.
- [2] Abrelpe: *Panorama dos resíduos sólidos no Brasil 2018-2019.*, 2019. <http://abrelpe.org.br/download-panorama-2018-2019/>, acessado em 19/04/2021.
- [3] ANDERBERG, R.: *Cluster Analysis for Applications*. 1973.
- [4] BARROSO, L. P., A.: *R. Análise de Multivariada*. UFLA, 2003.
- [5] BUSSAB, Wilton de Oliveira; MIAZAKI, E. S. A. D. F.: *Introducao a analise de agrupamentos*. Ime-Usp, 1990.
- [6] CONKE, L . S; NASCIMENTO, E . P.: *A coleta seletiva nas pesquisas brasileiras: uma avaliação metodológica*. Revista Brasileira Gestão Urbana, 10(1), 2018. https://www.scielo.br/scielo.php?script=sci_arttext&pid=S2175-33692018000100199&lng=pt&tlng=pt, acessado em 15/04/2021.
- [7] CORRAR, Luiz; PAULO, E. D. F. J. M.: *Análise multivariada para os cursos de administração, ciências contábeis e economia*. 2007.
- [8] DMAE: *Panorama dos resíduos sólidos no Brasil 2018-2019.*, 2021a. [https://www.uberlandia.mg.gov.br/prefeitura/orgaos-municipais/dmae/servicos-dmae/coleta-de-residuos/#:~:text=Coleta%20Seletiva%20\(Zona%20Urbana\)&text=0s%20servi%C3%A7os%20de%20coleta%20seletiva,frente%20sua%20resid%C3%Aancia%20ou%20com%C3%A9rcio](https://www.uberlandia.mg.gov.br/prefeitura/orgaos-municipais/dmae/servicos-dmae/coleta-de-residuos/#:~:text=Coleta%20Seletiva%20(Zona%20Urbana)&text=0s%20servi%C3%A7os%20de%20coleta%20seletiva,frente%20sua%20resid%C3%Aancia%20ou%20com%C3%A9rcio), acessado em 15/04/2021.
- [9] DMAEb: *Coleta seletiva reciclou mais de 1,3 mil toneladas de vidro em 2020*, 2021b. <https://www.uberlandia.mg.gov.br/2021/03/03/coleta-seletiva-reciclou-mais-de-13-mil-toneladas-de-vidro-em-2020/#:~:text=DMAE%20%23Reciclagem%20%23Vidro>, acessado em 15/04/2021.
- [10] DMAEc: *Coleta Seletiva*, 2021c. <https://www.uberlandia.mg.gov.br/prefeitura/orgaos-municipais/dmae/servicos-dmae/residuos-solidos/coleta-seletiva/>, acessado em 15/04/2021.
- [11] IBGE: *Instituto Brasileiro de Geografia e Estatística*, 2020. <https://www.ibge.gov.br/cidades-e-estados/mg/uberlandia.html>, acessado em 17/04/2021.

- [12] KAUFMANN, L., R. P. J.: *Finding groups in data: an introduction to cluster analysis*. New York: John Wiley, 1990.
- [13] KUGLER, M. JÚNIOR, J. T. . L. H. S.: *Desenvolvimento de uma Rede Neural LVQ em Linguagem VHDL para Aplicações em Tempo-Real*. proceedings of the VI Brazilian Conference on Neural Networks - VI Congresso Brasileiro de Redes Neurais, p. p. 103–108, 2003. <https://agenciabrasil.ebc.com.br/geral/noticia/2019-11/brasil-gera-79-milhoes-de-toneladas-de-residuos-solidos-por-ano>.
- [14] Malhotra, k.: *Pesquisa de Marketing Uma Orientação Aplicada*. Bookman, 2001.
- [15] MINGOTI, S. A.: *Análise de Dados Através de Métodos de Estatística Multivariada: uma abordagem aplicada*. Editora UFMG, 2005.
- [16] MOITA NETO, J. M.; MOITA, G. C.: *UMA INTRODUÇÃO À ANÁLISE EXPLORATÓRIA DE DADOS MULTIVARIADOS*. Universidade Federal do Piauí - 64.049-550 - Teresina – PI, 1998. <https://www.scielo.br/pdf/qn/v21n4/3193.pdf> acessado em 21/04/2021. , acessado em 15/08/2021.
- [17] Rao, C.R.: *Advanced statistical methods in biometric research*. 1952.
- [18] RETTIR, M . N ; THEY, N. . H.: *Introdução ao software estatístico R*. Centro de Estudos Costeiros, Limnológicos e Marinhos (Ceclimar), 2019. <https://www.lume.ufrgs.br/bitstream/handle/10183/188778/001087242.pdf?sequence=1>, acessado em 23/04/2021.
- [19] Romesburg, H. C.: *Cluster Analysis for Researchers*. Lulu Press, 2004.
- [20] Sneath, P. e Sokal, R.: *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. W H Freeman Co; 1st Edition., 1973.
- [21] SOUZA, L.: *Brasil gera 79 milhões de toneladas de resíduos sólidos por ano*. Revista Agência Brasil, 2020. <https://agenciabrasil.ebc.com.br/geral/noticia/2019-11/brasil-gera-79-milhoes-de-toneladas-de-residuos-solidos-por-ano>, acessado em 16/04/2021.
- [22] TEAM, R. C.: *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. 2021. 9. <https://www.R-project.org/>.