



Ministério da Educação
Universidade Federal de Ouro Preto
Escola de Minas
Departamento de Engenharia de Produção



ANÁLISE DE AGRUPAMENTO NO SOFTWARE MINITAB PARA OTIMIZAÇÃO DO TRANSPORTE TERRESTRE DE CARGAS LTL

ANA CAROLINA RODIANI ROLIM

Ouro Preto MG
2023

ANA CAROLINA RODIANI ROLIM

ANÁLISE DE AGRUPAMENTO NO SOFTWARE MINITAB PARA OTIMIZAÇÃO DO TRANSPORTE TERRESTRE DE CARGAS LTL

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como parte dos requisitos necessários para a obtenção do grau em Bacharel em Engenharia de Produção.

Orientador: Magno Silvério Campos

Ouro Preto - MG
26 de setembro de 2023



FOLHA DE APROVAÇÃO

Ana Carolina Rodiani Rolim

Análise de Agrupamento no software Minitab para otimização de transporte terrestres de cargas LTL

Monografia apresentada ao Curso de Engenharia de Produção da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Engenheira de Produção

Aprovada em 26 de setembro de 2023.

Membros da banca

Prof.ª Dr.ª **Magno Silvério Campos** - Orientador - Universidade Federal de Ouro Preto

Me. **Renato Fernandes Ferreira** - Examinador Convidado - Universidade Federal de Ouro Preto

Prof.ª Me. **Cristiano Luís Turbino de Franca e Silva** - Examinador Convidado - Universidade Federal de Ouro Preto

Magno Silvério Campos, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 26/09/2023.



Documento assinado eletronicamente por **Magno Silverio Campos, PROFESSOR DE MAGISTERIO SUPERIOR**, em 26/09/2023, às 20:06, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Renato Fernandes Ferreira, COORDENADOR(A) DE PROCESSOS E PROJETOS ORGANIZACIONAIS**, em 26/09/2023, às 20:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Cristiano Luis Turbino de Franca e Silva, PROFESSOR DE MAGISTERIO SUPERIOR**, em 27/09/2023, às 15:08, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0595431** e o código CRC **72EF7530**.

Este trabalho é dedicado aos meus pais, Olga e Paulo, por não medirem esforços para que eu pudesse chegar até aqui.

Agradecimentos

Cursar Engenharia de Produção foi uma jornada repleta de desafios e aprendizados. Agradeço a Deus e Nossa Senhora por me acompanharem até aqui com tantas bênçãos. Agradeço aos meus pais, Paulo e Olga, pelo amor e apoio incondicional. Vocês são minha vida! Às famílias Rodiani e Rolim por sempre estarem presentes. À Universidade Federal de Ouro Preto pelo ensino gratuito de excelência. Aos professores do DEPRO por compartilharem seus conhecimentos. Ao LABOR e CAEPRO pela oportunidade única de crescimento. À todas as amizedes que criei até aqui e me deram forças em todos os momentos. Ao Departamento de Relações Internacionais por me ajudar a realizar o meu maior sonho. À Montanuniversität Leoben por me acolher e me mostrar a vida do outro lado do mundo. Ao meu mentor, Magno, pelos conselhos em todos os âmbitos. Às empresas Usemobile, Gerdau e Fercam, minhas maiores escolas fora da universidade. À República Mexicanas pelo melhor lar que eu poderia ter durante esta jornada. Por fim, à cidade de Ouro Preto e suas tradições que para sempre terão meu coração.

Muito Obrigada!

*"The further I go
the closer to me I get."
Andrew Mccarthy*

Resumo

Um dos maiores desafios no mercado de transporte rodoviário de cargas é a tomada de decisão, que abrange desde a seleção de opções vantajosas até a definição de tarifas de frete adequadas e o gerenciamento da disponibilidade de caminhões. O foco principal deste estudo é a análise de agrupamento por meio do software Minitab visando aprimorar a tomada de decisões no transporte de cargas parciais terrestres, de forma a buscar compreender as inter-relações entre os elementos presentes em um conjunto de dados de cargas. O objetivo central é aprimorar a eficiência logística, reduzir os custos operacionais e aumentar a competitividade em um mercado intrinsecamente desafiador. Após análise dos resultados provenientes dos seis tipos de modelos, a autora concluiu que a abordagem que conduz ao maior número de agrupamentos homogêneos é aquela que se baseia no encadeamento médio, utilizando a distância euclidiana quadrática. Essa abordagem fornece aos profissionais do setor uma ferramenta para tomar decisões mais fundamentadas e eficazes ao lidar com a complexidade do transporte de cargas parciais em ambientes terrestres.

Palavras-chave: Tomada de Decisão, Análise de Agrupamento, Software Minitab, Ferramenta Estatística, Eficiência Logística, Cargas Parciais.

Abstract

One of the greatest challenges in the road freight transport market is decision-making, encompassing the selection of advantageous options to defining appropriate freight rates and managing truck availability. The main focus of this study is cluster analysis using the Minitab software, aiming to enhance decision-making in the transportation of partial loads on land, seeking to comprehend the interrelations among elements present in a dataset of cargoes. The central objective is to improve logistical efficiency, reduce operational costs, and enhance competitiveness in an inherently challenging market. After analyzing the results from the six types of models, the author concluded that the approach leading to the highest number of homogeneous clusters is the one based on average linkage, using quadratic Euclidean distance. This approach provides professionals in the field with a concrete tool for making more informed and effective decisions when dealing with the intricacies of transporting partial loads in terrestrial environments.

Keywords: Decision Making, Cluster Analysis, Minitab Software, Statistical Tool, Logistic Efficiency, Partial Loads.

Lista de Figuras

Figura 1 – Principais Métodos de Ligação	12
Figura 2 – Exemplo de Dendrograma	13
Figura 3 – Final Partition	18
Figura 4 – Dendrograma do Encadeamento Completo com a métrica de Distância Euclidiana	18
Figura 5 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Euclidiana	19
Figura 6 – Dendrograma do Encadeamento Completo com a métrica de Distância Euclidiana Quadrática	19
Figura 7 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Euclidiana Quadrática	20
Figura 8 – Encadeamento Completo com a métrica de Distância Manhattan	20
Figura 9 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Manhattan	21
Figura 10 – Dendrograma do Encadeamento Médio com a métrica de Distância Euclidiana	21
Figura 11 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância Euclidiana	22
Figura 12 – Dendrograma do Encadeamento Médio com a métrica de Distância Euclidiana Quadrática	22
Figura 13 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância Euclidiana Quadrática	23
Figura 14 – Dendrograma do Encadeamento Médio com a métrica de Distância de Manhattan	23
Figura 15 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância de Manhattan	24
Figura 16 – Agrupamentos Separados por Cores	26
Figura 17 – Cluster 1	27
Figura 18 – Cluster 2	27
Figura 19 – Cluster 3	28
Figura 20 – Cluster 4	28
Figura 21 – Cluster 5	28
Figura 22 – Cluster 6	29

Lista de tabelas

Tabela 1 – Composição dos custos de transporte	5
Tabela 2 – Modelo geral de um banco de dados para elaboração da análise de agrupamentos . . .	10
Tabela 3 – Conversão de Variáveis Categóricas de Texto para Escala Numérica	16
Tabela 4 – Comparação dos Níveis de Similaridade na Etapa Conclusiva	24

Lista de abreviaturas e siglas

3PL	Third Part Logistic
4PL	Fourth Part Logistic
ABNT	Associação Brasileira de Normas Técnicas
DEPRO	Departamento de Engenharia de Produção
EUA	Estados Unidos
FCT	Full Charter Trucks
ICMS	Imposto sobre Circulação de Mercadorias e Serviços
LLP	Lead Logistics Provider
LTL	Less-than-Truckload
PSL	Prestação de Serviços Logísticos
SM	Spot-Market
TL	Truck Load
UEA	União Europeia
UFOP	Universidade Federal de Ouro Preto

Sumário

1	INTRODUÇÃO	1
1.1	Objetivos	1
1.2	Justificativa	2
1.3	Organização do Texto	2
2	REVISÃO BIBLIOGRÁFICA	3
2.1	Transporte de Cargas	3
2.1.1	Transporte Rodoviário	3
2.2	Custos de Transporte	4
2.3	Terceirização do Transporte Rodoviário	5
2.3.1	Cargas TL e LTL	6
2.4	Estatística	6
2.4.1	Estatística na Tomada de Decisão	7
2.5	Análise de Agrupamento	8
2.6	Etapas da Análise de Agrupamento Hierárquica	9
2.6.1	Matriz de Distância ou Dissimilaridade	9
2.6.2	Método de Ligação	11
2.6.3	Dendrograma	12
2.6.4	Softwares Estatísticos	13
3	METODOLOGIA	14
3.1	Definição do Problema de Otimização	14
3.2	Elaboração do Banco de Dados	14
3.3	Configuração do Minitab	14
3.4	Seleção dos Métodos	14
3.5	Avaliação dos Agrupamentos e Proposta	15
4	DISCUSSÃO	16
4.1	Matriz de Variáveis	16
4.2	Aplicação no Minitab	17
4.2.1	Resultados do Minitab	18
4.3	Análise	21
4.3.1	Análise Final dos Clusters	25
5	CONSIDERAÇÕES FINAIS	30
5.1	Conclusão	30
5.2	Trabalhos Futuros	30
	Referências	32

Anexos	34
ANEXO A – BASE DE DADOS UTILIZADA	35

1 Introdução

O transporte de mercadorias é um componente significativo das operações logísticas. O aumento do comércio global, a concorrência acirrada entre os transportadores e o aumento das preocupações com a energia, são alguns dos motivos pelos quais as empresas estão colocando mais ênfase no gerenciamento e uso eficazes dos serviços de transporte. Em diversos setores da economia, os custos de transporte chegam a um quinto ou mesmo um quarto (madeira, petróleo, pedra, argila e produtos de vidro) do dólar médio de vendas. (SCHNEIDER, 1985 apud CHU, 2005) Ao considerar o mercado de transporte rodoviário de carga, é possível perceber sua importância para a economia de diversos países, incluindo o Brasil. Nesse cenário, as empresas que atuam nesse segmento enfrentam constantes desafios para otimizar seus processos e manter sua competitividade. Um dos principais desafios é a tomada de decisão, que, além de precisar ser tomada quase que de forma imediata, se apresenta como um processo crucial para o sucesso dessas empresas. Através da tomada de decisão, é possível selecionar as melhores alternativas para alcançar os objetivos desejados, desde o valor ideal para se cobrar o frete até a organização de disponibilidade de caminhões em cada frota.

1.1 Objetivos

O objetivo geral deste estudo consiste em analisar o uso da ferramenta estatística de análise de agrupamento no Software Minitab como uma forma de auxiliar a tomada de decisão no mercado de transporte rodoviário de cargas parciais.

Especificamente, os objetivos específicos deste estudo são realizar uma revisão abrangente da literatura sobre o transporte rodoviário de cargas parciais e da ferramenta estatística de análise de agrupamento, de forma a obter uma base sólida para a análise subsequente dos dados. Desenvolver o banco de dados manualmente gerado, incluindo informações detalhadas sobre as características das cargas parciais, tais como comprimento, acabamento externo, diâmetro externo, aplicação e destino. Esta etapa é crucial para garantir a qualidade e a consistência das informações que serão utilizadas na análise de agrupamento. Aplicar técnicas de análise de agrupamento no banco de dados no software Minitab, explorando variados métodos de encadeamento, critérios de distância e medidas de dissimilaridade. Busca-se também identificar agrupamentos de cargas parciais com características semelhantes, permitindo uma melhor compreensão das relações existentes entre os diferentes elementos presentes no conjunto de dados. Avaliar a eficácia da análise de agrupamento como uma ferramenta para auxiliar a tomada de decisão no transporte de cargas parciais terrestres. Propor recomendações e estratégias específicas para as empresas atuantes no setor de transporte rodoviário de cargas parciais nacionais, com base nos resultados obtidos por meio da análise de agrupamento. O objetivo é melhorar a eficiência logística, reduzir os custos operacionais e ampliar a competitividade das empresas no contexto de um mercado altamente desafiador e competitivo.

1.2 Justificativa

A presente pesquisa se justifica pela grande competitividade do mercado de transporte rodoviário de carga, no qual as empresas precisam tomar decisões praticamente de forma instantânea para se manterem em atividade. Nesse contexto, o uso de ferramentas estatísticas pode se apresentar como uma opção valiosa para ajudar as empresas a tomarem decisões mais bem fundamentadas, o que resultará em melhores resultados.

1.3 Organização do Texto

Esta monografia está organizada em cinco capítulos. O Capítulo 1 apresentou a contextualização, o problema de pesquisa, os objetivos e a justificativa para esta pesquisa. O Capítulo 2 apresentará uma revisão da literatura sobre o mercado de transporte rodoviário de cargas e análise de agrupamento. O Capítulo 3 explicará a metodologia utilizada para esta pesquisa. O Capítulo 4 analisará a parte prática das análises de agrupamento escolhidas no software Minitab.

Por fim, o Capítulo 5 trará as conclusões da pesquisa, reafirmando o objetivo geral e os objetivos específicos alcançados. Serão apresentadas as principais contribuições da pesquisa para a área de transporte rodoviário de carga e serão discutidas possíveis limitações do estudo. Além disso, serão propostas recomendações práticas para empresas que desejam utilizar ferramentas estatísticas para auxiliar na tomada de decisão nesse mercado.

2 Revisão Bibliográfica

2.1 Transporte de Cargas

A satisfação dos clientes depende em grande parte do transporte, que é fundamental para criar as utilidades de tempo e lugar. De fato, o deslocamento entre dois pontos cria a utilidade de lugar, enquanto a rapidez e consistência do transporte criam a utilidade de tempo (LAMBERT, 1994 apud SARDINHA, 2018). A principal parte dos sistemas logísticos empresariais é composta pelo transporte de cargas (WANKE; FLEURY, 2006). Segundo Fernandez et al. (2014), o setor de transporte representa um grande desafio para os economistas devido às suas características específicas, que são propícias ao estudo de diversos problemas econômicos clássicos. Entre esses problemas, podemos citar as externalidades, as possíveis economias de escala e os custos afundados, a natureza espacial do setor, a impossibilidade de armazenamento de produtos, as economias de densidade e os períodos de alta e baixa demanda, dentre outros.

Os autores também citam que o transporte de cargas envolve mais agentes e outras características particulares, como o armazenamento, o que torna mais difícil modelar e estimar com precisão quando comparado com o transporte de pessoas. Existem, essencialmente, cinco tipos de transporte: ferroviário, rodoviário, hidroviário, dutoviário e aéreo.

2.1.1 Transporte Rodoviário

O transporte rodoviário de cargas corresponde ao transporte por ruas, estradas e rodovias. (JACINTO et al., 2022). No Brasil, segundo a Confederação Nacional do Transporte (2023), o modal rodoviário possuía uma participação de aproximadamente 65% na matriz de transporte de cargas em Fevereiro de 2023. Modais como o ferroviário e cabotagem ocupavam respectivamente 15% e 11% do total da matriz. No transporte por rodovia, os caminhões são capazes de circular em todos os tipos de estradas, atendendo a qualquer combinação de origem e destino, movimentando produtos de tamanhos e pesos diversos e cobrindo diferentes distâncias, proporcionando velocidades satisfatórias em trajetos de curta distância. Entretanto, sua principal desvantagem é a capacidade inferior de carga, além da ineficiência em trajetos de longa distância, o que resulta em um alto custo variável, de acordo com Bowersox e Closs (2007 apud SARDINHA, 2018). Os caminhões podem ser classificados em 9 categorias:

1. **Caminhão plataforma:** utilizado para transportar contentores e cargas volumosas ou pesadas
2. **Caminhão baú:** possui estrutura semelhante à de um contêiner, protegendo a carga de intempéries
3. **Caminhão tremonha:** utilizado para transporte de cargas a granel, que são descarregadas por gravidade
4. **Caminhão aberto:** utilizado para cargas não perecíveis e de pequeno volume, podendo ser coberto com encerados em caso de chuva

5. **Caminhão refrigerado:** destinado a transportar cargas perecíveis, possui mecanismos próprios de refrigeração e controle de temperatura no compartimento
6. **Caminhão tanque:** carcaça é um reservatório dividido em tanques, utilizado para transportar derivados de petróleo e outros líquidos a granel
7. **Caminhão graneleiro ou silo:** carcaça própria para transportar granéis sólidos, descarregando por gravidade ou por portas especiais
8. **Caminhão especial:** projetado para cargas pesadas, podendo ser rebaixado e reforçado, e equipado com um guindaste para transporte de veículos
9. **Semirreboques:** carcaças sem propulsão própria, acopladas a caminhões-tratores ou cavalos-mecânicos para formar conjuntos articulados

O transporte rodoviário tem custos fixos reduzidos, pois o equipamento dos terminais e dos veículos não é muito dispendioso. No entanto, a construção e manutenção das vias onde os veículos circulam apresentam custos bastante elevados, especialmente quando há a necessidade de superar obstáculos naturais, como rios ou terrenos acidentados, que exigem a execução de obras complexas de engenharia, como construção de pontes, túneis, viadutos, entre outros. (SARDINHA, 2018)

2.2 Custos de Transporte

De maneira geral, as decisões relacionadas ao transporte de cargas têm um impacto significativo na relação entre custos fixos e variáveis, o que é um fator crucial na formulação de estratégias pelas empresas que contratam e prestam serviços de transporte. Por exemplo, os custos de operação do transporte aquaviário e ferroviário são quase todos fixos (KAUFMAN, 1953 apud WANKE; FLEURY, 2006) Cada empresa de transporte organiza suas tarifas de frete de forma individual, sendo que o cálculo do frete pode ser baseado no peso, volume ou na capacidade total do veículo utilizado. (ROSA, 2007 apud SARDINHA, 2018)

Cumprе ressaltar que as diferenças nos custos de produção entre países e regiões dentro de um país muitas vezes superam os custos logísticos necessários para o transporte até essas áreas, e isso pode determinar o comércio entre elas.(FERREIRA, 2012 apud SARDINHA, 2018) Poliak et al. (2021) chegaram à conclusão de que os operadores internacionais de transporte rodoviário de carga atuam atualmente no mercado comum da União Europeia sob condições uniformes para a realização de transportes internacionais. É possível considerar três perspectivas para a classificação dos custos de transporte de carga (IZADI, 2020):

1. Sob a perspectiva dos operadores de transporte de carga: Os custos referem-se aos gastos incorridos na prestação dos serviços, ou seja, custos operacionais;
2. Sob a perspectiva dos proprietários de carga: Os custos referem-se principalmente aos preços ou taxas pagas aos operadores de transporte de carga, sendo que alguns proprietários estão dispostos a pagar mais por transporte mais rápido e confiável, o que gera os custos de valor de tempo para o transporte de carga;

3. Sob a perspectiva nacional: Os custos de transporte de carga englobam aspectos sociais, ambientais e econômicos, que incluem tanto componentes financeiros quanto não financeiros, tangíveis e intangíveis.

Um exemplo da composição dos custos de transportes terrestres pode ser observado na Tabela 1. Pode-se notar que as despesas com combustível e o Imposto sobre Circulação de Mercadorias e Serviços (ICMS) possuem grande importância, compondo aproximadamente metade de todos os custos de transporte.

Tabela 1 – Composição dos custos de transporte

Descrição	Porcentagem
Combustível	33,19
ICMS	12,00
Pedágios	11,47
Pneus	10,42
Mão de Obra	9,96
Encargos	7,45
Manutenção	6,89
Gerenciamento	4,31
Seguro de Mercadoria	4,31

Fonte: Adaptado pela autora de (SILVA; CURTI, 2018).

É relevante também citar que, de acordo com as tendências observadas no transporte urbano de mercadorias, o tempo de transporte vazio está aumentando. Isso se deve, em parte, à entrega "just-in-time" e "porta a porta" de algumas encomendas, resultando em um maior número de viagens com cargas menores. Consequentemente, embora esse aumento na eficiência operacional do transporte possa parecer positivo, ele acaba resultando em um aumento nos custos e na poluição, reduzindo a eficiência geral do processo, principalmente devido ao aumento dos custos de transporte. No entanto, se não houver esforços para melhorar as cadeias logísticas, os custos de transporte acabarão aumentando significativamente. (JAIN; LINDSKOG; JOHANSSON, 2012 apud SARDINHA, 2018)

2.3 Terceirização do Transporte Rodoviário

Muitas atividades que antes eram realizadas internamente pelas empresas têm sido gradualmente transferidas a terceiros durante o processo de crescimento das mesmas. Em razão disso, as empresas passaram a adotar a estratégia de "cooperação" a fim de minimizar custos e maximizar a qualidade dos serviços, visando, consequentemente, a melhoria do desempenho. Embora a terceirização logística seja uma ferramenta inicialmente utilizada com o objetivo de reduzir custos, posteriormente ela se tornou um meio de criar mais valor agregado. (ARIF; JAWAB, 2018) De acordo com Nelas (2014 apud SARDINHA, 2018) o conceito de *outsourcing* está se tornando cada vez mais popular entre as empresas, baseado na ideia fundamental de que é crucial para a empresa concentrar-se no que faz melhor do que seus concorrentes, delegando o restante para especialistas no assunto.

A gestão logística pode ser realizada através de parcerias com empresas especializadas, conhecidas como Provedores de Serviços Logísticos (PSL). Esses provedores podem oferecer desde serviços básicos

de transporte, baseados na propriedade de ativos fixos, como armazéns e meios de transporte, nomeados como tradicionais, até funções mais sofisticadas, como assumir a responsabilidade sobre as operações logísticas dos clientes e gerenciar os contratos da empresa com outros PSL, nomeadamente 3PL. Os PSL do tipo 4PL/LLP são contratados para fornecer uma solução completa para as empresas, podendo até mesmo contratar serviços de terceiros para execução de determinadas atividades, mas continuam sendo responsáveis por todas as atividades logísticas e fazem a interface com o embarcador. (LANGLEY; CAPGEMINI, 2009 apud GUIDOLIN; FILHA, 2010)

Segundo GUIDOLIN e FILHA (2010) o setor mais significativo para a indústria de prestação de serviços logísticos é o automotivo. Além disso, empresas de alta tecnologia, como as de eletrônicos, necessitam de serviços sofisticados, tais como integração com os clientes no planejamento da gestão da cadeia de suprimentos e atuação nas atividades logísticas de toda a cadeia. Os contratos de serviços logísticos para essas empresas tendem a ser de menor prazo, em função do ciclo de vida dos produtos, e as empresas prestadoras desses serviços precisam ser flexíveis para atender às mudanças na configuração global da produção. O mercado de varejo/bens de consumo representa o terceiro maior mercado de serviços logísticos.

VOLPE (2011) concluiu em sua pesquisa "Análise do Transporte Rodoviário de Carga de uma Microempresa" que para microempresas, possuir um veículo rodoviário de carga próprio não é a opção mais indicada, e que a análise de custos entre transporte terceirizado e próprio é essencial para determinar a melhor escolha. O conhecimento sobre logística e custos logísticos é crucial para microempresas, já que a gestão logística é tão importante quanto a produção e outros setores da empresa.

2.3.1 Cargas TL e LTL

No transporte rodoviário de cargas terceirizado, há duas partes principais envolvidas: o carregador e o transportador. O carregador é responsável pela produção ou transformação dos bens que precisam ser transportados, enquanto o transportador possui ativos, como caminhões e armazéns, que podem ser utilizados no transporte dessas mercadorias do ponto de origem até o destino final. O embarcador pode contratar diretamente o transportador ou pode fazer isso por meio de intermediários. A terceirização de serviços logísticos é justificada como uma forma de aumentar a flexibilidade da empresa diante das flutuações do mercado.

Segundo Yin e Rallis (2018), o transporte rodoviário pode ser segmentado em carga completa, em inglês Truck Load (TL), carga fracionada, em inglês *less-than-truck-load* (LTL) e frotas privadas e dedicadas. Carga completa se refere ao modelo de transporte de ponto a ponto que, como o nome sugere, transportam cargas completas de um único remetente, ocupando o espaço completo do caminhão. Por outro lado, o modelo de carga fracionada lida com remessas menores de diferentes empresas carregadoras consolidadas no caminhão. Por fim, frotas privadas e dedicadas são de propriedade ou alugadas por um único remetente e, geralmente, transportam apenas suas próprias remessas.

2.4 Estatística

A etimologia da palavra "estatística" remonta ao latim "status", que significa "Estado". Inicialmente, o termo era usado para descrever um conjunto de dados relacionados ao Estado, tornando a estatística

uma ferramenta de administração para governantes, com o objetivo de controlar as finanças e a segurança nacional. A partir do século XIX, a estatística passou a ganhar importância em outras áreas do conhecimento humano. No século XX, ela se tornou uma ferramenta essencial para grandes empresas e organizações que buscavam a qualidade total, tornando-se um diferencial competitivo. Ao longo dos séculos, a estatística evoluiu com base em conjuntos de métodos e processos destinados a estudar e medir fenômenos coletivos. (FAVERO, 2017)

Segundo MONTGOMERY e RUNGER (2021), estatística é a ciência que nos ajuda a tomar decisões e a tirar conclusões na presença de variabilidade. As ferramentas estatísticas são empregadas para entender a variabilidade presente em sistemas ou fenômenos. A variabilidade, nesse contexto, significa que observações sucessivas não produzem resultados idênticos. A abordagem estatística permite incorporar essa variabilidade em processos de tomada de decisão, o que é útil nas mais diversas situações do cotidiano, como no caso do valor do frete de um transporte rodoviário de carga. Esse desempenho não é constante em transporte realizado, e diversos fatores podem influenciar a variabilidade observada, tais como a oferta e demanda de cargas, disponibilidade de caminhões e época do ano. A estatística nos oferece uma estrutura para descrever essa variabilidade e identificar as fontes potenciais de variabilidade mais importantes ou que têm maior impacto na composição do preço.

A estatística pode ser dividida em três grandes áreas: estatística descritiva (também conhecida como dedutiva), estatística probabilística e estatística inferencial (ou indutiva). A estatística descritiva permite a análise exploratória das variáveis estudadas, como sugere o próprio nome. Por meio de técnicas de ordenamento, medidas de posição, dispersão e medidas separatrizes da mediana, estudantes e profissionais de análise de crédito podem interpretar os fenômenos estudados em relação à sua frequência simples e relativa. Interpretar corretamente estatísticas como média, moda, mediana e desvio-padrão é de extrema importância, apesar de parecerem simples à primeira vista. (VIRGILLITO, 2017)

A estatística probabilística, por sua vez, recorre à teoria das probabilidades para explicar a frequência de eventos incertos e, assim, estimar ou prever a ocorrência de eventos futuros. (FAVERO, 2017). Por exemplo, a quantidade de transportes de carga necessária em uma empresa pode variar de acordo com a demanda específica na época, diversas empresas produzem praticamente o dobro durante o verão e não é possível saber com certeza quantos transportes serão necessários. Nesse caso, a probabilidade pode ser utilizada para indicar a possibilidade da ocorrência de cada evento possível, para auxiliar no planejamento realizado pela liderança da empresa. Por último, segundo Virgillito (2017), o objetivo de estudar inferência estatística é compreender o significado prático da estimação e dos estimadores, com o propósito de inferir resultados obtidos nas populações a partir de amostras coletadas. Essa área da estatística oferece ferramentas de previsão que podem ser aplicadas a eventos em diversas áreas do conhecimento. É de grande importância ser capaz de identificar características de interesse em uma amostra e estabelecer conexões com a população da qual foi retirada.

2.4.1 Estatística na Tomada de Decisão

O uso de dados na tomada de decisão desempenha um papel essencial na gestão estratégica de qualquer entidade organizacional. Os dados proporcionam uma percepção objetiva e fundamentada em fatos, possibilitando uma análise precisa do ambiente empresarial. Ao utilizar dados como alicerce para a tomada de decisão, as organizações aumentam suas probabilidades de êxito, reduzem riscos e otimizam

seus resultados, conferindo-lhes uma vantagem competitiva no mercado. Para compreender tendências, identificar padrões e antecipar mudanças, os responsáveis pela tomada de decisão devem adquirir, organizar e examinar dados pertinentes, o uso de ferramemas estatísticas é essencial para a garantia do sucesso deste processo.

Segundo [SANTOS et al. \(2016\)](#), para resolver os problemas, o gestor de uma empresa precisa não apenas estar ciente e compreender estatística, mas também expandir ainda mais sua visão. Cada empresa encontrará soluções para seus problemas com base no grau de conhecimento que possuem em cada área, e cada indivíduo assume a responsabilidade de agir de acordo com as expectativas de resultados estabelecidas. De acordo com [Bekman e Neto \(2009\)](#), para a formulação de um problema de decisão, é necessário fornecer uma descrição completa do mesmo, incluindo as seguintes informações:

1. **Relação de todas as opções possíveis:** Isso envolve identificar todas as alternativas disponíveis, tanto em termos de possíveis cursos de ação a serem tomados quanto em relação à coleta ou aquisição de novas informações relevantes.
2. **Lista de todos os eventos que podem ocorrer como resultado das possíveis decisões:** É essencial considerar todas as possíveis consequências que podem surgir de cada opção disponível.
3. **Cronologia das informações e tomada de decisões:** É importante estabelecer a ordem em que as informações pertinentes serão conhecidas pelo tomador de decisão e quando as decisões devem ser tomadas, levando em conta a temporalidade dos eventos.
4. **Quantificação das preferências do tomador de decisão:** As preferências do decisor em relação às consequências associadas a cada curso de ação devem ser quantificadas. Isso envolve atribuir valores ou pesos para avaliar o grau de importância ou utilidade de cada resultado possível.
5. **Julgamento probabilístico sobre a ocorrência dos eventos possíveis:** É necessário realizar uma avaliação probabilística, estimando as probabilidades de ocorrência dos diferentes eventos que podem surgir como resultado das decisões tomadas.

Dentro do contexto do controle de qualidade em processos industriais, é comum utilizar ferramentas e técnicas estatísticas para monitorar, controlar ou aprimorar os processos de produção. No entanto, estudos recentes indicam que, para certas organizações, principalmente as de grande porte, a estatística tem um potencial que vai além do simples objetivo operacional de buscar a melhoria contínua. Ela oferece vantagens competitivas adicionais decorrentes da garantia da qualidade de produtos e processos. ([SANTOS; ANTONELLI, 2011](#))

2.5 Análise de Agrupamento

Segundo [Favero \(2017\)](#), a análise de agrupamentos é uma coleção de técnicas exploratórias altamente úteis que podem ser aplicadas com o objetivo de identificar a existência de comportamentos semelhantes entre várias observações (como indivíduos, empresas, municípios, países, entre outros) em relação a determinadas variáveis. Seu principal propósito é criar grupos ou clusters nos quais a homogeneidade interna prevaleça. Essas técnicas, também conhecidas como análise de conglomerados ou análise de clusters, visam alocar as observações em um número relativamente pequeno de grupos que

sejam homogêneos internamente e heterogêneos entre si, representando assim o comportamento coletivo das observações em relação às variáveis específicas.

Em uma situação geral, em que se tenha uma amostra com n dados de um conjunto de p variáveis de processo, X_1, X_2, \dots, X_p , a análise de agrupamentos tem como objetivo:

1. Classificar os dados em g grupos com base em critérios de similaridade ou diferença entre os dados ou entre os grupos de dados.
2. Permitir a interpretação dos resultados obtidos, que consistem no conjunto dos g grupos de dados.

O resultado da Análise de Agrupamentos dependerá dos critérios adotados para agrupar os dados. O agrupamento é realizado de modo a minimizar as disparidades entre as amostras dentro de cada agrupamento (cluster), enquanto maximiza as disparidades entre as amostras de agrupamentos distintos. (ALBUQUERQUE, 2013)

Segundo Favero (2017), existem dois métodos possíveis para realizar a análise de agrupamento: os hierárquicos e os não hierárquicos. Os métodos hierárquicos, que serão explorados em maior detalhe no presente estudo, permitem identificar a ordem e alocação das observações, oferecendo ao pesquisador a possibilidade de estudar, avaliar e decidir sobre a quantidade de agrupamentos a serem formados. Por outro lado, nos métodos não hierárquicos, parte-se de uma quantidade conhecida de clusters e, em seguida, é elaborada a alocação das observações nesses clusters, seguida pela avaliação da representatividade de cada variável na formação dos mesmos. É possível realizar o agrupamento tanto de variáveis quanto de observações. No caso do agrupamento de variáveis, o objetivo é reunir variáveis em agrupamentos que compartilham características comuns. Esses agrupamentos permitem reduzir o número de variáveis para análise, sendo uma abordagem apropriada quando não há informações iniciais sobre como formar os grupos. Por outro lado, o agrupamento de observações consiste em juntar observações que possuem características semelhantes em grupos. Essa análise também é apropriada quando não há informações iniciais sobre como formar os grupos.

2.6 Etapas da Análise de Agrupamento Hierárquica

Esta seção descreverá as etapas envolvidas na realização da análise de agrupamento hierárquica, que incluem:

2.6.1 Matriz de Distância ou Dissimilaridade

A primeira etapa na análise de agrupamento consiste em estabelecer uma matriz de dissimilaridades entre as observações que estão sendo analisadas. Cada linha da matriz representa uma observação única, enquanto as colunas representam diferentes medidas ou variáveis. É importante ter duas ou mais colunas numéricas para realizar o cálculo das distâncias.

No entanto, no caso em que as variáveis são categóricas, é necessário converter os valores de texto para uma escala numérica antes de utilizar essas variáveis na análise de agrupamento. Por exemplo, uma equipe de Marketing está realizando um estudo de pesquisa de mercado sobre as preferências dos consumidores em relação a um novo produto, e uma das variáveis de interesse é o "Nível de Preferência", que possui

as seguintes categorias: "Altamente Preferir", "Levemente Preferir", "Neutro" e "Não Preferir". Para realizar uma análise de agrupamento nessa variável, é preciso recodificar essas categorias em valores numéricos recodificando essas categorias em valores numéricos para realizar o agrupamento de observações. Uma abordagem comum é atribuir valores como +2, +1, -1 e -2 para as categorias, respectivamente. (MINITAB, 2023) Dessa forma, a matriz de distâncias será composta tanto por medidas numéricas contínuas como por valores numéricos atribuídos às variáveis categóricas, permitindo a aplicação adequada dos algoritmos de agrupamento e a obtenção dos resultados desejados.

Tabela 2 – Modelo geral de um banco de dados para elaboração da análise de agrupamentos

Observação	Variável 1	Variável 2	Variável k
i	x_{1i}	x_{2i}	x_{ki}
1	x_{11}	x_{21}	x_{k1}
2	x_{12}	x_{22}	x_{k2}
3	x_{13}	x_{23}	x_{k3}
4	x_{14}	x_{24}	x_{k4}
n	x_{1n}	x_{2n}	x_{kn}

Fonte: Elaborado pela Autora (2023)

Na Tabela 2, pode-se notar um banco de dados apresentando n observações e, para cada observação i ($i = 1, \dots, n$), valores correspondentes a cada uma das j ($j = 1, \dots, k$). Existem várias medidas de dissimilaridade que podem ser utilizadas na análise de agrupamento hierárquico, sendo a Distância Euclidiana 2.1, a mais comum. Nesta equação, $d(p, q)$ representa a medida de dissimilaridade entre dois pontos p e q em um espaço de k dimensões. A soma é realizada sobre todos os componentes, de $j=1$ a k , e o quadrado é aplicado a cada diferença.

$$d(p, q) = \sqrt{\sum_{j=1}^k (x_{jp} - x_{jq})^2} \quad (2.1)$$

Já distância euclidiana quadrática 2.2 pode ser considerada como uma opção adicional à distância euclidiana, adequada para situações em que as variáveis apresentam pouca dispersão nos seus valores. O cálculo da distância euclidiana ao quadrado proporciona uma interpretação mais intuitiva dos resultados da análise, além de facilitar a atribuição das observações aos grupos correspondentes. Essa abordagem é particularmente valiosa para aprimorar a compreensão dos resultados obtidos e otimizar o processo de agrupamento.

$$d(p, q) = \sum_{j=1}^k (x_{jp} - x_{jq})^2 \quad (2.2)$$

A distância de Manhattan 2.3, também conhecida como distância absoluta ou distância em bloco, não leva em consideração a geometria triangular associada à expressão original de Pitágoras. Em vez disso, ela considera apenas as diferenças entre os valores de cada variável.

$$d(p, q) = \sum_{j=1}^k |x_{jp} - x_{jq}| \quad (2.3)$$

Diferentes medidas de distância podem levar a resultados distintos na análise de agrupamento, o que pode resultar na alocação de observações em agrupamentos homogêneos diferentes. Portanto, é crucial que o pesquisador, embasado em suas escolhas, considere as razões que o levaram a utilizar uma medida específica em detrimento de outras. A utilização de múltiplas medidas ao analisar o mesmo conjunto de dados pode ajudar a fundamentar essa decisão, permitindo a comparação dos resultados obtidos. (FAVERO, 2017)

2.6.2 Método de Ligação

Após a criação da matriz de dissimilaridade, é imprescindível especificar o método de ligação a ser utilizado. Esse método determina como os agrupamentos serão formados durante o processo de agrupamento hierárquico. Os métodos de ligação único, completo e médio estão entre os principais métodos de ligação utilizados. (ANDERBERG, 1973 apud DONI, 2004) O método de Encadeamento Único 2.4, também conhecido como *nearest neighbor*, prioriza as menores distâncias para formar novos agrupamentos durante o processo de aglomeração, incorporando observações ou grupos. Este método é recomendado quando as observações são relativamente diferentes e deseja-se formar agrupamentos com um mínimo de homogeneidade. No entanto, sua eficácia diminui quando as observações ou grupos estão próximos entre si.

$$d_{mj} = \min(d_{kj}, d_{lj}) \quad (2.4)$$

Por outro lado, o método de Encadeamento Completo 2.5, também conhecido como *furthest neighbor*, segue uma abordagem oposta, privilegiando as maiores distâncias entre as observações ou grupos para formar novos agrupamentos. É recomendado quando não há uma grande separação entre as observações e o pesquisador deseja identificar heterogeneidades entre elas. (FAVERO, 2017)

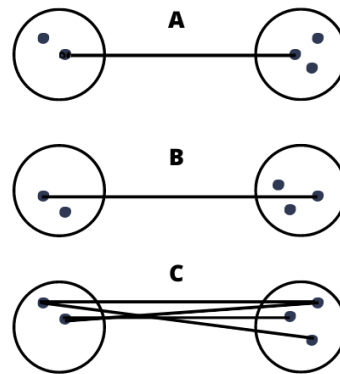
$$d_{mj} = \max(d_{kj}, d_{lj}) \quad (2.5)$$

Já o método de Encadeamento Médio, também conhecido como *average linkage*, une dois grupos com base na distância média entre todos os pares de observações pertencentes a esses grupos. Embora haja alterações no cálculo das medidas de distância entre os agrupamentos, o método de encadeamento médio preserva a ordenação das observações em cada grupo oferecida pelo método de encadeamento único, desde que haja uma considerável separação entre as observações.

$$d_{mj} = \frac{N_k d_{kj} + N_l d_{kl}}{N_m} \quad (2.6)$$

Na Figura 1, são apresentadas as relações entre os métodos de encadeamento e os respectivos métodos de ligação utilizados na análise de agrupamento. A relação **A** está associada ao método de encadeamento único, enquanto a relação **B** está relacionada ao método de encadeamento completo. Por fim, a relação **C** está vinculada ao método de encadeamento médio. (MINITAB, 2023)

Figura 1 – Principais Métodos de Ligação



Fonte: Elaborado pela Autora (2023)

A escolha do método de ligação tem um impacto significativo na estrutura dos agrupamentos resultantes. Cada método de ligação considera diferentes critérios para determinar a proximidade ou dissimilaridade entre os agrupamentos, resultando em diferentes estratégias de agrupamento hierárquico. Portanto, a seleção cuidadosa do método de ligação é crucial para obter agrupamentos adequados e coerentes com os objetivos da análise. (FAVERO, 2017)

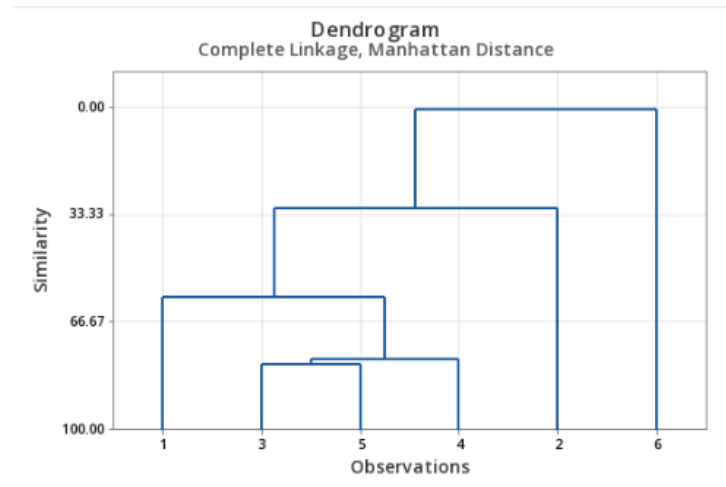
2.6.3 Dendrograma

Finalmente, com base nos estágios de aglomeração e nas distâncias entre os agrupamentos formados, é possível criar um gráfico chamado de dendrograma ou fenograma. Esse gráfico resume o processo de aglomeração e mostra a alocação de cada observação em cada agrupamento. De acordo com FAVERO (2015), o dendrograma fornece uma visualização clara da estrutura hierárquica dos agrupamentos e pode auxiliar na interpretação dos resultados da análise de agrupamento.

A Figura 2 representa um dendrograma de seis pontos correspondendo aos agrupamentos com hierarquia explícita. Os pontos são agrupados com maior semelhança quando estão mais próximos uns dos outros no plano. O dendrograma, um gráfico hierárquico, visualiza de forma explícita a hierarquia dos agrupamentos. No eixo x, os pontos de dados individuais são dispostos (sem ordem específica, exceto para evitar cruzamentos de linhas). Já o eixo y representa a distância entre os agrupamentos. Na parte inferior do dendrograma ($y = 0$), cada ponto está em um agrupamento separado. À medida que o valor de y aumenta, diferentes agrupamentos são fundidos dentro da restrição de distância: primeiramente, 3 e 5 são agrupados, seguidos pela junção de 4. Em seguida, o agrupamento 354 é mesclado com 1, e assim por diante, até que todos os agrupamentos estejam combinados no topo do dendrograma. Os números indicam as junções nos dendrogramas e correspondem aos círculos numerados no diagrama superior.

Além da representação padrão em forma retangular, o dendrograma pode ser visualizado de outras maneiras, como na forma circular ou na forma filogenética. Essas abordagens alternativas permitem explorar diferentes perspectivas da hierarquia dos agrupamentos e facilitam a interpretação dos padrões de similaridade entre os elementos. Cada formato oferece uma representação única do dendrograma, fornecendo insights adicionais sobre a estrutura dos agrupamentos e suas relações. A escolha da forma de

Figura 2 – Exemplo de Dendrograma



Fonte: Elaborado pela Autora (2023)

representação depende das necessidades específicas da análise e das preferências do pesquisador.

2.6.4 Softwares Estatísticos

Para Okagbue et al. (2020), softwares estatísticos são programas de computador especializados e projetados para realizar a análise de dados estatísticos. Estes são disponibilizados como pacotes independentes, extensões, linguagens de programação ou macros/complementos. Os softwares podem ser utilizados por meio de uma interface gráfica do usuário ou por meio de script, com a finalidade de analisar dados provenientes de um processo cuidadosamente organizado de observação e experimentação. A análise dos dados pode resultar em um resultado que pode ser interpretado para gerar conhecimento.

O Software Minitab teve início como uma adaptação voltada para estudantes do sistema Omnitab do National Bureau of Standards, criado por Joseph Hilsenrath e posteriormente desenvolvido por David Hogben e Sally Peavy. O nome Minitab foi derivado dessa origem. O Software Estatístico foi originalmente desenvolvido em 1972 para estudantes em cursos introdutórios de estatística na Universidade Estadual da Pensilvânia. Entretanto, por ter sido desenvolvido como um programa portátil, ou seja, que poderia ser executado em diferentes tipos de computadores, o Minitab foi amplamente distribuído para outras universidades. (RYAN; JOINER; CRYER, 2013)

O Minitab permite que todos os usuários possam prever resultados melhores, projetar produtos aprimorados e otimizar processos, resultando em um aumento de receitas e redução de custos. Este se destaca por sua abordagem integrada e exclusiva, fornecendo software e serviços que promovem a excelência empresarial em qualquer lugar, utilizando recursos de computação em nuvem. Entre os principais testes estatísticos disponíveis no software estão os testes-t, testes de proporção para uma ou duas amostras, teste de normalidade, teste qui-quadrado e testes de equivalência. Adicionalmente, o software possibilita a utilização de métodos tradicionais no *Minitab Statistical Software*, integrando com as linguagens de código aberto R ou Python, ou ainda aprimorando seus recursos com algoritmos de aprendizado de máquina, como Árvores de Classificação e Regressão (CART®), *TreeNet* e *Random Forests*.

3 Metodologia

A metodologia desta pesquisa foi a de pesquisa direta com dados aleatórios. Esta é uma metodologia que envolve a coleta de informações por meio de interação direta com os participantes ou fontes relevantes, porém, com a particularidade de que os dados coletados são gerados de forma aleatória, não correspondendo a observações reais. Nesse contexto, a pesquisa busca simular cenários e contextos específicos relacionados ao mercado de transporte rodoviário de cargas parciais, usando dados fictícios que se assemelham às características do setor. Essa abordagem permite explorar como a ferramenta estatística de análise de agrupamento no Software Minitab pode ser aplicada em diferentes contextos simulados, proporcionando insights valiosos sobre seu potencial uso na tomada de decisões nesse mercado.

3.1 Definição do Problema de Otimização

O problema de otimização consiste em encontrar a melhor forma de definir os grupos de cargas parciais que possam ser transportadas juntas em uma mesma rota. Cada carga parcial é caracterizada por um conjunto de atributos. O objetivo é agrupar as cargas de forma a minimizar a quantidade de rotas e a distância total percorrida pelos veículos de transporte, levando em consideração as especificações de cada carga.

3.2 Elaboração do Banco de Dados

Para fins de pesquisa, a base de dados foi criada aleatoriamente. Esta seguiu critérios específicos para garantir sua relevância e aplicabilidade ao problema em questão. Características como magnetização, acabamento externo, urgência e destino foram consideradas para 100 tipos de cargas distintas.

3.3 Configuração do Minitab

Primeiramente, o banco de dados foi importado para o Minitab no formato de planilha. As variáveis foram devidamente configuradas, garantindo a adequada representação dos dados. Nenhum tratamento especial foi aplicado aos dados antes da análise, assegurando a integridade e fidelidade dos resultados obtidos.

3.4 Seleção dos Métodos

Há diferentes métodos de encadeamento disponíveis no Minitab, cada um com abordagens distintas. Neste estudo foi optado por selecionar dois métodos de encadeamento comuns, encadeamento completo (Complete Linkage) e encadeamento médio (Average Linkage) e três métricas de distância: euclidiana, euclidiana quadrática e Manhattan. A escolha desses métodos deve-se em função de sua habilidade em lidar com variados tipos de dados e sua relevância para a otimização do transporte de cargas.

3.5 Avaliação dos Agrupamentos e Proposta

Baseando-se nos resultados da análise de agrupamento, a autora conclui o trabalho propondo uma abordagem de otimização para o transporte de cargas parciais terrestres. Através da formação de grupos de cargas similares, torna-se possível criar rotas de transporte mais eficientes, reduzindo o número de veículos utilizados e minimizando a distância total percorrida.

4 Discussão

4.1 Matriz de Variáveis

Inicialmente, procede-se à importação da matriz de variáveis para o software Minitab. Essa matriz foi estruturada com o propósito de representar 100 distintos tipos de cargas, cada uma caracterizada por seis variáveis categóricas: magnetização, acabamento externo, urgência e destino e encontra-se no Anexo 1. Em virtude da limitação do Minitab em aceitar variáveis categóricas para a análise em questão, foi necessário converter os valores de texto para uma escala numérica, a fim de viabilizar o tratamento dos dados (conforme apresentado na Tabela 3).

Tabela 3 – Conversão de Variáveis Categóricas de Texto para Escala Numérica

N	A	B	C	D
1	sim	não oxidável	crítico	a menos de 100km do hub
2	não	oxidável	importante	entre 100 km e 300 km do hub
3			significativo	entre 300 km e 600 km do hub
4			secundário	mais de 600 km do hub
5			irrelevante	

Fonte: Elaborado pela Autora (2023)

As colunas A, B, C, e D representam, respectivamente, as especificações de magnetização da carga, acabamento externo, urgência e destino. Já a coluna N representa o novo valor numérico das variáveis categóricas após a conversão. É importante separar cargas de acordo com sua magnetização uma vez que cargas magnetizadas podem gerar campos magnéticos que interferem em dispositivos eletrônicos sensíveis, como computadores, celulares, sistemas de navegação por GPS e equipamentos de comunicação. Essa interferência pode resultar em falhas temporárias ou permanentes nos dispositivos.

Deve-se atentar ao acabamento externo das cargas uma vez que as cargas parciais oxidáveis contêm substâncias suscetíveis a reações químicas com o oxigênio do ar ou outros elementos presentes no ambiente. Essas reações podem resultar em situações de combustão, corrosão ou formação de produtos tóxicos e perigosos. A separação meticulosa entre cargas oxidáveis e não oxidáveis reduz significativamente o risco de ocorrência de reações indesejadas.

Já a separação com base na urgência e na distância do destino desempenha um papel fundamental na garantia de uma logística eficaz e ágil. Cargas com diferentes níveis de urgência requerem tratamento diferenciado na entrega, tornando essencial a priorização das mais urgentes para evitar atrasos e prejuízos aos clientes e seus respectivos processos produtivos. Além disso, a distância do destino tem um impacto direto no tempo de trânsito e no planejamento das rotas, tornando imprescindível a consolidação das cargas de acordo com a proximidade geográfica para otimizar a alocação dos recursos e minimizar o tempo de transporte. Essa abordagem estratégica possibilita que as empresas deem preferência às entregas mais críticas e maximizem a eficiência das operações, resultando em um serviço mais confiável e competitivo no mercado.

4.2 Aplicação no Minitab

Uma vez aplicado o algoritmo na análise de agrupamento utilizando o Minitab, nota-se especialmente as etapas de amalgamação, a tabela de partição final e o dendrograma.

As etapas de amalgamação indicam as fases de consolidação de agrupamentos durante a execução de um algoritmo de agrupamento. Os estágios de consolidação referem-se especificamente à sequência de unificação de agrupamentos no contexto do agrupamento hierárquico, resultando em uma estrutura hierárquica que retrata as relações de semelhança entre agrupamentos em diferentes níveis. Cada estágio é registrado, propiciando a visualização da progressão da consolidação e a compreensão da gênese dos agrupamentos. O grau de similaridade entre as observações contidas em cada agrupamento é diretamente proporcional ao nível de similaridade, ao passo que a distância entre as observações estabelece uma relação inversa com o nível de distância. Idealmente, a aspiração é atingir agrupamentos com níveis de similaridade significativos e, simultaneamente, níveis de distância notavelmente reduzidos. No entanto, é de suma importância ponderar esse objetivo com a necessidade de preservar um número pragmático e coerente de agrupamentos, para assegurar tanto a aplicabilidade quanto a compreensibilidade dos resultados obtidos.

A Final Partition denota a alocação definitiva das observações em agrupamentos específicos após a execução de um algoritmo de agrupamento. Em sua essência, corresponde à disposição das observações em conjuntos determinados pelo método de agrupamento escolhido. Dentro desta tabela, é possível observar três métricas:

1. **Within Cluster Sum of Squares (WCSS):** Avalia a dispersão dos dados dentro de cada cluster após a análise de agrupamento, mensurando a coesão interna de cada grupo por meio da soma dos quadrados das distâncias entre as observações e o centroide do cluster correspondente. Quanto menor essa métrica, maior a proximidade entre as observações dentro do mesmo cluster, o que indica uma maior homogeneidade e qualidade no agrupamento efetuado.
2. **Average Distance from Centroid:** Avalia a média das distâncias entre as observações de um determinado cluster e o centroide associado a esse grupo. O centroide é definido como o ponto médio das coordenadas das observações internas ao cluster. Essa métrica quantifica a densidade dos dados dentro desse cluster específico. Quanto menor for a média das distâncias entre as observações e o centroide, mais próximas as observações estarão do centro do cluster, indicando maior coesão e compactidade no agrupamento. Isso implica que as observações dentro do cluster são mais similares entre si, em comparação com observações de outros clusters.
3. **Maximum Distance from Centroid:** Se refere à maior distância entre as observações de um cluster específico e o centroide correspondente a esse cluster. Essa métrica quantifica a dispersão máxima dos dados dentro do cluster, fornecendo insights sobre o quão amplamente as observações estão distribuídas em relação ao centroide. Quando essa medida é maior, algumas observações no cluster estão mais afastadas do centroide, indicando uma maior dispersão dos dados. Por outro lado, valores menores indicam que as observações estão mais próximas do centroide, refletindo maior coesão e agrupamento compacto.

Foram analisados os modelos de Encadeamento Completo e Médio considerando as métricas de distância Euclidiana, Euclidiana Quadrática e Manhattan em cada modelo.

4.2.1 Resultados do Minitab

Ao aplicar os diferentes tipos de modelos no software, foi possível notar que os valores observados para o Within-Cluster Sum of Squares (WCSS), a Distância Média e a Distância Máxima do Centróide eram idênticos em todos os casos, como é apresentado na Figura 3. Esta convergência é influenciada pelas características inerentes do conjunto de dados e pela configuração subjacente dos agrupamentos, resultando em uma formação interna de agrupamentos comparativamente consistente, independente das metodologias adotadas.

Figura 3 – Final Partition

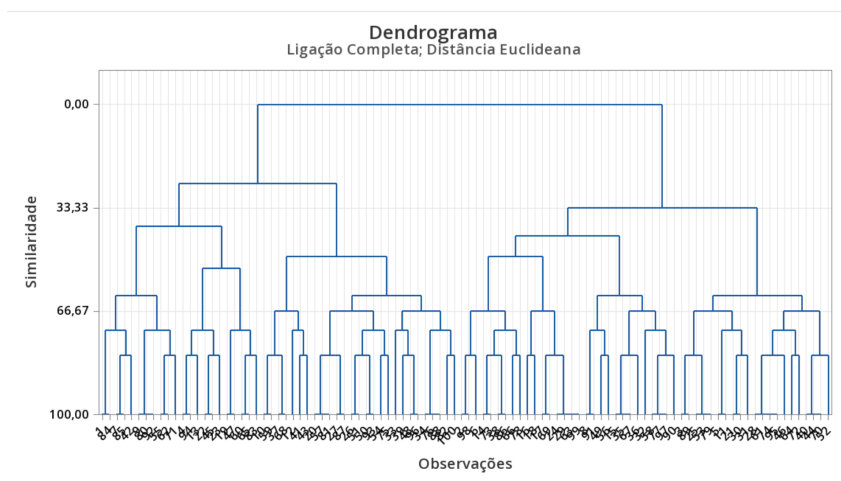
Partição Final

	Número de observações	Dentro da soma de quadrados do agrupado	Distância média do centróide	Distância máxima do centróide
Agrupado1	100	369,41	1,84959	2,81707

Fonte: Elaborado pela Autora (2023)

A Figura 4 representa o dendrograma do primeiro modelo analisado, de Encadeamento Completo com a métrica de Distância Euclidiana. Já a Figura 5 retrata de maneira visual as flutuações dos níveis de similaridade e distância ao longo de todas as etapas de amalgamação. Observa-se uma acentuada diminuição no índice de similaridade a partir da Etapa 91, apresentando uma redução contínua nas subseqüente dez etapas, declinando de 61,51 para o patamar de 0. Simultaneamente, o nível de distância registrou um perfil baixo até a mesma Etapa 91, seguido por um aumento substancial, atingindo um valor de 5,19 na última etapa, praticamente duplicando em relação aos valores anteriores. Na etapa em questão, verificou-se a presença de nove agrupamentos distintos.

Figura 4 – Dendrograma do Encadeamento Completo com a métrica de Distância Euclidiana



Fonte: Elaborado pela Autora (2023)

A Figura 6 representa o dendrograma do Encadeamento Completo com as métricas de Distância Euclidiana Quadrática. Para esta métrica de distância também notam-se mudanças substanciais na etapa

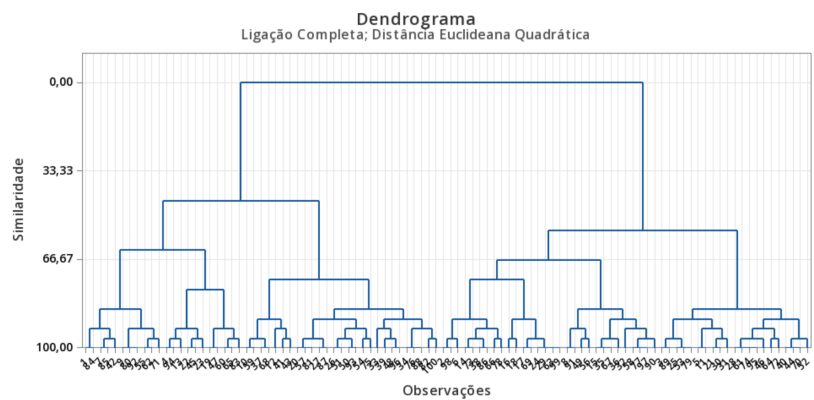
Figura 5 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Euclidiana



Fonte: Elaborado pela Autora (2023)

91, como é possível observar na Figura 7. A medida de similaridade apresenta uma redução significativa, exibindo uma diminuição contínua nas dez etapas subsequentes, declinando de 85,185 para atingir o patamar de 0. Já o nível de distância experimenta um notável aumento a partir da etapa 91, transitando de 4 para 27.

Figura 6 – Dendrograma do Encadeamento Completo com a métrica de Distância Euclidiana Quadrática

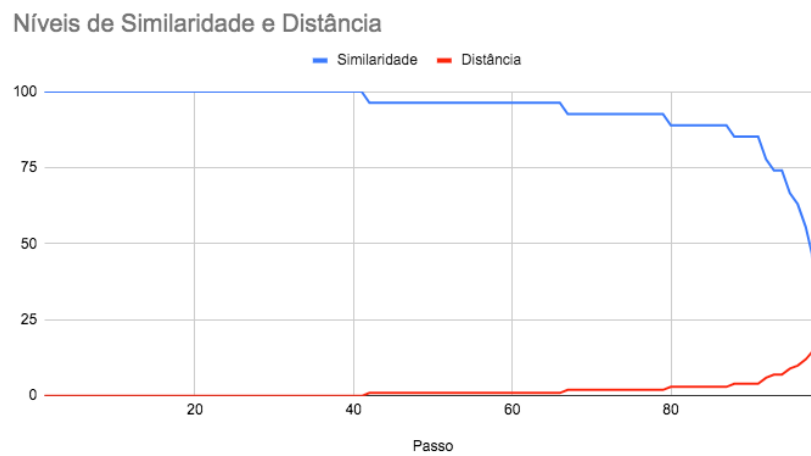


Fonte: Elaborado pela Autora (2023)

O dendrograma do Encadeamento Completo para a distância de Manhattan pode ser observado na Figura 8. Divergindo das métricas de distância Euclidiana e Euclidiana Quadrática, o cenário é distinto para a distância Manhattan, em que a trajetória da similaridade revela um padrão de declínios graduais a partir da etapa 41 até a etapa 87, momento em que se torna mais pronunciado, diminuindo a cada estágio subsequente. No que tange à distância, a etapa 87 se destaca, porém com um aumento discreto, evoluindo de 4 para 9, como pode-se observar na figura 9. Na Etapa 87, haviam 13 clusters.

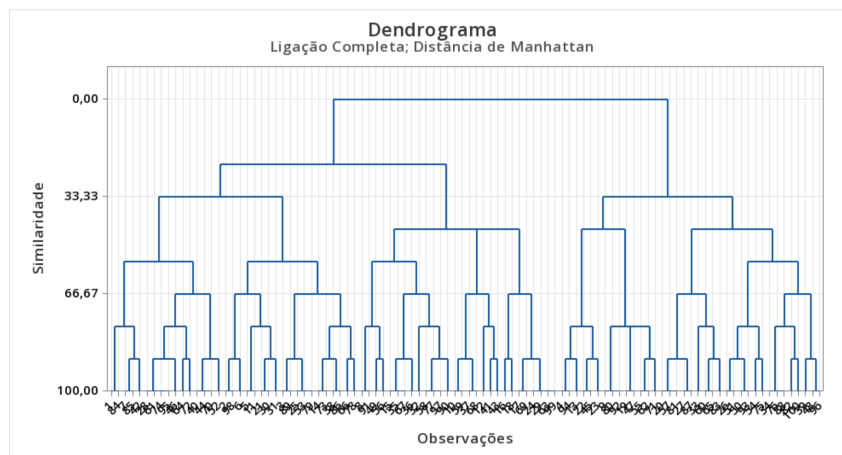
Para o Encadeamento Médio para a distância Euclidiana, pode-se observar o dendrograma na

Figura 7 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Euclidiana Quadrática



Fonte: Elaborado pela Autora (2023)

Figura 8 – Encadeamento Completo com a métrica de Distância Manhattan

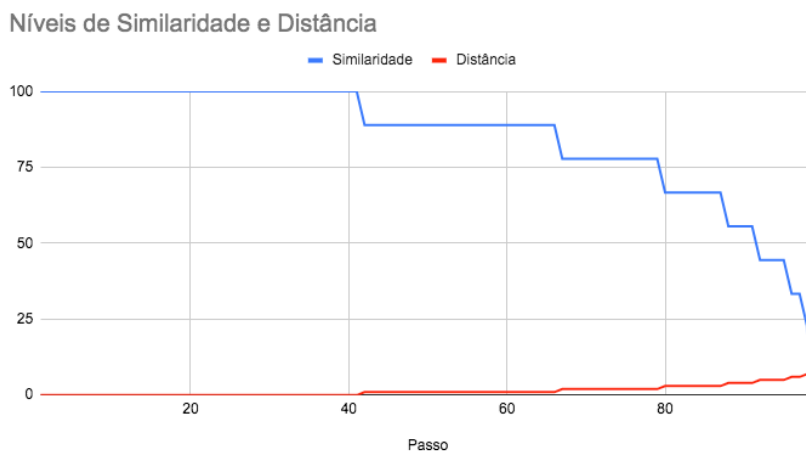


Fonte: Elaborado pela Autora (2023)

Figura 10. No contexto desta métrica de distância empregada no método de encadeamento completo, pode-se discernir uma notável tendência de redução gradual no nível de distância. Especificamente, essa redução é evidente até a Etapa 94, em que se registra uma queda de 66,221 para 38,685, conforme exemplificado na Figura 11. Além disso, na mesma etapa, o nível de distância revela uma transição, com um aumento quantificável de 1,7552 para 3,18604.

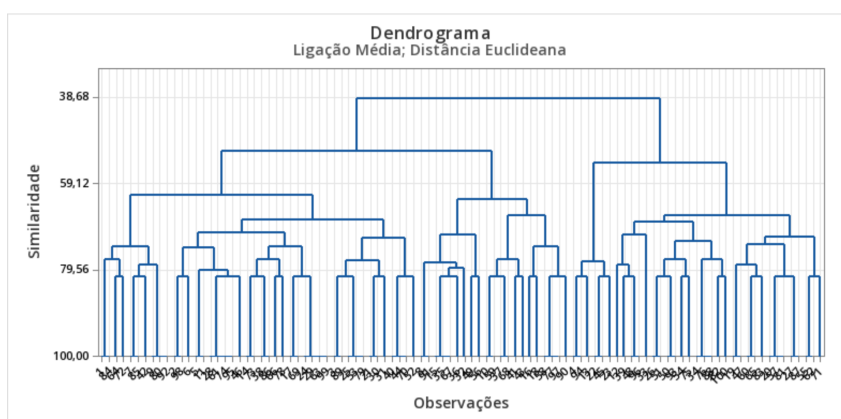
A Figura 12 representa o dendrograma do Encadeamento Médio com as métricas de Distância Euclidiana Quadrática. Igualmente relevante para esta métrica de distância são as mudanças substanciais identificadas na etapa 94, conforme visualizado na Figura 13. A mensuração de similaridade revela uma redução de significativa magnitude, manifestando um declínio contínuo nas dez etapas subsequentes, diminuindo progressivamente de 87,838 para atingir o patamar de 59,608. O nível de distância também sofre um notório incremento a partir da etapa 94, aumentando de 3,2837 para 10,9059.

Figura 9 – Níveis de Similaridade e Distância por etapa do Encadeamento Completo com a métrica de Distância Manhattan



Fonte: Elaborado pela Autora (2023)

Figura 10 – Dendrograma do Encadeamento Médio com a métrica de Distância Euclidiana



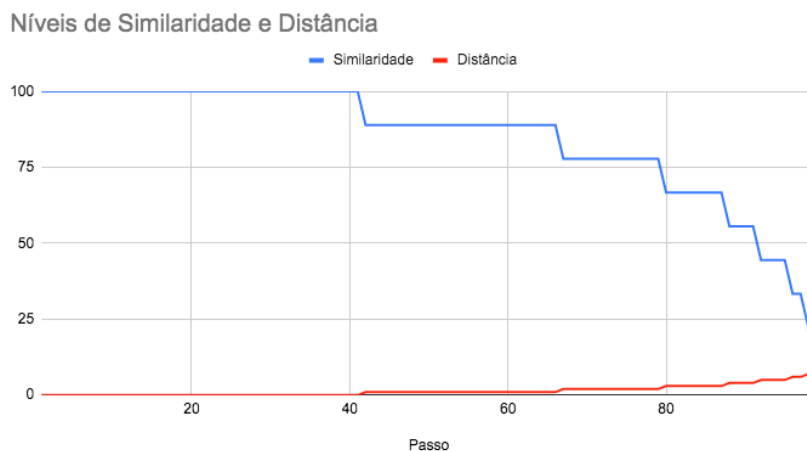
Fonte: Elaborado pela Autora (2023)

A Figura 14 representa o dendrograma do Encadeamento Médio com as métricas de Distância de Manhattan. As oscilações nos índices de similaridade e distância relativos à métrica de distância Manhattan estão representadas na Figura 15. A medida de similaridade revela uma notável redução, manifestando uma diminuição contínua nas seis etapas subsequentes, declinando de 68,107 para atingir o patamar de 48,332. Por outro lado, o nível de distância registra um discreto aumento a partir da etapa 94, evoluindo de 2,87037 para 4,6501.

4.3 Análise

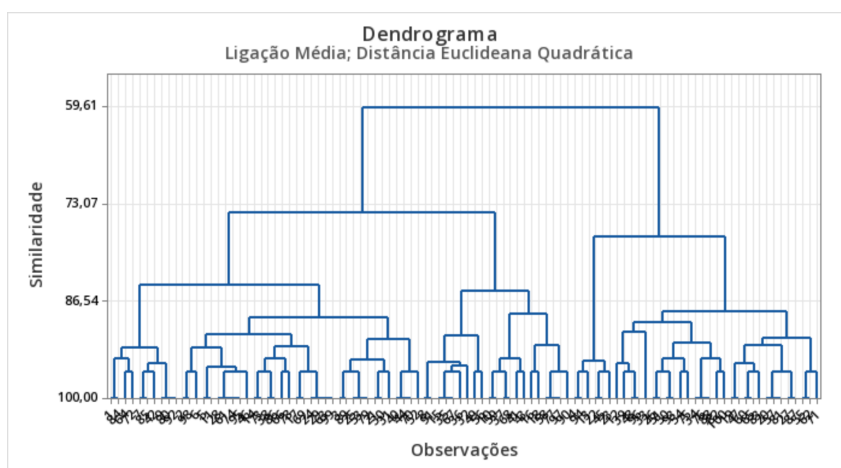
A seleção de solução final de um agrupamento consiste em buscar obter a estrutura mais simples possível que ainda represente agrupamentos homogêneos. (HAIR et al., 2009) Neste trabalho, a autora conclui encontrar o melhor método para análise de agrupamento de cargas parciais em transportes terrestres

Figura 11 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância Euclidiana



Fonte: Elaborado pela Autora (2023)

Figura 12 – Dendrograma do Encadeamento Médio com a métrica de Distância Euclidiana Quadrática



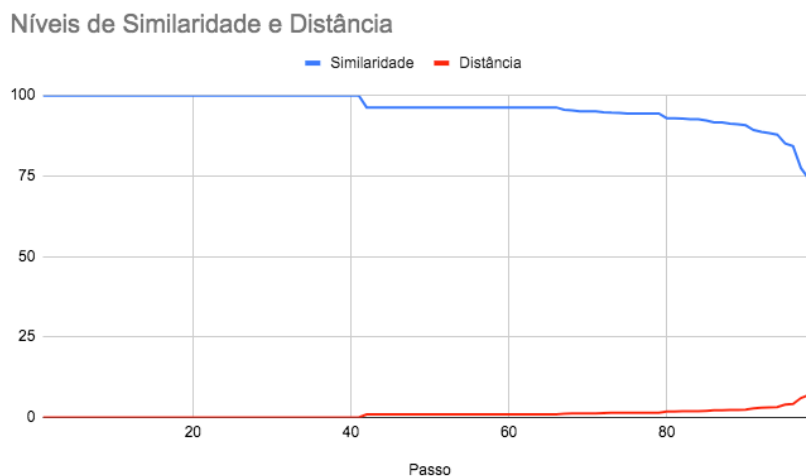
Fonte: Elaborado pela Autora (2023)

aquele que apresenta o maior número de agrupamentos homogêneos possível. Essa conclusão é possível através da análise das flutuações dos níveis de similaridade em cada agrupamento. Acentuados aumentos na heterogeneidade sugerem a união de dois agrupamentos relativamente distintos na etapa em questão. A Tabela 4 compara os níveis de similaridade de cada método de ligação e distância na última etapa, a qual foi considerada pela autora como contendo o maior número de grupos homogêneos possível.

A partir desta análise, a autora julga que o método mais eficaz é o de encadeamento médio com a utilização da distância euclidiana quadrática, apresentando somente 6 agrupamentos e um nível de similaridade de 87.838 entre estes. A representação do agrupamento selecionado encontra-se na Figura 16, na qual os clusters foram discriminados por meio da aplicação de cores distintas para cada um.

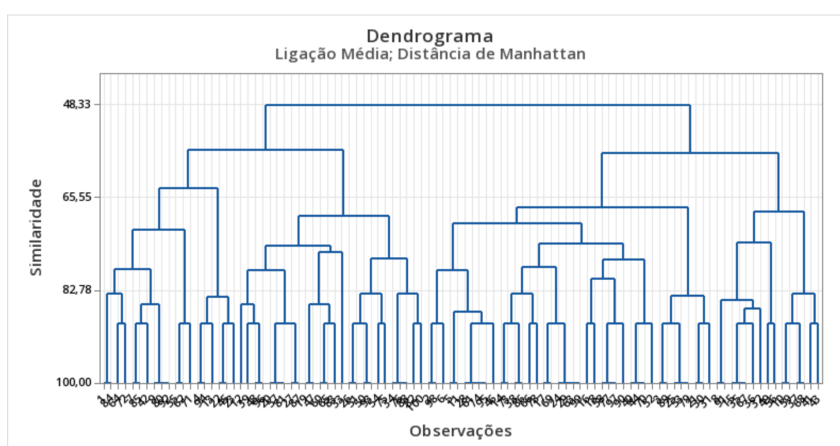
No primeiro cluster, representado na Figura 17, nota-se uma intercalação entre objetos magnetizados e não magnetizados em relação à carga, sugerindo que a magnetização não é determinante nesse

Figura 13 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância Euclidiana Quadrática



Fonte: Elaborado pela Autora (2023)

Figura 14 – Dendrograma do Encadeamento Médio com a métrica de Distância de Manhattan



Fonte: Elaborado pela Autora (2023)

aspecto. A predominância do acabamento oxidável sugere sua relevância na seleção, apesar de uma exceção não oxidável. O destino aponta para uma distinção entre objetos para curtas distâncias e aqueles entre 100 km e 300 km. A urgência revela-se como critério significativo para agrupamento, principalmente em aplicações relevantes.

No segundo cluster, representado pela Figura 18, a similaridades entre as cargas emergem de análises. Magnetização parece não influenciar a relevância, uma vez que tanto objetos magnetizados quanto não magnetizados constam neste cluster. O acabamento oxidável também é comum, mas com exceções não oxidáveis. A categoria de urgência é evidente, com muitos objetos classificados como irrelevantes e secundários. Os destinos variam entre 100 km e 600 km.

Já o terceiro cluster, representado pela Figura 19, apresenta similaridades que revelam padrões

Figura 15 – Níveis de Similaridade e Distância por etapa do Encadeamento Médio com a métrica de Distância de Manhattan



Fonte: Elaborado pela Autora (2023)

Tabela 4 – Comparação dos Níveis de Similaridade na Etapa Conclusiva

Modelo	Distância	Nível de Similaridade	Nº de Agrupamentos
Completo	Euclidiana	61.51	9
Completo	Euclidiana Quadrática	85.185	9
Completo	Manhattan	66.67	13
Médio	Euclidiana	66.221	6
Médio	Euclidiana Quadrática	87.838	6
Médio	Manhattan	68.107	6

Fonte: Elaborado pela Autora (2023)

específicos: a magnetização, geralmente magnetizada, emerge como uma característica dominante em objetos com urgências irrelevantes ou secundárias, destinados a distâncias maiores do que 600 km, sugerindo uma tendência de uso em contextos mais distantes.

Para o quarto cluster, Figura 20, nota-se um foco na característica de magnetização, predominantemente magnetizada. As cargas possuem urgência significativa, tanto primárias quanto secundárias. A distância é notável, com objetos destinados a distâncias superiores a 600 km ou entre 300 km e 600 km.

No penúltimo cluster, Figura 21, as similaridades que destacam-se são as cargas não serem magnetizadas e possuírem acabamento não oxidável. Neste cluster estão presentes objetos onde a importância e a criticidade da aplicação são proeminentes, além de serem destinados a distâncias menores que 100 km.

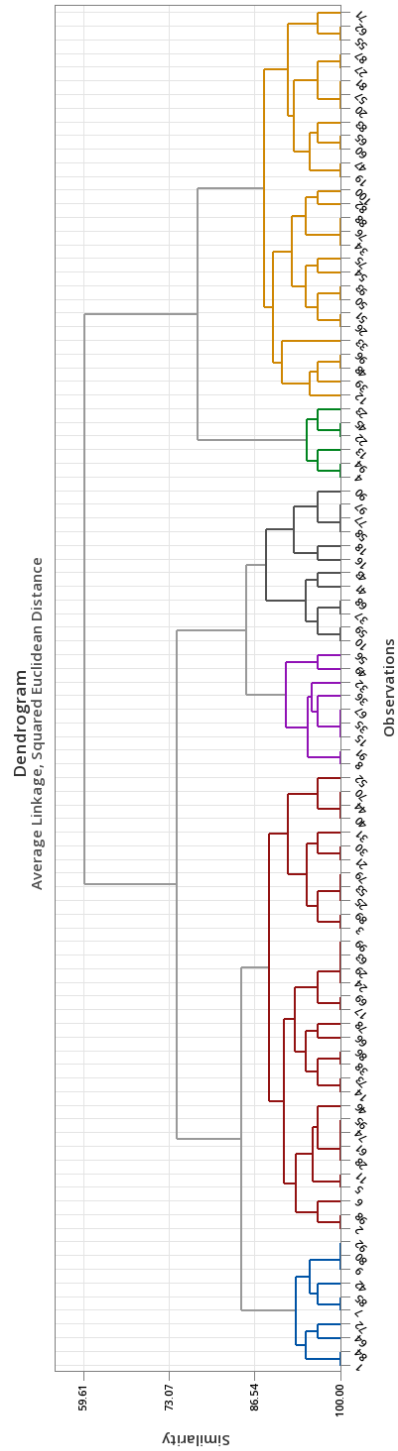
No último cluster, Figura 22, predominam cargas não magnetizadas com acabamento oxidável ou não oxidável, sendo destinados a distâncias entre 100 km e 600 km. A aplicação crítica ou importante é uma característica destacada nesse conjunto. A análise sugere que a não magnetização e a aplicação crucial ou importante são os fatores determinantes nesse agrupamento, com uma faixa de distâncias médias.

4.3.1 Análise Final dos Clusters

A comparação entre todos os clusters revela padrões intrigantes dentro das diversas categorias do conjunto de dados. O Cluster 1 demonstra uma mistura de itens magnetizados e não magnetizados, todos com acabamento externo oxidável e importância significativa na aplicação, com distâncias variando de menos de 100 km a entre 100 km e 300 km. Por outro lado, o Cluster 2 abrange itens diversos em termos de magnetização e acabamento externo, predominantemente com significado secundário ou irrelevância em suas aplicações, e distribuídos em diferentes faixas de distância. O Cluster 3 se destaca como um agrupamento exclusivamente composto por itens magnetizados com diversos acabamentos externos, em sua maioria exibindo importância secundária e destinados a distâncias superiores a 600 km. O Cluster 4 apresenta itens magnetizados e não oxidáveis com importância significativa ou secundária na aplicação, destinados a distâncias superiores a 600 km. Por fim, o Cluster 5 reúne itens principalmente não oxidáveis e com diferentes graus de importância na aplicação, variando de importante a crítico, e tipicamente destinados a distâncias inferiores a 100 km. O Cluster 6 apresenta itens com diferentes graus de oxidabilidade e importância em suas aplicações, sendo que itens críticos e importantes frequentemente possuem uma faixa de destino mais ampla, superior a 300 km. Essas descobertas evidenciam a complexidade das relações entre carga, acabamento externo, importância na aplicação e destino, lançando luz sobre as estruturas intrincadas inerentes ao conjunto de dados.

Diante dessa intrincada complexidade revelada pelos padrões identificados, a avaliação da validade da divisão das cargas entre os clusters recai sobre o analista. A tarefa de discernir a pertinência e eficácia dessa segmentação requer uma análise criteriosa entre os atributos de cada cluster, incluindo carga, acabamento externo, relevância da rápida entrega e destino. A determinação da relevância dessa divisão constitui um passo essencial para extrair percepções acionáveis do conjunto de dados e informar decisões informadas no contexto em questão.

Figura 16 – Agrupamentos Separados por Cores



Fonte: Elaborado pela Autora (2023)

Figura 17 – Cluster 1

Cluster	Carga	Carga	Acabamento externo	Aplicação	Destino
1	1	Sim	Oxidável	Significativo	Menos de 100 km
1	7	Não	Oxidável	Significativo	Menos de 100 km
1	9	Não	Oxidável	Significativo	Entre 100km e 300km
1	42	Não	Não Oxidável	Significativo	Menos de 100 km
1	64	Sim	Oxidável	Secundário	Menos de 100 km
1	72	Não	Oxidável	Secundário	Menos de 100 km
1	80	Não	Oxidável	Significativo	Entre 100km e 300km
1	84	Sim	Oxidável	Significativo	Menos de 100 km
1	85	Não	Oxidável	Significativo	Menos de 100 km
1	92	Não	Oxidável	Significativo	Entre 100km e 300km

Fonte: Elaborado pela Autora (2023)

Figura 18 – Cluster 2

2	2	Não	Oxidável	Irrelevante	Entre 100km e 300km
2	3	Sim	Não Oxidável	Irrelevante	Menos de 100 km
2	5	Sim	Oxidável	Irrelevante	Entre 100km e 300km
2	6	Não	Oxidável	Irrelevante	Entre 300km e 600km
2	11	Sim	Oxidável	Irrelevante	Entre 100km e 300km
2	14	Não	Não Oxidável	Irrelevante	Entre 100km e 300km
2	17	Não	Não Oxidável	Secundário	Entre 300km e 600km
2	21	Não	Oxidável	Irrelevante	Menos de 100 km
2	24	Não	Não Oxidável	Secundário	Entre 100km e 300km
2	25	Não	Não Oxidável	Irrelevante	Menos de 100 km
2	28	Sim	Oxidável	Secundário	Entre 100km e 300km
2	29	Não	Não Oxidável	Secundário	Entre 100km e 300km
2	30	Não	Oxidável	Irrelevante	Menos de 100 km
2	31	Sim	Oxidável	Irrelevante	Menos de 100 km
2	38	Sim	Não Oxidável	Irrelevante	Entre 100km e 300km
2	40	Sim	Não Oxidável	Secundário	Menos de 100 km
2	44	Sim	Não Oxidável	Secundário	Menos de 100 km
2	46	Não	Oxidável	Secundário	Entre 100km e 300km
2	52	Sim	Não Oxidável	Secundário	Entre 100km e 300km
2	53	Não	Não Oxidável	Irrelevante	Menos de 100 km
2	61	Sim	Oxidável	Secundário	Entre 100km e 300km
2	63	Não	Não Oxidável	Secundário	Entre 100km e 300km
2	66	Não	Não Oxidável	Irrelevante	Entre 300km e 600km
2	69	Não	Não Oxidável	Secundário	Entre 300km e 600km
2	70	Sim	Não Oxidável	Secundário	Menos de 100 km
2	73	Não	Não Oxidável	Irrelevante	Entre 100km e 300km
2	74	Sim	Oxidável	Secundário	Entre 100km e 300km
2	78	Sim	Não Oxidável	Irrelevante	Entre 300km e 600km
2	79	Não	Não Oxidável	Irrelevante	Menos de 100 km
2	86	Sim	Não Oxidável	Irrelevante	Entre 100km e 300km
2	89	Sim	Não Oxidável	Irrelevante	Menos de 100 km
2	95	Sim	Oxidável	Secundário	Entre 100km e 300km
2	98	Não	Oxidável	Irrelevante	Entre 100km e 300km
2	99	Não	Não Oxidável	Secundário	Entre 100km e 300km

Fonte: Elaborado pela Autora (2023)

Figura 19 – Cluster 3

3	8	Sim	Não Oxidável	Irrelevante	mais de 600km
3	15	Sim	Oxidável	Irrelevante	mais de 600km
3	32	Sim	Oxidável	Secundário	mais de 600km
3	35	Sim	Oxidável	Irrelevante	mais de 600km
3	36	Sim	Oxidável	Irrelevante	Entre 300km e 600km
3	49	Não	Não Oxidável	Secundário	mais de 600km
3	56	Não	Não Oxidável	Irrelevante	mais de 600km
3	67	Sim	Oxidável	Irrelevante	mais de 600km
3	91	Sim	Não Oxidável	Irrelevante	mais de 600km

Fonte: Elaborado pela Autora (2023)

Figura 20 – Cluster 4

4	10	Sim	Oxidável	Significativo	mais de 600km
4	16	Sim	Não Oxidável	Significativo	Entre 100km e 300km
4	18	Sim	Não Oxidável	Significativo	Entre 300km e 600km
4	37	Não	Oxidável	Significativo	mais de 600km
4	41	Sim	Não Oxidável	Significativo	mais de 600km
4	43	Não	Não Oxidável	Significativo	mais de 600km
4	58	Sim	Não Oxidável	Secundário	Entre 300km e 600km
4	59	Sim	Oxidável	Significativo	mais de 600km
4	68	Não	Oxidável	Significativo	mais de 600km
4	77	Sim	Não Oxidável	Secundário	Entre 300km e 600km
4	90	Sim	Não Oxidável	Secundário	mais de 600km
4	97	Sim	Não Oxidável	Secundário	Entre 300km e 600km

Fonte: Elaborado pela Autora (2023)

Figura 21 – Cluster 5

5	4	Não	Não Oxidável	Importante	Menos de 100 km
5	13	Sim	Não Oxidável	Importante	Menos de 100 km
5	22	Não	Não Oxidável	Crítico	Menos de 100 km
5	23	Sim	Não Oxidável	Crítico	Menos de 100 km
5	45	Não	Não Oxidável	Crítico	Menos de 100 km
5	94	Não	Não Oxidável	Importante	Menos de 100 km

Fonte: Elaborado pela Autora (2023)

Figura 22 – Cluster 6

6	12	Não	Não Oxidável	Significativo	Entre 300km e 600km
6	19	Não	Oxidável	Crítico	Entre 100km e 300km
6	20	Não	Oxidável	Importante	Entre 300km e 600km
6	26	Sim	Oxidável	Crítico	Entre 300km e 600km
6	27	Não	Oxidável	Crítico	Entre 300km e 600km
6	33	Não	Não Oxidável	Crítico	mais de 600km
6	34	Sim	Oxidável	Importante	mais de 600km
6	39	Não	Não Oxidável	Importante	mais de 600km
6	47	Não	Oxidável	Crítico	Entre 100km e 300km
6	48	Não	Não Oxidável	Importante	Entre 300km e 600km
6	50	Sim	Não Oxidável	Crítico	Entre 300km e 600km
6	51	Sim	Oxidável	Crítico	Entre 300km e 600km
6	54	Sim	Oxidável	Importante	Entre 300km e 600km
6	55	Sim	Oxidável	Importante	Entre 100km e 300km
6	57	Não	Oxidável	Importante	Entre 300km e 600km
6	60	Não	Não Oxidável	Crítico	Entre 100km e 300km
6	62	Sim	Oxidável	Importante	Entre 100km e 300km
6	65	Não	Não Oxidável	Crítico	Entre 100km e 300km
6	71	Não	Oxidável	Importante	Entre 100km e 300km
6	75	Sim	Não Oxidável	Importante	Entre 300km e 600km
6	76	Sim	Oxidável	Importante	mais de 600km
6	81	Não	Oxidável	Importante	Entre 300km e 600km
6	82	Sim	Não Oxidável	Importante	mais de 600km
6	83	Não	Não Oxidável	Crítico	Entre 300km e 600km
6	87	Não	Oxidável	Crítico	Entre 300km e 600km
6	88	Sim	Oxidável	Importante	mais de 600km
6	93	Sim	Não Oxidável	Crítico	Entre 300km e 600km
6	96	Não	Não Oxidável	Importante	Entre 300km e 600km
6	100	Sim	Não Oxidável	Crítico	mais de 600km

Fonte: Elaborado pela Autora (2023)

5 Considerações Finais

5.1 Conclusão

De acordo com Hair et al. (2009) a escolha da solução definitiva em uma análise de agrupamento requer considerável discernimento por parte do pesquisador e é amplamente vista como um processo altamente subjetivo. Apesar dos avançados métodos desenvolvidos para auxiliar na avaliação das opções de agrupamento, a responsabilidade final de determinar o número de agrupamentos a serem adotados como solução definitiva permanece nas mãos do pesquisador.

Após a realização minuciosa dos seis tipos de modelos de análise de agrupamento, empregando os métodos de encadeamento completo e médio, considerando distintas métricas de distância euclidiana, euclidiana quadrática e manhattan, a autora conclui por identificar a abordagem ótima para a análise de agrupamento de cargas parciais no contexto de transportes terrestres. Através da avaliação da coesão dos agrupamentos formados, pode-se concluir que o método que resulta no maior número de agrupamentos homogêneos é caracterizado pelo encadeamento médio em combinação com a distância euclidiana quadrática. Ao adotar tal metodologia, os profissionais da área da logística podem embasar suas escolhas de maneira mais assertiva e objetiva, otimizando assim os processos de separação e alocação de cargas, ao mesmo tempo que contribuem substancialmente para a otimização das operações logísticas e o aperfeiçoamento do desempenho global do setor de transporte de cargas parciais.

À medida que o tamanho da amostra aumenta, surge uma maior complexidade na interpretação dos resultados, resultando em desafios na representação visual e na análise apropriada dos agrupamentos identificados. Diante desse cenário, em contextos que envolvem conjuntos de dados de grande magnitude, é prudente considerar a adoção de ferramentas ou abordagens alternativas na análise de agrupamento, a fim de assegurar a obtenção de resultados precisos e eficazes.

5.2 Trabalhos Futuros

Uma direção promissora para pesquisas futuras seria a exploração da implementação de algoritmos genéticos para otimizar o transporte de cargas parciais terrestres. A adoção de técnicas evolucionárias possibilitaria a busca por soluções ótimas ou próximas do ótimo em desafios complexos de alocação de cargas. Mediante a aplicação de modelagem genética e a geração iterativa de alternativas, tais algoritmos podem ser adaptados para enfrentar restrições de capacidade veicular, variados tipos de carga e otimização de rotas, resultando em abordagens altamente eficazes e altamente adaptáveis.

Além disso, uma abordagem adicional promissora seria a integração de sistemas de informações geográficas (SIG) para aprimorar a otimização do transporte de cargas parciais. Ao unir dados geoespaciais, como informações de tráfego, distâncias, topografia e pontos de interesse, seria possível enriquecer a modelagem do problema. Essa integração permitiria uma alocação de cargas mais precisa, abrangendo não apenas a capacidade dos veículos, mas também os traçados das rotas e as condições em tempo real do trânsito. Mediante a fusão das ferramentas SIG com algoritmos de otimização, seria viável desenvolver

estratégias de alocação mais dinâmicas e eficientes, adaptáveis aos cenários em constante evolução das condições geográficas e de tráfego.

Referências

- ALBUQUERQUE, M. A. *ANÁLISE DE AGRUPAMENTO HIERÁRQUICA E INCREMENTAL - ESTUDO DE CASO EM CIÊNCIAS FLORESTAIS*. Dissertação (Mestrado) — UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO, 2013. 9
- ANDERBERG, M. R. Cluster analysis for applications. *New York: Academic Press*, 1973. 11
- ARIF, J.; JAWAB, F. Outsourcing of logistics' activities: Impact analysis on logistics service performance. 2018. 5
- BEKMAN, O. R.; NETO, P. L. de O. C. *Análise estatística da decisão*. 2ª edicao. ed. [S.l.]: Editora Blucher, 2009. 8
- BOWERSOX, D. J.; CLOSS, D. Logística empresarial - o processo de integração da cadeia de suprimento. *São Paulo: Atlas*, 2007. 3
- CHU, C. A heuristic algorithm for the truckload and less-than-truckload problem. *European Journal of Operational Research*, p. 657–667, 2005. 1
- CONFEDERACAO NACIONAL DO TRANSPORTE. *Boletim Unificado*. Brasília, 2023. 34 p. 3
- DONI, M. V. *ANÁLISE DE CLUSTER: MÉTODOS HIERÁRQUICOS E DE PARTICIONAMENTO*. Dissertação (Mestrado) — UNIVERSIDADE PRESBITERIANA MACKENZIE, 2004. 11
- FAVERO, L. P. *Análise de Dados: Técnicas Multivariadas Exploratórias com SPSS e STATA*. [S.l.]: Grupo GEN, 2015. 12
- FAVERO, P. L. *Manual de Análise de Dados - Estatística e Modelagem Multivariada com Excel®, SPSS® e Stata®*. [S.l.]: Grupo GEN, 2017. 7, 8, 9, 11, 12
- FERNANDEZ, X. L. et al. Logistic platforms and efficiency of road haulage in europe (2004-2012). *Revista de Economia Mundial, num. 38*, p. 127–14, 2014. 3
- FERREIRA, S. S. L. *Gestão de armazéns: Implementação de um sistema de picking na indústria alimentar*. Dissertação (Mestrado) — Escola Superior Agrária, Santarém., 2012. 4
- GUIDOLIN, S.; FILHA, D. M. Cadeia de suprimentos: o papel dos provedores de serviços logísticos. *Banco Nacional De Desenvolvimento Econômico E Socialy*, p. 433–484, 2010. 6
- HAIR, J. F. et al. *Análise multivariada de dados*. 6. ed. ed. [S.l.]: Bookman, 2009. 21, 30
- IZADI, M. N. . O. T. A. Cost models and cost factors of road freight transportation: A literature review and model structure. *Fuzzy Information and Engineering*, 2020. 4
- JACINTO, B. et al. Inovação e logística: o compartilhamento de frota para o segmento de transporte rodoviário de cargas luth transportes. 8, 2022. 3
- JAIN, S.; LINDSKOG, E.; JOHANSSON, B. N. Supply chain carbon footprint tradeoffs using simulation. *Proceedings of the Winter simulation Conference*, 2012. 5
- KAUFMAN, J. J. The wage-price relationships in the railroad industry: a comment. *The Journal of Business of the University of Chicago*, p. 48–50, 1953. 4
- LAMBERT, D. M. Custos logísticos, produtividade e análise de desempenho. 1994. 3

- LANGLEY, J.; CAPGEMINI, J. Third-party logistic study 2009. results and findings of 2009. *14th annual study*, 2009. 6
- MINITAB, L. *Suporte ao Minitab® 21*. 2023. Disponível em: <https://support.minitab.com/pt-br/minitab/21/nav_statistical-modeling-multivariate-how-to/>. 10, 11
- MONTGOMERY, D. C.; RUNGER, G. C. *Estatística Aplicada e Probabilidade para Engenheiros*. [S.l.]: Grupo GEN, 2021. 7
- NELAS, F. *Factores de motivação das equipas de vendas em Outsourcing na indústria Farmacêutica*. Dissertação (Mestrado) — IPAM Lisboa, 2014. 5
- OKAGBUE, F. I. et al. Trends and usage pattern of spss and minitab software in scientific research. *International Conference on Recent Trends in Applied Research*, 2020. 13
- POLIAK, M. et al. Competitiveness of price in international road freight transport. *Journal of Competitiveness*, p. 83–98, 2021. 4
- ROSA, A. C. *Gestão de transporte na logística de distribuição física: uma análise da minimização do custo operacional*. Dissertação (Mestrado) — Departamento de Economia, Contabilidade e Administração da Universidade de Taubaté, São-Paulo, 2007. 4
- RYAN, B. F.; JOINER, B. L.; CRYER, J. D. *MINITAB Handbook: Update for Release*. 6ª edição. ed. [S.l.]: Cengage, 2013. 13
- SANTOS, A. B.; ANTONELLI, S. C. Aplicação da abordagem estatística no contexto da gestão da qualidade: um survey com indústrias de alimentos de são paulo. *Gest. Prod.* 18 (3), p. 509–524, 2011. 8
- SANTOS, B. M. et al. A importância e o uso da estatística na Área empresarial: uma pesquisa de campo com empresas do município de elói mendes - mg. *SEGET*, XIII, 2016. 8
- SARDINHA, F. D. J. A gestão de transportes na cadeia de logística. 2018. 3, 4, 5
- SCHNEIDER, L. M. New era in transportation strategy, transportation strategy. p. 118–126, 1985. 1
- SILVA, R.; CURI, M. Custos logísticos: um estudo sobre a composição do frete rodoviário entre zonas aduaneiras. *Brazilian Journal Of Development*, p. 3773–3788, 2018. 5
- VIRGILLITO, B. S. *Estatística Aplicada*. [S.l.]: Editora Saraiva, 2017. 7
- VOLPE, P. M. Análise do transporte rodoviário de carga de uma microempresa. 2011. 6
- WANKE, P.; FLEURY, P. Estrutura e dinâmica do setor de serviços no brasil. In: _____. [S.l.: s.n.], 2006. cap. Transporte de Cargas no Brasil: Estudo Exploratório das Principais Variáveis Relacionadas aos Diferentes modais e às suas estruturas de custos, p. p. 409–464. 3, 4
- YIN, B. B.; RALLIS, C. W. A study of shipper performance in the less-than-truckload market. 2018. 6

Anexos

ANEXO A – Base de Dados Utilizada

Carga	Acabamento externo	Urgência	Destino
1	2	3	1
2	2	5	2
1	1	5	1
2	1	2	1
1	2	5	2
2	2	5	3
2	2	3	1
1	1	5	4
2	2	3	2
1	2	3	4
1	2	5	2
2	1	3	3
1	1	2	1
2	1	5	2
1	2	5	4
1	1	3	2
2	1	4	3
1	1	3	3
2	2	1	2
2	2	2	3
2	2	5	1
2	1	1	1
1	1	1	1
2	1	4	2
2	1	5	1
1	2	1	3
2	2	1	3
1	2	4	2
2	1	4	2
2	2	5	1
1	2	5	1
1	2	4	4
2	1	1	4
1	2	2	4
1	2	5	4
1	2	5	3
2	2	3	4
1	1	5	2
2	1	2	4
1	1	4	1
1	1	3	4
2	1	3	1
2	1	3	4
1	1	4	1

2	1	1	1
2	2	4	2
2	2	1	2
2	1	2	3
2	1	4	4
1	1	1	3
1	2	1	3
1	1	4	2
2	1	5	1
1	2	2	3
1	2	2	2
2	1	5	4
2	2	2	3
1	1	4	3
1	2	3	4
2	1	1	2
1	2	4	2
1	2	2	2
2	1	4	2
1	2	4	1
2	1	1	2
2	1	5	3
1	2	5	4
2	2	3	4
2	1	4	3
1	1	4	1
2	2	2	2
2	2	4	1
2	1	5	2
1	2	4	2
1	1	2	3
1	2	2	4
1	1	4	3
1	1	5	3
2	1	5	1
2	2	3	2
2	2	2	3
1	1	2	4
2	1	1	3
1	2	3	1
2	2	3	1
1	1	5	2
2	2	1	3
1	2	2	4
1	1	5	1

1	1	4	4
1	1	5	4
2	2	3	2
1	1	1	3
2	1	2	1
1	2	4	2
2	1	2	3
1	1	4	3
2	2	5	2
2	1	4	2
1	1	1	4