



**Business Intelligence**

**PUC**  
RIO

*Juliana Quintanilha da Silveira Politi*

*Análise de Categorias: Aplicação da Análise  
de Clusters para apoio à tomada de decisão  
em suprimentos*

*Monografia de Final de Curso*

**17/07/2017**

***Monografia apresentada ao Departamento de Engenharia Elétrica  
da PUC/Rio como parte dos requisitos para a obtenção do título  
de Especialização em Business Intelligence.***

***Orientadores: Leonardo Alfredo Forero Mendonza***

---

## **RESUMO**

A Gestão do Processo de Suprimento de Bens e Serviços tem se tornado uma atividade cada vez mais estratégica para os negócios. O aprimoramento deste processo pode trazer muitos benefícios à organização por meio da utilização mais eficiente de recursos financeiros, bem como influenciar na qualidade do produto final. Para se atingir bons resultados, é importante que haja monitoramento da qualidade do relacionamento com fornecedores e prestadores de serviços. Este estudo apresenta uma análise a respeito de um conjunto de prestadores de serviço de uma do ramo do petróleo, por meio da técnica de clusterização. O algoritmo utilizado é o AGNES (Agglomerative Nesting), o qual consiste em uma abordagem aglomerativa para definição dos diferentes grupos de fornecedores. A análise das características intrínsecas de cada grupo permitirá a definição de estratégias de relacionamento com cada segmento.

Palavras-chave:

Análise de Agrupamento; Clusterização Hierárquica; Aglomerativo; Fornecedores; Cadeia de Suprimentos

## **ABSTRACT**

Managing the Procurement Process has become an increasingly strategic business activity. Improving this process can bring many benefits to the organization through more efficient use of financial resources as well as influence the quality of the final product. To achieve good results, it is important to exist a monitoring system of the relationship quality with suppliers and service providers. This study presents an analysis about a set of service providers from an oil industry business, through the clustering technique. The algorithm used is the AGNES (Agglomerative Nesting), which consists in an agglomerative approach to define the different groups of suppliers. The analysis of the intrinsic characteristics of each group will allow the definition of relationship strategies with each segment.

Keywords

Cluster Analysis, Hierarchy, Agglomerative, Supply Chain Management, Suppliers.

## SUMÁRIO

1.	INTRODUÇÃO .....	5
1.1.	MOTIVAÇÃO .....	5
1.2.	OBJETIVO .....	6
1.3.	ESTRUTURAÇÃO DO TRABALHO .....	6
2.	GESTÃO DA CADEIA DE SUPRIMENTOS .....	7
2.1.	GERENCIAMENTO E DESENVOLVIMENTO DE FORNECEDORES .....	9
2.2.	ETAPAS DO DESENVOLVIMENTO DE FORNECEDORES E PRESTADORES DE SERVIÇO .....	12
2.3.	GESTÃO DO RELACIONAMENTO .....	13
3.	ANÁLISE DE AGRUPAMENTO DE DADOS .....	15
3.1.	PRÉ-PROCESSAMENTO DE DADOS .....	16
3.2.	SELEÇÃO DA MEDIDA DE SIMILARIDADE .....	16
3.3.	EXECUÇÃO DO ALGORITMO DE AGRUPAMENTO .....	18
3.3.	AVALIAÇÃO DE CLUSTERS E VALIDAÇÃO DOS RESULTADOS .....	22
3.4.	INTERPRETAÇÃO DOS RESULTADOS .....	24
4.	ESTUDO DE CASO .....	25
5.	CONSIDERAÇÕES FINAIS .....	31
6.	REFERÊNCIAS BIBLIOGRÁFICAS .....	33

**LISTA DE FIGURAS**

Figura 1- Matriz de Categorias..... 11

Figura 2 - Dendograma.....20

Figura 3 - Dendograma Resultado da Aplicação.....27

Figura 4 - Índice de Silhouette .....27

**LISTA DE TABELAS**

Tabela 1 - Resultado da Clusterização .....28

# **1. INTRODUÇÃO**

Diante do cenário competitivo do mercado atual, as organizações tem buscado a melhoria contínua na qualidade de seus processos internos. Uma das formas de obter vantagem nesta competição é por meio do gerenciamento da cadeia de suprimentos, o qual considera a gestão e a coordenação do fluxo de materiais e informações desde o fornecedor até o consumidor final, buscando a integração de recursos humanos e físicos, com o objetivo de alcançar a satisfação do cliente final.

Neste contexto, a Gestão do Processo de Suprimento de Bens e Serviços torna-se um item estratégico para os negócios, pois impacta diretamente na qualidade do produto final, na eficiência da utilização dos recursos financeiros, assim como no desempenho operacional. O setor responsável nas empresas pela atividade de Suprimento tem como principais responsabilidades a escolha de fornecedores adequados e a negociação de preços.

Na busca pela melhor utilização dos recursos financeiros, muitas empresas estão em constante verificação a respeito de gastos desnecessários, minimização de perdas e eliminação de desperdícios. Para manter a saúde financeira do negócio, um dos caminhos é a redução de custos. A Gestão de Suprimentos pode ser a chave para uma solução diferenciada para que as contratações sejam realizadas de maneira otimizada e inteligente criando, dessa forma, um diferencial competitivo frente à concorrência.

## **1.1. MOTIVAÇÃO**

A motivação para elaboração deste trabalho decorre da identificação das dificuldades encontradas pelas diferentes empresas na segmentação de sua base de fornecedores em diferentes categorias, a fim de possibilitar a definição de ações estratégicas do setor de Suprimentos.

A gestão da base de fornecedores envolve diversas dimensões o que pode trazer ainda mais complexidade na avaliação dos fornecedores com os quais se deve estabelecer um relacionamento mais estreito.

A utilização de métodos de apoio à decisão neste processo torna-se de grande valia para se realizar tais avaliações e auxilia na definição de objetivos e do foco dos planos de suprimentos.

## **1.2. OBJETIVO**

O objetivo deste trabalho foi a aplicação de metodologia de agrupamento, por meio de um algoritmo de aprendizado não supervisionado, demonstrando assim o valor dos métodos de apoio à decisão para à área de Suprimentos. O modelo proposto visa subsidiar decisões para gestão de fornecedores do portfólio por meio da identificação de atributos e características que permitiram determinar o perfil de cada empresa e formas de relacionamento.

## **1.3. ESTRUTURAÇÃO DO TRABALHO**

Este trabalho está dividido em seis seções. No Capítulo 2 é apresentada a contextualização do problema, por meio do esclarecimento de conceitos relativos ao gerenciamento da cadeia de suprimentos.

No capítulo 3, é feita uma breve exposição da base teórica utilizada neste trabalho, iniciando pela descrição da análise de agrupamentos e aprofundando para os principais métodos utilizados para agrupamento ou clusterização.

No capítulo 4, é apresentada aplicação análise de clusters a um problema real, mostrando seus resultados e conclusões,

No capítulo 5 são apresentadas das considerações finais sobre o estudo e a proposição de trabalhos futuros. O capítulo 6 apresenta as referências bibliográficas utilizadas neste estudo.

## 2. GESTÃO DA CADEIA DE SUPRIMENTOS

A gestão da cadeia de suprimentos é a gestão da interconexão das empresas que se relacionam por meio de ligações à montante e à jusante entre os diferentes processos, que produzem valor na forma de produtos e serviços para o consumidor final. (Slack, 2002). Analisar a cadeia de suprimentos como um todo pode aumentar a eficiência, a identificação de gargalos, balanceando capacidade e coordenando o fluxo de materiais.

Segundo o Conselho de Profissionais da Gestão da Cadeia de Suprimentos (em inglês, Council of Supply Chain Management Professionals - CSCMP), a definição de gestão da cadeia de suprimentos é “o planejamento e a gestão de todas as atividades associadas à logística interna e interorganizacional, bem como a coordenação e colaboração entre todos os parceiros da cadeia, sejam eles fornecedores, prestadores de serviço ou consumidores”.

*A área de suprimentos vem adquirindo um papel cada vez mais estratégico à medida que o conceito de gestão da cadeia de suprimentos - Supply Chain Management é trazido à discussão. O foco da gestão de suprimentos se expandiu e passou a incluir também o gerenciamento das interações com seus fornecedores. Recentemente, o termo “gerenciamento da cadeia de valor” tem sido utilizado como uma evolução do conceito de gerenciamento da cadeia de suprimentos. Isto porque o foco da gestão da cadeia de valor é a criação de valor ao usuário final, agregando mais benefícios com menores custos, por meio da maior integração e do maior compartilhamento de informações entre as empresas da cadeia produtiva. (Walters ET Lancaster, 2000).*

A gestão da cadeia deve garantir que os esforços individuais dos participantes, além de visar a objetivos próprios, também visem a objetivos comuns da cadeia (Rodrigues e Sellitto, 2009). Para Christopher (2002), a gestão de cadeias de suprimentos é a extensão, para toda a cadeia externa de fornecimento, do gerenciamento logístico interno. Devido às relações interorganizacionais que surgem, esforços de redução de tempos de atravessamento e de aumento de qualidade têm produzido ganhos de competitividade individuais, que podem resultar em ganhos compartilhados por toda a cadeia.

Ao mesmo tempo que oferecem vantagens quanto ao resultado operacional das empresas participantes, as cadeias de suprimentos ou redes de

fornecedores também representam um desafio para os gestores da cadeia e da empresa focal, devido à influência que o desempenho dos fornecedores exerce sobre a competitividade desta. (Rodrigues e Sellitto, 2009).

A função compras em uma empresa tem como objetivo estabelecer contratos com fornecedores de materiais e serviços. Para o sucesso no desempenho desta função é necessário conhecimento a respeito dos processos da empresa além da capacitação dos fornecedores

A escolha de fornecedores é parte fundamental para a atividade de suprimentos de bens e serviços, devendo ser avaliada a qualificação das empresas disponíveis no mercado para atender às necessidades. No processo de seleção é realizada a coleta e análise das informações sobre fornecedores para permitir a seleção daquele que melhor atenda às exigências identificadas do negócio.

Quanto maior o grau de integração entre fornecedor e cliente, maior a redução de custos ao longo da cadeia. Por isso, cada dia mais a relação empresa/fornecedor vêm tendo um importante papel nos negócios, levando esta relação a ter um constante e acelerado crescimento.

O setor de suprimentos deve buscar aumentar o número de fornecedores em potencial a serem consultados, de maneira que se tenha certeza de que o melhor negócio foi executado em benefício da empresa. O processo de seleção das fontes de fornecimento não se restringe ao momento em que é necessária a aquisição de determinado material. A atividade deve ser exercida de forma permanente e contínua, através de algumas etapas. (Pozo, 2001):

- a) levantamento e pesquisa de mercado: estabelecida a necessidade da aquisição para determinado material, é necessário levantar e pesquisar fornecedores em potencial. O levantamento poderá ser realizado através de cadastros de fornecedores do órgão de compra; edital de convocação; guias comerciais e industriais; catálogos de fornecedores; revistas especializadas; entre outros;
- b) análise e classificação: compreende a análise dos dados cadastrais do fornecedor e a respectiva classificação quanto aos tipos de materiais a fornecer, bem como, a eliminação daqueles fornecedores que não satisfazem as exigências da empresa;
- c) avaliação de desempenho: esta etapa é efetuada pós-cadastramento e nela faz-se o acompanhamento do fornecedor

quanto ao cumprimento do contrato, servindo não raras vezes como elemento de eliminação das empresas fornecedoras.

A necessidade de melhorar o desempenho dos fornecedores tem tornado relevante a proposição de métodos para a sua medição (Canto, 2004). O resultado de tais medições definirá as ações que serão tomadas no seu relacionamento com o fornecedor. O fornecedor, por sua parte, deve monitorar os mesmos fatores utilizados pelo cliente para ter certeza de que os está satisfazendo (Kotler et Armstrong, 1998). Canto (2004) argumenta que os indicadores devem ser capazes de medir os aspectos importantes do fornecimento e constituem uma forma coesa de avaliar o desempenho dos fornecedores ao longo do processo de suprimentos.

## **2.1. GERENCIAMENTO E DESENVOLVIMENTO DE FORNECEDORES**

A evolução dos conceitos de gerenciamento da cadeia de suprimentos trouxe a visão de que para se manter competitiva no mercado, cada empresa deve construir e manter relações com fornecedores competentes e extrair o maior valor possível destas relações. Dessa forma, a competência especializada dos fornecedores pode ter uma influência substancial na capacidade inovadora da empresa compradora e na sua habilidade de oferecer produtos com alta qualidade, contribuindo para elevação de suas vantagens competitivas. (Braga, 2009)

Dependendo do nicho de mercado em que a empresa atua, os gastos com suprimento por ser significativos em relação ao custo final de produção, reforçando a necessidade de integração com esses fornecedores e prestadores de serviço, por meio do estabelecimento de parcerias de negócio.

O sucesso de iniciativa para desenvolvimento de mercado fornecedor depende de diversos fatores. Dentre eles, tem-se o envolvimento da alta gerência da empresa, da avaliação do desempenho dos fornecedores, emprego de equipes multifuncionais, ampla comunicação entre as partes envolvidas, perspectiva de longo prazo e foco no custo total de produção.

O envolvimento da alta gerência é determinante porque muitas interações com os fornecedores poderão ocorrer no nível mais elevado das organizações, devido a decisões estratégicas que podem ocorrer entre os parceiros. Além disso, muitas ações a serem executadas nos níveis funcionais dependerão de recursos para sua execução.

Quanto ao sistema de avaliação de fornecedores, além de produzir melhorias no desempenho dos fornecedores/prestadores de serviços por si só, é importante para comunicar aos fornecedores quais foram os resultados por eles obtidos quando da avaliação efetuada, trazendo transparência e credibilidade à relação.

Neste trabalho de desenvolvimento, a existência de equipes multifuncionais é fundamental, pois certamente envolverá questões ligadas à engenharia, qualidade, finanças e outras funções.

O conceito de custo total de propriedade também é importante para orientar as decisões sobre o programa de desenvolvimento de fornecedores, devido à necessidade de se considerar todos os custos envolvidos no programa e comparar com as expectativas de ganhos que se possa ter com este empreendimento.

Os primeiros resultados do programa de gerenciamento da base de fornecedores pode ser a sua redução. Na medida em que as parcerias entre contratante e prestador de serviço/fornecedor evoluem, são mantidos apenas aqueles fornecedores/prestadores de serviço que podem fazer diferença no negócio da empresa. Por outro lado, com uma base menor de fornecedores, o esforço de gerenciamento também é reduzido e modificado para estimular relacionamentos cada vez mais estreitos entre as empresas mantendo uma perspectiva de longo prazo.

Um exemplo de integração com fornecedores foi o modelo empresarial Kereitsu, desenvolvido pela indústria automobilística no Japão. Neste modelo, há uma coalizão de empresas unidas por certos interesses econômicos que neste caso era decorrente da aquisição de uma fatia da empresa fornecedora pela montadora.

Além da manutenção e do estreitamento do relacionamento com os fornecedores/prestadores de serviço que já fazem parte da base de fornecedores, a identificação de novas potenciais fontes de suprimentos é uma atividade que deve ser cuidadosa, especialmente quando as fontes de fornecimento são poucas ou quase nenhuma. Estabelecer iniciativas para garantir a inteligência de mercado, pode facilitar muito a tarefa de busca de novos parceiros.

Para analisar o mercador fornecedor potencial, assim como a base de fornecedores existente na empresa, há alguns fatores que podem ser observados. Dentre eles, é importante avaliar a disponibilidade de recursos financeiros do prestador de serviço/fornecedor, principalmente em tempos de

crise financeira. A facilidade na captação de recursos financeiros pode causar reflexos na capacidade de aquisição de máquinas e tecnologia para acompanhar o crescimento do contratante, suportar esforços de pesquisa e desenvolvimento de novos produtos ou processos, manter os níveis atuais de recursos de informação tecnológica ou até mesmo investir em estoques para suportar as operações em andamento com o contratante (Braga, 2009).

O nível de desenvolvimento técnico está relacionado à constante preocupação das empresas por inovação dos seus produtos e melhorias de qualidade para permanecerem competitivas. Novas tecnologias podem ser obtidas através de desenvolvimento interno ou em conjunto com fornecedores que realizam o investimento de pesquisa e desenvolvimento.

Por se tratar de um processo custoso para a organização, a definição de quais produtos e serviços que devem ser contemplados na investigação do mercado fornecedor, deve ser avaliada com base na identificação dos itens mais importantes para empresas, que mereçam os esforços da pesquisa sofisticada que será realizada. Uma alternativa bastante utilizada é a matriz de categorias, podendo cada linha de produtos ser enquadrada em um dos quatro grandes quadrantes da matriz (adaptação Kraljic, 1983).

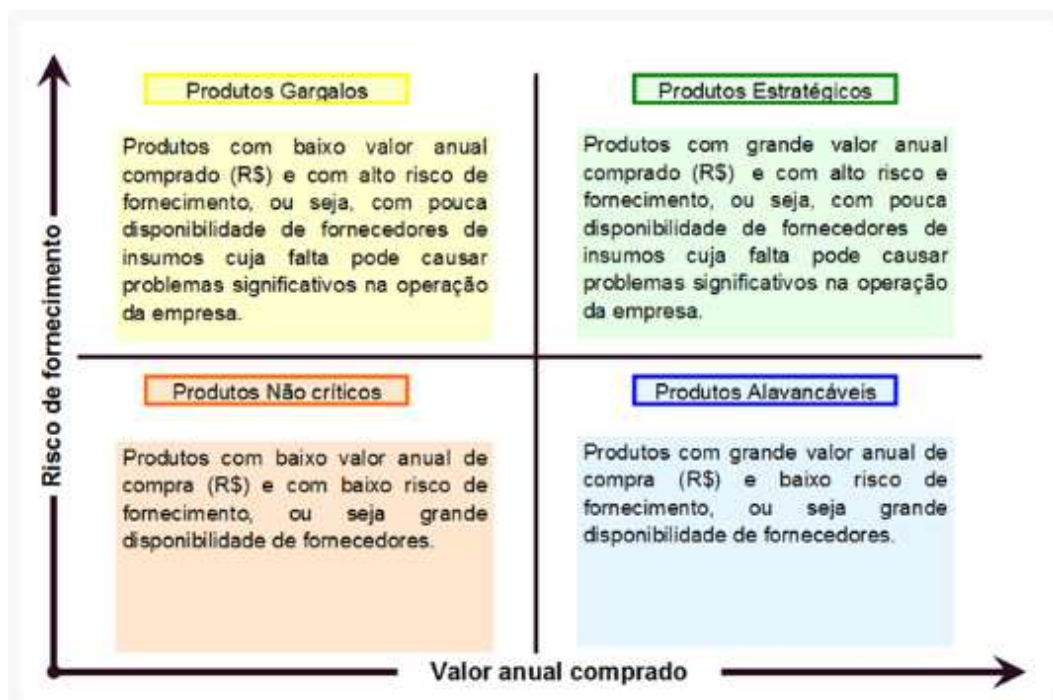


Figura 1- Matriz de Categorias

Desse modo, o esforço de inteligência sobre o mercado fornecedor deve-se iniciar pelos os itens estratégicos, depois com os gargalos e, finalmente, com os itens de alavancagem. Não se justifica o emprego deste levantamento de dados para os itens considerados como não-críticos, devido a sua pouca

representatividade e baixo risco de fornecimento. Para estes últimos, a pesquisa simplificada utilizando as fontes de informações tradicionais atende aos objetivos do processo de suprimento.

## **2.2. ETAPAS DO DESENVOLVIMENTO DE FORNECEDORES E PRESTADORES DE SERVIÇO**

O estabelecimento de um plano para desenvolvimento da base de fornecedores e prestadores de serviço é uma tarefa que precisa ser tratada de forma estruturada, de modo a reduzir-se o risco do insucesso. A seguir estão descritas de maneira sucinta as principais etapas que podem compor este plano:

### **1) Identificação da importância relativa dos produtos/serviços**

Conforme já descrito no item 2.1. , a primeira etapa deste roteiro trata da identificação dos itens para os quais qualquer alteração no fornecimento tem reflexos significativos sobre as operações da empresa. A classificação das famílias ou categorias de produtos/serviços, por meio da matriz de categorias, pode ser realizada neste processo de identificação.

### **2) Medição dos fornecedores e identificação dos candidatos ao programa**

Esta etapa busca identificar, dentro de cada categoria de produtos/serviços, os fornecedores que apresentam desempenho abaixo de um determinado nível mínimo estabelecido pela empresa contratante. Para isso, há a necessidade de se avaliar os fornecedores em determinados parâmetros como capacidade técnica, qualidade, capacidade gerencial, responsabilidade ambiental e saúde financeira, dentre outros quesitos.

Há diversas metodologias possíveis para se fazer esta avaliação, dentre elas a análise de agrupamentos, descritas no item 3 deste trabalho. Esta classificação pode ser utilizada também para auxiliar a empresa a racionalizar sua base de fornecedores/prestadores de serviços, eliminando aqueles que apresentem desempenho bem abaixo do nível mínimo esperado.

### **3) Identificação das áreas de interesse**

A etapa anterior permitiu identificar a área na qual há deficiência nos fornecedores/prestadores de serviços da base. Dando sequência ao processo, o

propósito desta é especificar a causa do problema identificado. Os problemas de desempenho do fornecedor/prestador de serviço podem ocorrer nas dimensões de capacidades requeridas, porém, estes mesmos problemas também podem ser enquadrados em fontes de recursos da empresa fornecedora, como, por exemplo, produto, processos ou sistemas operacionais.

Depois de identificadas as causas dos problemas e definidas as áreas de desenvolvimento, são realizadas as estimativas de custos, metas e cronogramas que suportarão a defesa dos mesmos nas esferas de decisão da empresa compradora.

#### 4) Análise de risco

Nesta fase, são realizadas as análises financeiras dos possíveis projetos planejados na fase anterior, definição de expectativas de ganhos com os resultados a serem obtidos com determinado projeto e estabelecimento do mecanismo de compartilhamento dos custos com o fornecedor.

#### 5) Definição com o fornecedor sobre os projetos a serem implementados

Após a identificação de oportunidades e análises financeiras realizadas, a alta gerência do contratante e do fornecedor/prestador de serviço buscam a definição dos trabalhos a serem executados. Nesta etapa, são avaliados alguns aspectos como viabilidade dos projetos, metas a serem cumpridas, programa de trabalho para a implementação de mudanças, importância estratégica das mudanças e impactos no negócio.

### **2.3. GESTÃO DO RELACIONAMENTO**

A importância da gestão do relacionamento para o sucesso no gerenciamento e desenvolvimento de fornecedores é muito grande, sendo responsável por grande parte do seu sucesso. A gestão de relacionamento exige paciência e persistência, com alta frequência de comunicação e um forte programa de avaliação de desempenho e feedback com o fornecedor/prestador de serviços.

A gestão de relacionamentos colaborativos com fornecedores/prestadores de serviços pode ser efetuada através de diversas estratégias de suprimentos. O alto nível de comprometimento em superar os

problemas de desempenho, buscar alternativas de redução de custo, resolver problemas de qualidade, trazer inovações tecnológicas e também criar produtos e serviços não disponíveis no mercado pode ser a chave deste bom de uma boa relação.

Adicionalmente, as empresas precisam ter uma comunicação verdadeira com seus fornecedores a respeito do desempenho deles. Se os resultados das medições de desempenho e as expectativas de ações não forem comunicados aos fornecedores, estas ações não acontecerão. É muito comum as empresas enviarem os resultados de desempenho sem nenhum comentário sobre as próximas etapas para melhoria contínua, destruindo parte do propósito do processo.

O relacionamento cooperativo é um relacionamento contínuo, planejado, baseado na confiança mútua, podendo em alguns casos ser denominado de parceria.

A parceria de suprimentos foi definida pela Confederation of British Industry e o Department of Trade and Industry (apud Baily et al., 2000, p. 205), como o compromisso entre clientes/fornecedores, independente da dimensão, com um relacionamento a longo prazo baseado em objetivos claros e mutuamente definidos, visando ao esforço em termos de capacidade e competitividade mundial.

Estes relacionamentos de parceria exigem o desenvolvimento de contratos de compra de longo prazo. Para os fornecedores/prestadores de serviço, uma das principais vantagens recebidas é uma posição de destaque em novos programas de desenvolvimento em futuras decisões de suprimento.

### 3. ANÁLISE DE AGRUPAMENTO DE DADOS

A análise multivariada de dados é um conjunto de técnicas matemáticas capaz de processar dados de diversas variáveis simultaneamente, gerando informações úteis em processos decisórios (Hair et al., 1998). Fazem parte deste conjunto diferentes técnicas, cada uma com sua fundamentação teórica e sua aplicabilidade.

Quando o interesse é analisar como o conjunto de dados se relaciona, quais são as semelhanças e o grau desta, podem-se utilizar os métodos de análise de agrupamento, ou análise de clusters.

A clusterização é a classificação não-supervisionada de dados, formando agrupamentos ou clusters. Esta análise envolve a organização de um conjunto de padrões em clusters (ou grupos), de acordo com alguma medida de similaridade. Dessa forma, elementos pertencentes a um dado grupo devem possuir mais semelhanças entre si do que em relação aos elementos pertencentes aos outros grupos.

Diferentemente da classificação supervisionada, em que são fornecidos elementos pré-classificados e o problema a ser resolvido é rotular novos elementos, na classificação não-supervisionada, o problema é agrupar um conjunto de padrões não-rotulados em clusters que apresentem alguma propriedade comum.

Um problema de clusterização pode ser definido formalmente da seguinte forma: existe um conjunto  $X$  com  $n$  elementos  $X = \{X_1, X_2, \dots, X_n\}$ . O objetivo das técnicas é a obtenção de um conjunto de  $k$  clusters,  $C = \{C_1, C_2, \dots, C_k\}$ , tal que os elementos contidos em um cluster  $C_i$  possuam uma maior similaridade entre si do que com os elementos de qualquer um dos demais clusters do conjunto  $C$ . O conjunto  $C$  é considerado uma clusterização com  $k$  clusters caso as seguintes condições sejam satisfeitas:

$$\bigcup_{i=1}^k C_i = X$$

$$C_i \neq \emptyset, \text{ para } 1 \leq i \leq k$$

$$C_i \cap C_j = \emptyset, \text{ para } 1 \leq i, j \leq k \text{ e } i \neq j$$

Por essas condições, um elemento não pode pertencer a mais de um cluster e cada cluster tem que ter ao menos um elemento.

A clusterização possibilita a identificação de novas hipóteses a respeito dos inter-relacionamentos dos dados e de sua estrutura. As etapas que envolvem um processo de clusterização podem ser detalhadas no item a seguir:

### **3.1. PRÉ-PROCESSAMENTO DE DADOS**

Esta etapa é constituída pela preparação e ajuste dos dados para a realização do agrupamento. Os tipos de atributos influenciam no processo de clusterização e eles podem ser binários, discretos ou contínuos. Atributos binários indicam a presença ou não de determinada característica. Os atributos discretos possuem um conjunto finito de valores possíveis. Os atributos contínuos podem assumir qualquer valor real dentro de um intervalo pré-definido.

Os agrupamentos são influenciados também pela escala nos valores dos atributos. A escala define a significância relativa dos valores dos atributos.

O pré-processamento dos dados envolve a normalização dos mesmos, que deve considerar a medida de similaridade que será aplicada. No caso da distância euclidiana, utilizada em muitos casos, é atribuída maior importância aos atributos que podem assumir valores em um intervalo maior que aqueles com menor intervalo de valores. Com o objetivo de minimizar este viés, a normalização pode ser usada para manter os atributos dentro de uma faixa de valores comum.

O sistema de coordenadas em que os dados estão representados também é importante para o processo de clusterização. Para alguns grupos de dados, o sistema de coordenadas cartesianas pode não ser o mais adequado, podendo ser utilizado o sistema de coordenadas polares, por exemplo.

A dimensionalidade do conjunto de dados também tem influência na clusterização. Um número menor de atributos reduz consideravelmente a complexidade dos algoritmos e o tempo de execução dos mesmos. Quando o conjunto de dados possui muitos atributos, podem ser utilizados métodos de seleção de atributos para avaliar aqueles que melhor descrevem o conjunto de dados, melhorando a eficiência dos algoritmos.

### **3.2. SELEÇÃO DA MEDIDA DE SIMILARIDADE**

A medida de similaridade deve ser escolhida com muita cautela considerando as características do conjunto de dados, como tipo e escala dos

atributos. Há na literatura muitas propostas de medidas de similaridade, dentre as quais estão as medidas de distância, de correlação e de associação.

As medidas de distância e de correlação são aplicáveis à atributos contínuos, enquanto que as medidas de associação são utilizadas para atributos discretos.

As medidas de distância, na realidade, são uma medida de dissimilaridade, pois quanto maior o valor da distância, menor o nível de semelhança entre os elementos e quanto menor a distância, maior a similaridade fornecida por uma função de distância definida entre pares de dados ou padrões.

Há diversas maneiras de calcular a distância entre dois elementos em um espaço multidimensional. Algumas formas de calcular estas medidas estão brevemente descritas abaixo:

- 1) Distância de Manhattan: é a distância entre dois pontos no espaço euclidiano, como sendo a soma dos comprimentos das projeções dos segmentos de reta entre os pontos dos eixos das coordenadas, ou seja, a soma das diferenças absolutas dos valores dos atributos.
- 2) Distância Euclidiana: é a medida mais utilizada em clusterização. É definida como a soma da raiz quadrada da diferença entre as coordenadas dos dois elementos, no espaço euclidiano.
- 3) Distância de Chebyshev: é definida pela máxima diferença absoluta entre os valores dos atributos. Ela é apropriada para os casos que dois elementos são considerados diferentes caso possuam qualquer atributos diferente.

As medidas de correlação consideradas também os valores de diversos atributos, porém consideram o padrão desses valores e não sua magnitude. Algumas das medidas de correlação foram desenvolvidas, algumas delas estão apresentadas a seguir.

- 1) Correlação de Pearson: mede o grau de correlação e a direção dessa correlação entre duas variáveis e assume valores entre  $[-1,1]$ . Se este índice for igual a 1 significa correlação perfeita e igual a -1, indica uma correlação inversa, ou seja se uma das variáveis

aumenta a outra diminui. Se igual a zero, significa que não há dependência linear entre as variáveis.

- 2) Correlação de Spearman: considera a substituição dos valores dos atributos dos elementos por seus índices relativos a sua magnitude. A correlação é calculado sobre os vetores a partir desses índices.

### **3.3. EXECUÇÃO DO ALGORITMO DE AGRUPAMENTO**

Os algoritmos utilizados para o agrupamento dos dados podem ser classificados de acordo com a abordagem utilizada na definição dos clusters: particionamento, redes-auto-organizáveis, baseado em densidade, hierárquico ou baseado em grafos.

#### **3.2.1. Algoritmos de Clusterização por Particionamento**

Neste tipo de clusterização o conjunto de dados é dividido em um número determinado de clusters. A escolha do número de cluster varia conforme o problema a ser abordado e pode interferir na eficiência do algoritmo. O objetivo é maximizar a similaridade entre elementos de um mesmo cluster e minimizar a similaridade entre elementos de clusters diferentes.

É estabelecida uma função objetivo, a qual pode considerar critérios locais (a partir de informações estruturais dos dados) ou critérios globais (utiliza um elemento do conjunto como representante de cada cluster e classifica os demais dados com base na similaridade).

Um dos algoritmos mais conhecidos que utiliza a técnica do particionamento é “K-means”. Neste algoritmo, em cada cluster é definido um elemento para ser o “centro”. A função objetivo do algoritmo é minimizar o erro quadrático da variação entre os dados em relação ao centro de cada cluster.

A definição do número de clusters ideal pode ser facilitada quando se tem um maior conhecimento sobre a situação em estudo, porém pode ser feita a partir da execução do algoritmo considerando diferentes números para K.

A partir da definição dos centros, a semelhança entre um dado e um cluster é calculada a partir da distância entre os dois. O elemento é incluído no cluster com o qual tiver a menor distância. Em seguida, os centros de cada cluster são recalculados para representarem adequadamente cada cluster. Em seguida são avaliadas as distâncias entre o novo centro e os demais membros

do grupo, buscando minimizá-las até que não haja mais mudança nos centros dos clusters.

### **3.2.2. Algoritmos de Clusterização por Densidade**

Um dado com  $d$  atributos pode ser representado como um ponto em um espaço  $d$ -dimensional e os clusters correspondem a sub-conjuntos de dados que estejam próximos. Dessa forma, os clusters localizam-se em regiões de maior densidade no espaço de métricas e são separados por regiões de baixa densidade. Estes métodos utilizam critério de clusterização local, por considerarem a densidade de ligações entre os dados. A possibilidade de encontrar clusters de formas arbitrárias e o fato de não precisar da definição do número de clusters como parâmetro inicial são as principais vantagens dos métodos baseados em densidade.

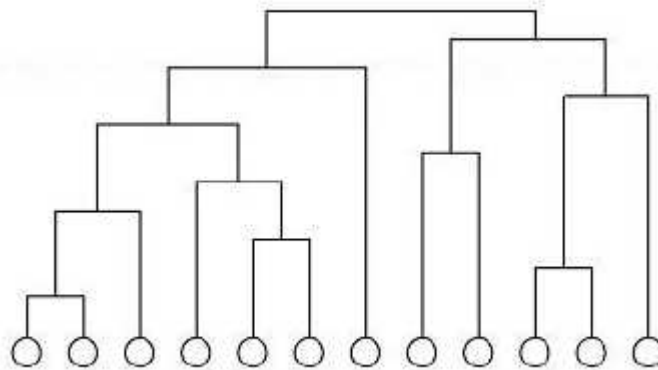
Jain (1988) foi o primeiro a explorar uma abordagem baseada em densidade para identificar clusters em dados espaciais. No método Jain o conjunto de dados é particionado para um número de células que não se sobrepõem e os histogramas são construídos. Células com contagem relativamente alta frequência de pontos são os centros dos grupos potenciais e as fronteiras entre grupos caem nos "vales" do histograma. É um método que identifica clusters de qualquer forma, mas a memória requisitada e o tempo de execução para o armazenamento e a pesquisa nos histogramas multidimensionais podem ser enormes.

O algoritmo DBSCAN (Density Based Spatial Clustering of Applications with Noise), desenvolvido por Ester et al. (1996), utiliza uma abordagem baseada em centro. Cada elemento de um cluster tem uma vizinhança de um determinado raio que contenha um número mínimo de pontos. Após a identificação das regiões densas, os pontos podem ser classificados de acordo com sua posição em relação à mesma: ponto central ou ponto periférico. Se dois pontos centrais são próximos, eles são colocados no mesmo cluster.

### **3.2.3. Algoritmos de Clusterização Hierárquica**

Estes algoritmos organizam um conjunto de dados em uma estrutura hierárquica de acordo com a proximidade entre os indivíduos. Como resultado da execução de um algoritmo de clusterização hierárquica obtém-se uma árvore

binária ou dendograma. Esta árvore divide, de maneira iterativa, a base de dados em subconjuntos menores, conforme pode ser visto na figura a seguir:



**Figura 2 - Dendograma**

Esta representação fornece descrições informativas e visualização para as estruturas de grupos em potencial. A raiz do dendograma representa o conjunto de dados inteiro e as extremidades (nós-folhas) representam os elementos. Ao cortar o dendograma em diferentes níveis de é possível avaliar o número adequado de clusters para este conjunto de dados.

No eixo vertical do dendograma estão representadas as distâncias entre os elementos ou grupos de elementos, isto é, indica o nível de similaridade. No eixo horizontal são marcados os indivíduos. As linhas verticais partem dos indivíduos, e têm altura correspondente ao nível em que os indivíduos são considerados semelhantes. O *dendrograma* pode ser formado de duas maneiras:

- Abordagem aglomerativa (bottom-up): Considera-se, inicialmente, cada elemento como sendo um cluster, totalizando  $n$  clusters. Em cada etapa, calcula-se a distância entre cada par de clusters. Verifica-se, então, dois clusters com a menor distância e é feita a união dos mesmos. Este processo continua até que todos os elementos estejam em um único cluster (o nível mais alto da hierarquia).
- Abordagem divisiva (top-down): o processo é o inverso da abordagem bottom-up. O início considera todos os elementos em um único cluster. Em cada etapa, um cluster é escolhido e dividido em dois clusters menores. Este processo continua até que se tenham  $n$  clusters.

Os métodos aglomerativos são de uso mais comum. Estes são constituídos de em sua maioria em métodos de encadeamento. Os métodos de encadeamento compreendem:

- Método do encadeamento único (Single Linkage), que se baseia na distância mínima, regra do vizinho mais próximo. Os dois primeiros objetos agrupados são os que apresentam menor distância entre si.. Em cada estágio a distância entre dois grupos é definida como a distância entre seus dois pontos mais próximos.
- Método do encadeamento completo (Complete Linkage) é semelhante ao encadeamento único, embora se baseie na distância máxima entre os objetos ou o método do vizinho mais afastado. Neste, a distância entre dois grupos é calculada entre seus dois pontos mais afastados.

Os estudos de Kauffman e Rosseeuw (1990) apresentaram diversos métodos de clusterização, dentre eles dois métodos hierárquicos divisivos DIANA (Divisive ANALysis) e MONA (MONotetic Analysis) e um método hierárquico aglomerativo, o AGNES (AGlomerative NESTing), o qual será utilizado neste estudo.

O algoritmo AGNES, é um dos algoritmos de clusterização mais usados e assim como outros algoritmos aglomerativos, é executado a partir de uma série de sucessivas fusões entre clusters. O par de clusters mais semelhantes é identificado a partir da matriz de similaridade e selecionado para agrupamento. Caso mais de um par de clusters seja identificado como mais semelhante, a seleção aleatória entre esses pares é utilizada para determinar qual deles será agrupado na iteração corrente.

#### **3.2.4. Redes Auto-Organizáveis**

Um tipo específico de rede neural, a rede auto-organizável (Self Organizing Maps – SOM), desenvolvida por Kohonen (1982) pode separar dados em grupos desconhecidos previamente. Estas redes são compostas por um conjunto de neurônios, onde cada dado tem seus atributos ligados a todos os neurônios da rede. O aprendizado ocorre conforme os dados são apresentados à rede.

Um elemento serve de protótipo do cluster e pode ser definido como um exemplar deste cluster. Novos elementos podem ser distribuídos para clusters cujo elemento é mais similar baseado em alguma medida de distância. Os atributos de um elemento atribuído a um cluster podem ser preditos dos atributos do exemplar do cluster.

As redes neurais apresentam uma série de vantagens na clusterização como robustez ao ruído, capacidade de generalização, aprendizado adaptativo a partir de exemplos e processamento paralelo.

### **3.2.5. Clusterização baseada em Grafos**

Os algoritmos de clusterização baseados em grafos buscam representar um conjunto de dados em um grafo, onde cada vértice representa um elemento do conjunto de dados e a existência de uma aresta conectando dois vértices é feita com base na proximidade entre os dois dados. A maneira mais simples de estabelecer as ligações entre os vértices é conectar cada vértice aos vértices restantes, onde o peso indica a similaridade entre os dois dados e um cluster é definido como um subgrafo do grafo inicial. Para tal, adota-se uma medida de similaridade no processo de agrupamento, o que pode fazer com que o algoritmo apresente alguma dificuldade em determinar clusters de formas variadas.

Embora esta técnica possibilite a obtenção de informações a partir de uma base de dados, depende da avaliação e validação dos dados por parte do decisor.

## **3.3. AVALIAÇÃO DE CLUSTERS E VALIDAÇÃO DOS RESULTADOS**

Avaliar os grupos formados pelo algoritmo e validar estes resultados pode ser uma tarefa de alta complexidade. O fato de não haver uma classe pré-definida para os dados traz a necessidade da utilização de heurísticas e suposições para a identificação das estruturas embutidas nestes dados, que devem levar em consideração a qualidade dos cluster encontrados. Esta qualidade está associada ao fato de cada cluster apresentar corretamente uma informação.

A validação do agrupamento pode ser realizada com base em índices estatísticos que verificam, de maneira quantitativa, as estruturas encontradas.

Os critérios para avaliar os grupos formados podem ser relativos, internos ou externos.

Os critérios internos e externos tem como objetivo medir o quanto o resultado obtido confirma uma hipótese pré-estabelecida. Dessa forma, são utilizados testes de hipóteses para determinar se uma estrutura obtida é apropriada para os dados.

O critério externo testa se os pontos do conjunto de padrões são aleatoriamente estruturados ou não. Essa análise é baseada na hipótese nula expressada como uma representação de uma estrutura aleatória de um conjunto de padrões. Segundo Liu et al., (2010) um exemplo da validação externa é a entropia, que avalia a “pureza” dos grupos com base nos rótulos das classes.

O critério interno tem como objetivo avaliar o agrupamento resultante de um algoritmo utilizando somente quantidades e características inerentes ao conjunto de padrões. Diferente do critério externo, no critério interno o número de grupos é desconhecido, uma vez que os padrões não precisam estar rotulados.

Os critérios relativos verificam o melhor agrupamento que um algoritmo pode obter ou se o tipo de algoritmo utilizado é o mais adequado para o conjunto de dados em análise.

O índice Silhouette, que será empregado neste estudo, foi proposto por Rousseeuw (1987) e tem como objetivo avaliar o resultado de um algoritmo de agrupamento de acordo com o quanto os grupos formados são compactos e bem separados entre si. Segundo Rousseeuw (1987), esse índice é apropriado para 32 medidas de proximidade que utilizam escala proporcional (ratio scale) como é o caso da distância Euclidiana. Esse índice utiliza as médias de proximidades para avaliar o agrupamento. O índice Silhouette é dado pela equação:

$$Silhouette = \frac{1}{NG} \sum_l \left\{ \frac{1}{|G_l|} \sum_{P_i \in G_l} \frac{b(P_i) - a(P_i)}{\max[b(P_i), a(P_i)]} \right\}$$

A função  $a(P_i)$  dada pela equação abaixo calcula a distância média entre o padrão  $P_i$  e os demais padrões pertencentes ao mesmo grupo. A função  $b(P_i)$  dada pela equação a seguir calcula as distâncias médias entre o padrão  $P_i$  e os padrões pertencentes a cada um dos grupos que  $P_i$  não pertence e seleciona aquela com menor valor.

$$a(P_i) = \frac{1}{|G_i| - 1} \sum_{P_j \in G_i} d(P_i, P_j)$$

$$b(P_i) = \min_{j, j \neq i} \left[ \frac{1}{|G_j|} \sum_{P_j \in G_j} d(P_i, P_j) \right]$$

Os valores de índice Silhouette mais altos significam melhores configurações de grupos. Os valores do índice Silhouette variam no intervalo [-1,1].

### 3.4. INTERPRETAÇÃO DOS RESULTADOS

A última fase do processo de clusterização consiste em examinar os clusters encontrados pelo algoritmo para buscar significados relacionados ao contexto da aplicação. Esta etapa exige especialmente da participação de um especialista no tema para que sejam identificados os conceitos que explicam os elementos pertencentes a um mesmo cluster. Dessa forma, o especialista avalia de maneira subjetiva os grupos encontrados, buscando o significado prático da estrutura definida.

Esta etapa é de alta complexidade, pois um algoritmo de clusterização sempre encontrará grupos em conjuntos de dados, independentemente dos conceitos envolvidos. Por outro lado, os clusters nem sempre descrevem um agrupamento adequado ou estão relacionados ao objetivo da aplicação.

## 4. ESTUDO DE CASO

A intensidade do relacionamento entre empresa contratante e o prestador de serviços pode variar de um nível muito baixo de interação a um nível de interação máximo, que corresponde a uma integração vertical entre empresas.

A gestão de uma parceria requer tempo e dedicação das pessoas envolvidas no gerenciamento das atividades terceirizadas. Por isso, esse investimento deve ter como contrapartida a perspectiva de um retorno que justifique o esforço. (Ribeiro, 2002)

A fim de avaliar o nível de interação adequado entre contratante e contratado, as empresas envolvidas devem avaliar se existem motivos suficientes para formação da parceria e se os ambientes e filosofias gerenciais são compatíveis. Embora os motivadores para a formação das parcerias possam ser diferentes, é importante que cada empresa conheça os motivadores da outra a fim de que haja transparência no relacionamento. Existem ainda características culturais que são de difícil mudança no curto prazo e ao mesmo tempo cruciais para garantir os ganhos da parceria.

Os motivadores que estimulam a contratação do terceiro podem ser a redução de ativos, de custos operacionais, aumento da cobertura geográfica, melhoria da qualidade dos serviços, entre outros. Há também as características do prestador de serviços consideradas importantes em um relacionamento, dentre elas, a disposição para investir conforme necessidade do cliente, capacidade de compartilhar custos e benefícios, ética, capacidade de inovar, flexibilidade negocial, solidez financeira e nível tecnológico

Segundo Ribeiro (2002), o nível de interação desejado pode ser uma parceria de curto prazo, parceria de longo prazo ou parceria sem prazo limite. O agrupamento dos prestadores de serviço possibilitará avaliar o melhor nível de interação e definirá questões como planejamento gerencial, controles operacionais, troca de informações e amplitude das atividades terceirizadas.

Para identificar o nível de interação ideal do relacionamento entre contratante e contratada, este estudo propõem uma metodologia para análise de fornecedores/prestadores de serviço em uma empresa da indústria do petróleo, transformando dados cadastrais em inteligência operacional. A metodologia proposta e testada é baseada na técnica de análise de clusters (*cluster analysis*).

Uma metodologia de agrupamento de fornecedores pode ajudar a definir formas de relacionamento e a selecionar a que fornecedores cada política deve ser destinada.

O conjunto de dados analisado foi obtido por meio do banco de dados da empresa, o qual concentra uma série de informações de cunho técnico, financeiro e de sistemas de gestão de cada prestador de serviço. Foram selecionados 67 prestadores de serviços de um mesmo nicho de mercado.

Neste estudo, foram considerados como atributos para o agrupamento o nível de qualificação técnica (variando de 0 a 10), o patrimônio líquido das empresas (em R\$) e o nível de qualidade dos sistemas de gestão de meio ambiente (variando de 0 a 10), segurança e saúde. A seleção dos atributos considerou a importância dos mesmos para a empresa contratante e buscou eliminar redundâncias, trazendo maior simplicidade ao modelo.

O objetivo da aplicação é definir grupos/cluster formados por prestadores de serviço que possuem características semelhantes em relação aos atributos selecionados.

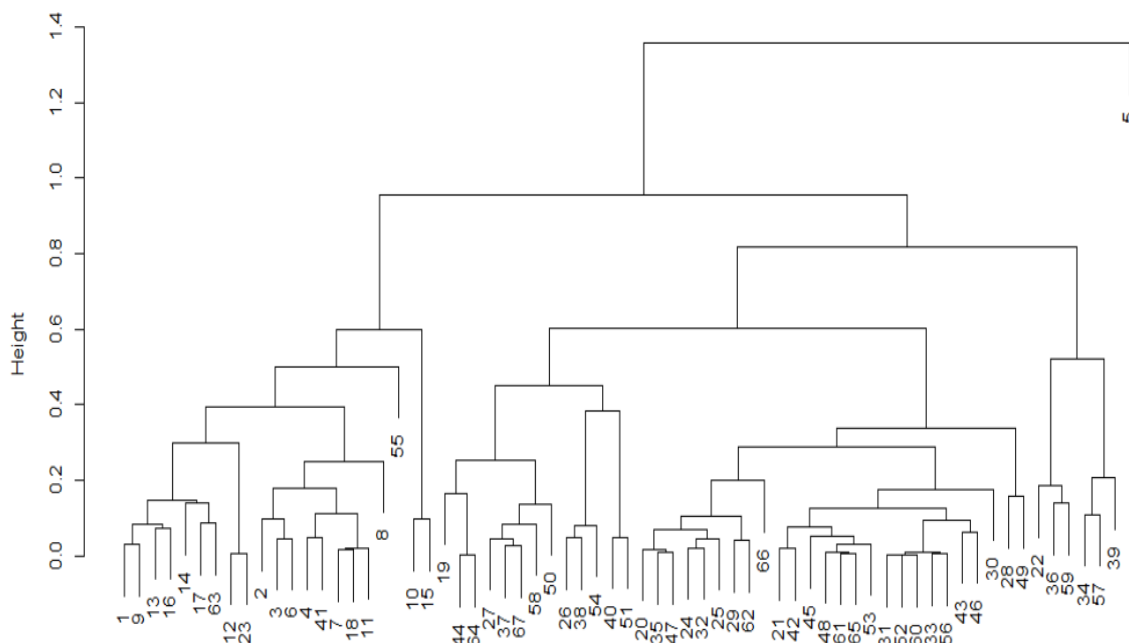
Em relação à escolha do modelo para agrupamento, escolheu-se o método de classificação não supervisionado, com algoritmo do tipo aglomerativo AGNES (Agglomerative Nesting). Como medida de similaridade foi utilizada a distância euclidiana. Para mitigar os impactos relativos à diferença de escala entre os atributos, foi realizado pré-processamento dos dados por meio da normalização dos mesmos.

Após a execução do algoritmo, foi necessária aplicação de métodos auxiliares de validação dos clusters formados, avaliando qual seria o número de grupos mais adequado a este conjunto de dados. Nesta etapa foi utilizado o Índice de Silhouette.

A interpretação dos resultados contou com o conhecimento de especialista na área de Suprimento de Bens e Serviços, o que contribuiu para melhor identificação das características dos grupos formados.

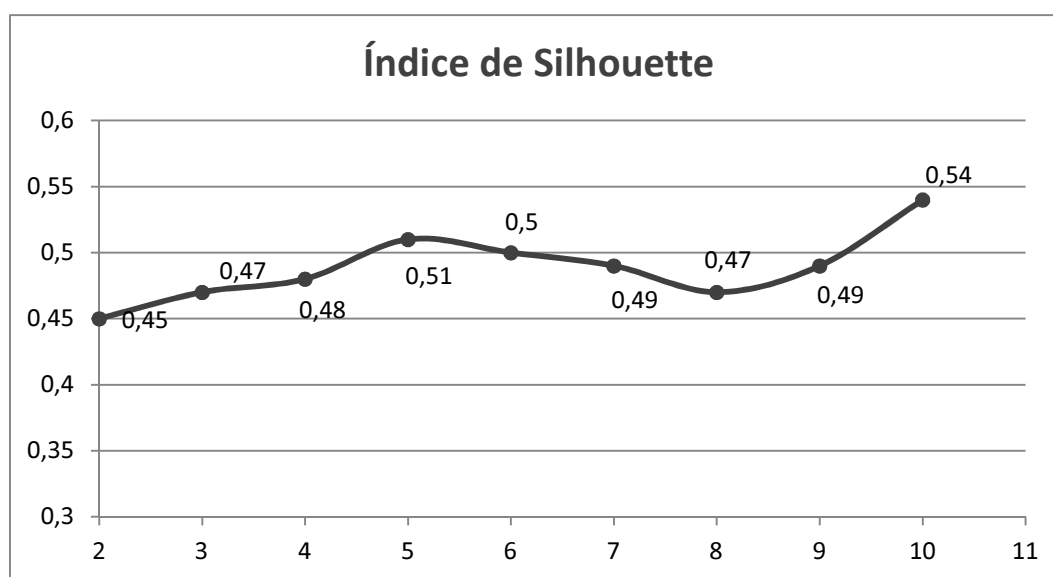
Para realização da clusterização foi utilizado o software R Studio versão 0.99.903, com o auxílio do pacote de “cluster”. Os resultados obtidos foram satisfatórios e possibilitaram melhor conhecimento a respeito das informações armazenadas no cadastro da empresa, podendo servir de subsídio para atuação estratégica da área de Suprimentos.

Após a execução do algoritmo, foi obtido dendograma a seguir:



**Figura 3 - Dendrograma Resultado da Aplicação**

Em seguida, foi calculado o Índice de Silhouette de cada agrupamento possível, variando de 2 a 10 grupos, cujo resultado pode ser observado no gráfico abaixo:



**Figura 4 - Índice de Silhouette**

A análise do gráfico acima permitiu definir junto ao especialista o número de cluster ideal para este conjunto de dados. Os maiores valores do índice ocorreram para cinco e dez grupos. Para facilitar a análise e definição das estratégias de relacionamento, considerou-se que um menor o número de

clusters seria mais adequado ao contexto da aplicação. Por este motivo, optou-se pelo agrupamento em cinco diferentes clusters.

A composição destes clusters e as características intrínsecas de cada um deles foram interpretadas a partir dos atributos que foram alvo da análise. A tabela 1 a seguir detalha a identificação de cada cluster e a quantidade de elementos.

Cluster	Características Principais	Nº de componentes
G1	Nível Técnico Alto e demais atributos variados.	21
G2	Níveis técnicos e de gestão de SMS Altos, com patrimônio líquido muito alto.	1
G3	Nível Técnico Médio, Nível de gestão de SMS médio e baixo, patrimônio líquido médio e baixo.	13
G4	Nível Técnico Baixo, Nível de Gestão de SMS baixo, patrimônio líquido variado.	26
G5	Nível Técnico médio e baixo, Nível de gestão de SMS alto, patrimônio líquido variado.	6

**Tabela 1 - Resultado da Clusterização**

A partir da definição destes grupos, foi possível escolher a estratégia de relacionamento com cada um deles, considerando os diferentes perfis de parceria definidos por Ribeiro (2002) e incluindo uma nova categoria: não realizar parceria.

Vale destacar que o cluster G2 é composto por apenas um elemento, o qual foi considerado pelo algoritmo como um elemento com características diferenciadas em relação aos demais. Este prestador de serviço deve ser considerado de maneira especial em relação aos demais, pois apresenta ao mesmo tempo nível técnico elevado e sistema de gestão de SMS alto, além de boa saúde financeira. Isso significa que as duas empresas já possuem alinhamento em relação ao que consideram importante e um planejamento em conjunto das operações pode trazer ainda mais benefícios para ambas. Para este caso, deveria ser considerado um perfil de parceria sem limite na qual o planejamento estratégico é feito em conjunto pelas duas empresas, considerando os diferentes níveis gerenciais e sistemas de custeio são focados na parceria. Inclui a realização conjunta do planejamento de demanda, planos de mudança em processos, introdução de novos produtos, atendimento a novos mercados, mudanças na rede logística.

Além disso, no que se refere aos controles operacionais o desenvolvimento conjunto de indicadores de desempenho também pode ser uma ação a ser tomada, visando o atingimento de objetivos comuns. No que se refere a troca de informações, deve-se aumentar o nível de sofisticação possibilitando a antecipação de problemas e aumentando o nível de comunicação entre a contratante e a prestadora de serviços.

O cluster G1 é composto principalmente por prestadores de serviço com elevado nível técnico, que necessitam, porém, melhorar aspectos relativos aos sistemas de gestão. Além disso, com relação à situação financeira, este cluster apresentou variedade, o que significa dizer que é composto por empresas com diferentes capacidades de investimento. Por este motivo, para este grupo, considerou-se mais adequado o estabelecimento de parceria de longo prazo, buscando a melhoria no aspecto da gestão de SMS.

Neste caso, há maior foco no processo e o processo de custeio deve ser separado para cada empresa, porém já é possível traçar planos em conjunto. Neste caso não há possibilidade de alterações das operações sem que haja consulta da empresa contratante. O nível de troca de informações pode ser maior, desde que seja avaliado o tipo de informação a ser compartilhada.

O cluster G3 apresentou nível técnico intermediário, com nível de gestão de SMS média e baixo e patrimônio líquido médio e pequeno. Isto significa dizer que este grupo de prestadores de serviços devem ser alvo de ações de melhoria na qualidade técnica dos serviços prestados e nos sistemas de gestão de SMS. Para este cluster o tipo de parceria mais indicado seria o de curto prazo no qual são compartilhados planos existentes, os sistemas de custeio são focados em cada empresa e o foco do planejamento está em tarefas e projetos específicos. Neste tipo de parceria não há desenvolvimento conjunto de indicadores de desempenho e pode haver sugestões de ambas as partes para mudanças na operação. Quanto à comunicação, às empresas podem trocar informações com menor frequência e compartilhar somente as informações necessárias à prestação dos serviços.

O cluster G4 apresentou um conjunto de prestadores de serviço de baixo nível técnico e de gestão de SMS, com os mais diversos níveis de patrimônio líquido. Para este grupo, não há interesse da contratante em estabelecer qualquer tipo de parceria. Neste caso, a não ser que haja melhoria em pelo menos um dos atributos, o relacionamento será somente institucional e contratual evitando um nível maior de interação.

O cluster G5 apresentou um grupo de prestadores de serviços de nível técnico médio e baixo, com nível do sistema de gestão alto e patrimônio líquido variado. Neste caso, devido ao alinhamento à questão de SMS, a contratante poderá estabelecer uma parceria de curto prazo, compartilhando planos existentes, os sistemas de custeio são focados em cada empresa e o foco do planejamento está em tarefas e projetos específicos. A companhia contratante poderá desenvolver também ações para melhoria no nível técnico da prestação de serviço.

É importante destacar que não há um tipo ideal de parceria, nem todo acordo entre contratante e contratado deve tornar-se uma relação deste tipo. O melhor relacionamento é aquele que está ajustado aos benefícios que a terceirização trará ao contratante, aliado à sinergia entre os ambientes culturais e características de ambos os parceiros.

## 5. CONSIDERAÇÕES FINAIS

O trabalho foi desenvolvido com o objetivo de aplicar o método de agrupamento AGNES, a um problema de gestão de relacionamento da cadeia de suprimentos. O modelo proporcionou uma visão mais apurada da realidade, extraíndo informação de qualidade a partir de dados cadastrais que antes não eram utilizados. Foi possível identificar atributos e características dos prestadores de serviço gerando subsídios para a gestão adequada do relacionamento.

A partir da aplicação da clusterização hierárquica, foi possível classificar um conjunto de prestadores de serviço segundo diferentes características, observando a multidimensionalidade do problema. Dessa forma, foi possível determinar o perfil de cada empresa e definir quais seriam as melhores formas de relacionamento.

A aplicação apresentada considera diversos aspectos, desde características técnicas até capacidade financeira das empresas, permitindo ao decisor avaliar qual seria a melhor ação estratégica a tomar: parceira de curto prazo, parceria de longo prazo, parceria sem limite e não formar parcerias.

Foi identificado um cluster de empresas para com as quais a empresa contratante não deve estabelecer parcerias tendo em vista o baixo nível técnico e de atendimento à questões relativas à saúde, meio ambiente e segurança.

Por outro lado, foi também identificada uma empresa que possui um desempenho excepcional, sendo avaliada positivamente para todos os aspectos. Isto significa que se trata de um prestador de serviço chave e com o qual a contratante deve estabelecer um relacionamento de parceria prazo limite, por meio do desenvolvimento de planos estratégicos em conjunto, gerando assim ganhos para ambas as empresas.

Vale destacar que por se tratar de uma metodologia de aprendizado não supervisionado, as informações extraídas após aplicação do algoritmo de clusterização não era conhecidas a priori e os grupos foram formados a partir dos dados inseridos no programa. Por este motivo, o método aqui empregado é bastante dependente de um especialista no assunto, o qual participa de todo o processo de definição do modelo, execução e análise dos resultados.

Estudos futuros poderiam envolver um maior número de atributos atrelados a cada empresa, possibilitando uma visão ainda mais ampla do processo de gestão de relacionamento com fornecedores. Além disso, há

possibilidade de explorar outros métodos de agrupamento disponíveis na literatura em comparação a este a fim de realizar a comparação dos resultados.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

BAILY, P. et al. Compras: princípios e administração. 1. ed. São Paulo: Atlas, 2000.

BRAGA, A. Gerenciamento e Desenvolvimento de Fornecedores. Instituto ILOS. 2009. Disponível em: <<http://www.ilos.com.br/web/gerenciamento-e-desenvolvimento-de-fornecedores-parte-1/>>. Acesso em 17/06/2017.

CANTO, R. Proposta de revisão do modelo de avaliação do desempenho dos fornecedores da GKN do Brasil Ltda. Dissertação (Mestrado em Engenharia de Produção) – Escola de Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.

CHRISTOPHER, M. Logística e gerenciamento da cadeia de suprimentos: estratégias para redução de custos e melhoria dos serviços. São Paulo: Pioneira, 2002.

ESTER, M., KRIEGEL, H.P., SANDER, J., XU, X.. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Institute for Computer Science, University of Munich. Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining, 1996.

HAIR, J.F.J., Anderson, R.E., Tatham, R.L., Black, W.C. Multivariate Data Analysis, 5th edn, Prentice Hall, Upper Saddle River, New Jersey, 1998.

JAIN, A.K., DUBES, R.C. Algorithms for Clustering Data, Prentice Hall, 1988.

KAUFMAN, L., ROUSSEAU, P.J. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons, 1990.

KOHONEN, T. Self-Organized Formation of Topologically Correct Feature Maps. Biological Cybernetics 43(1), 59-69, 1982.

KOTLER, P.; ARMSTRONG, G. Princípios de marketing. Rio Janeiro: Prentice-Hall, 1998.

KRAJILIC, P. Purchasing must become supply management. Harvard Business Review 61 (5), 109–117, 1983.

LIU, Y., LI, Z., XIONG, H., GAO, X. and WU, J. Understanding of Internal Clustering Validation Measures. In: Proc. of The 10th International IEEE Conference on Data Mining (ICDM), Sydney, NSW, pp. 911–916, 2010.

POZO, H. Administração de Recursos Materiais e Patrimoniais: uma abordagem logística. São Paulo: Atlas, 2001.

RIBEIRO, Aline, Gestão do Relacionamento com Prestadores de Serviço Logístico, Revista Tecnológica, 2002.

RODRIGUES, D. M., SELLITTO, M. A. Análise do desempenho de fornecedores de uma empresa de manufatura apoiada em análise de aglomerados. *Produção*, v. 19, n. 1, p. 055-069, 2009.

ROUSSEEUW, P. J. Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Computational and Applied Mathematics*. 20: 53–65. 1987.

SLACK, N., CHAMBERS, S., JOHNSTON, R. *Administração da Produção*. 2ª ed. São Paulo: Atlas. 2002.

WALTERS, D., LANCASTER, G. Implementing value strategy through the value chain, *Management Decision*, Vol. 38 Issue: 3, pp.160-178, 2000.