

**Autoria:**

Capitão-Tenente (Intendente da Marinha) Lincoln Costa da Silva

Capitão-Tenente (Intendente da Marinha) Rogerio do Vale Cerqueira Junior

Capitão de Mar e Guerra (RM1-Intendente da Marinha) Hilton de Araújo Lopes (CIANB)

Capitão de Fragata Marcos dos Santos (CASNAV)

# UTILIZAÇÃO DE TÉCNICAS DE MINERAÇÃO DE DADOS PARA DETECTAR POSSÍVEIS RELACIONAMENTOS ENTRE EMPRESAS PARTICIPANTES DE LICITAÇÕES NAS FORÇAS ARMADAS

**Resumo:** As aquisições públicas, normalmente conduzidas por meio de processos licitatórios, podem ser alvo de práticas inidôneas, dentre as quais se destaca o cartel, um dos objetos de análise das auditorias governamentais. Em face do crescente aumento de informações, a tarefa dos auditores se torna cada vez mais desafiadora. Nessa seara, este artigo buscou demonstrar como técnicas de Mineração de Dados (MD) podem auxiliar as Unidades de Auditoria Interna das Forças Armadas na verificação, entre empresas participantes de licitações públicas, de possíveis relações que caracterizem práticas de cartel. Para tal, utilizou-se o *software* estatístico R, que gerou regras de associação entre licitantes. Analisando os resultados, concluiu-se que o uso de técnicas de MD pode contribuir expressivamente nas atividades dos auditores dessas Unidades de Auditoria, permitindo-lhes melhor planejar as avaliações e direcionar os esforços de análise.

**Palavras-chave:** Cartel. Licitações. Auditoria. Mineração de Dados.

## 1 INTRODUÇÃO

As compras governamentais, realizadas, em regra, segundo procedimentos formais de licitação, são instrumentos na implementação de políticas públicas, voltadas ao atendimento das necessidades dos cidadãos e à garantia de direitos sociais. Em média, no período de 2006 a 2017, as compras públicas

demandaram gastos que corresponderam a 12,5% do Produto Interno Bruto (PIB) brasileiro (RIBEIRO; INÁCIO JÚNIOR, 2019).

Devido a essa magnitude, as licitações públicas podem ser alvo<sup>1</sup> de uma prática ilícita que frauda a competitividade, afasta a Administração da obtenção da proposta mais vantajosa e, por conseguinte, dificulta a

gestão eficiente. Trata-se de prática especializada de cartel, em que há uma concertação com o objetivo de vencer o certame e beneficiar os participantes da relação inidônea.

Nesse sentido, os acordos anticompetitivos em licitações constituem fator de alto risco para a eficiência, eficácia, efetividade e integridade das compras públicas, porquanto frequentemente resultam no aumento dos preços, redução da qualidade e/ou restrição da oferta de bens e serviços (ANDRADE, 2019). No intuito de mitigar essa prática, os processos de licitação e contratos têm sido, cada vez mais, objeto de especial atenção das atividades de auditoria governamental.

A fim de apoiar as ações de controle, a Organização para a Cooperação e Desenvolvimento Econômico (OCDE) recomenda o emprego de ferramentas proativas de detecção de cartéis, tais como a análise de dados (OCDE, 2019). Assim, no auxílio à detecção de relações anticoncorrenciais em licitações, a utilização de ferramentas de Mineração de Dados (MD) pode trazer grandes benefícios, pela capacidade de, em um universo extenso de dados, contribuir para a geração de conhecimento útil aos auditores.

Nesse contexto, surge o seguinte problema para o presente estudo: como a MD, em apoio às atividades desenvolvidas pelas Unidades de Auditoria Interna Governamental (UAIG) dos Comandos Militares, pode auxiliar na descoberta de possíveis relacionamentos indevidos entre participantes de licitações das Unidades Administrativas de Serviços Gerais (UASG) subordinadas a cada um desses Comandos?

Buscando responder a essa pergunta, este artigo tem como objetivo geral demonstrar como técnicas de MD auxiliam na verificação de possíveis relacionamentos indevidos entre empresas participantes de pregões, com vistas a apoiar as atividades de auditoria das UAIG dos Comandos Militares. Para atingir esse objetivo, outros objetivos, de cunho específico,

foram traçados, quais sejam: i) levantar, na literatura, as formas pelas quais empresas executam práticas anticoncorrenciais em licitações; ii) selecionar as técnicas de MD que são aplicáveis à auditoria em licitações; e iii) utilizar um *software* estatístico a fim de proceder à MD.

Dessa forma, a pesquisa justifica-se por tratar de práticas caracterizadas como crime contra a ordem econômica e concorrencial, sendo de dano substancial à economia de um país. Além disso, é oportuna por abordar tema de interesse das auditorias governamentais, com foco no Ministério da Defesa, responsável pela contratação média anual de mais de R\$ 13 bilhões nos últimos quatro anos. Contribui, portanto, de maneira relevante, dada a possibilidade de aumento da eficiência operacional no planejamento e execução de auditorias governamentais, prevenção de danos oriundos de fraudes em licitações públicas e incremento do sistema de *accountability*.

Quanto à estrutura, o artigo foi organizado em cinco seções. A presente parte se trata de uma breve contextualização do tema, apresentação do problema de pesquisa e definição dos objetivos do artigo. Na segunda parte, será apresentado o fundamento teórico que estabeleceu as bases para o estudo. Na terceira, será exposta a metodologia aplicada na condução do trabalho com vistas aos objetivos. Na parte seguinte, será realizada uma análise dos resultados obtidos, seguida da proposta de solução para o problema de pesquisa. Na última parte, serão apresentadas considerações finais, limitações e sugestões para pesquisas futuras.

## 2 REFERENCIAL TEÓRICO

### 2.1 Licitação

Licitação pode ser definida como um procedimento utilizado para se contratar com o poder público. Ela tem como dois de seus objetivos garantir a observância do princípio

constitucional da isonomia e selecionar a proposta mais vantajosa para a Administração Pública Federal (BRASIL, 1993).

Conforme disposto na Lei nº 8.666/93, as modalidades de licitação são: concorrência, tomada de preços, convite, concurso e leilão. Adicionalmente, por meio da Lei nº 10.520/02, instituiu-se o pregão a fim de proceder à aquisição de bens e serviços comuns. Em complemento a essa Lei, o Decreto nº 5.450/05 regulamentou a forma eletrônica do pregão, cujo ambiente operacional é o Portal de Compras do Governo Federal (Comprasnet).

De acordo com a Tabela 1, nota-se que a análise de licitações da modalidade pregão é fator de relevância para o referido estudo, pois: i) essa modalidade concentra o maior percentual de gastos realizados pelo governo no período de 2015 a 2018; e ii) é possível utilizar técnicas de MD para as verificações de todos os pregões conduzidos por órgãos governamentais, uma vez que os dados abertos do Sistema Integrado de Administração e Serviços Gerais (SIASG) estão integrados sob a forma de armazém de dados, isto é, *Data Warehouse* (DW).

Dentre outros fatores, o volume das necessidades e a materialidade dos gastos públicos em processos licitatórios despertam o interesse de licitantes que, por vezes, não são idôneos. Conforme afirma a OCDE (2016), as

licitações acabam por fomentar um ambiente propício à formação de cartéis, prática que visa beneficiar um grupo seletivo em detrimento da livre concorrência.

## 2.2 Cartel

Segundo Nunes (2016), o cartel é definido como um acordo entre agentes econômicos que atuam no mesmo mercado e visam construir uma limitação artificial e ilegítima da concorrência. Dentre as práticas anticoncorrenciais, o cartel pode ser considerado como a mais danosa e que mais implica consequências negativas à sociedade e também à economia como um todo (ANDRADE, 2019). A respeito desse conceito, Albuquerque e Leal (2015) complementam que, normalmente, a prática do cartel ocorre de forma regionalizada.

Dentre as situações que caracterizam a referida prática, destacam-se três, quais sejam: i) divisão de mercado, caracterizada quando um conjunto de licitações é dividido entre os membros do cartel, que, desta forma, não concorrem entre si em cada uma delas; ii) rodízio, em que existe um acordo de revezamento entre as empresas a respeito de quem vencerá o certame; e iii) fixação de preços, em que se observam acordos entre os concorrentes para aumentar, fixar ou manipular os preços (WEISHAAR, 2013).

**Tabela 1: Utilização média das modalidades de licitação nas contratações do Governo Federal no período de 2015 a 2018**

<i>Modalidade de Licitação</i>	<i>Valor médio de gastos entre 2015 e 2018</i>	<i>Percentual</i>
<i>Concorrência</i>	R\$ 11.754.888.813,06	8,13%
<i>Concurso</i>	R\$ 8.197.071,90	0,01%
<i>Convite</i>	R\$ 29.330.996,12	0,02%
<i>Pregão</i>	R\$ 131.806.158.615,41	91,17%
<i>Tomada de Preços</i>	R\$ 970.160.960,23	0,67%
<i>Total</i>	R\$ 144.568.736.456,72	100,00%

Fonte: Elaborado pelos autores com base nos dados disponíveis no Portal Brasileiro de Dados Abertos (2019).

Em relação à prática do cartel envolvendo a manipulação de preços na modalidade prego, o Tribunal de Contas da União (TCU), em seu Acórdão nº 1793/2011, destaca a atuação de empresas na condição de “coelho”. Isto é, aquelas que reduzem os preços na fase de lances a fim de desencorajar a participação de outros licitantes, desistindo posteriormente do certame para beneficiar outra empresa que esteja envolvida no conluio, que, por sua vez, é contratada ainda que não tenha apresentado a proposta mais vantajosa.

Ao longo do tempo, notam-se, cada vez mais, o desenvolvimento significativo da Defesa da Concorrência e o aumento da repressão às práticas contrárias à Ordem Econômica. Tais evoluções contribuem sobremaneira para a repercussão, na sociedade, das decisões do Conselho Administrativo de Defesa Econômica (CADE), autarquia federal, vinculada ao Ministério da Justiça e Segurança Pública.

Nesse contexto, Andrade (2019) observa que a prática do cartel representa tanto ilícito administrativo punível pelo CADE, nos termos da Lei nº 12.529/11, quanto crime, passível de punição com pena de 2 a 5 anos de reclusão, conforme preconiza a Lei nº 8.137/90.

Além disso, cabe ressaltar que os processos licitatórios já concretizados e dotados de tal vício poderão também ser anulados, baseado no artigo 90 da Lei nº 8.666/93, que reconhece como crime o ato de “frustrar ou fraudar, mediante ajuste, combinação ou qualquer outro expediente, o caráter competitivo do procedimento licitatório, com o intuito de obter, para si ou para outrem, vantagem decorrente da adjudicação do objeto da licitação” (BRASIL, 1993).

Na tentativa de coibir práticas de cartel, a contribuição dos agentes públicos de compras desempenha papel expressivo. Ante a proximidade com os processos de aquisição, eles podem proporcionar valiosas provas e evidências na apuração dos cartéis que atuam

nas licitações do governo. Dessa forma, as organizações que atuam como UAIG no âmbito dos Comandos Militares, em parceria com as UASG, ganham relevante importância na prevenção e fiscalização de práticas anticoncorrenciais.

### 2.3 Auditoria Interna Governamental

O Sistema de Controle Interno do Poder Executivo Federal (SCI), com base no artigo 74 da Constituição Federal de 1988, possui como órgão central a Controladoria-Geral da União (CGU) e foi instituído, entre outras finalidades, com o objetivo de:

*(...) **comprovar a legalidade** e avaliar os resultados, quanto à eficácia e eficiência, da gestão orçamentária, financeira e patrimonial nos órgãos e entidades da administração federal, bem como da aplicação de recursos públicos por entidades de direito privado (BRASIL, 1988, grifo nosso).*

Integrantes do SCI, as UAIG estão previstas na Instrução Normativa nº 3/2017 da CGU, que aprova o Referencial Técnico da Atividade de Auditoria Interna Governamental do Poder Executivo Federal.

Como parte dessa estrutura, encontram-se a Secretaria de Controle Interno (Ciset) do Ministério da Defesa e as unidades setoriais a ela vinculadas, que são o Centro de Controle Interno da Marinha (CCIMAR), o Centro de Controle Interno do Exército (CCIEEx) e o Centro de Controle Interno da Aeronáutica (CENCIAR).

Segundo o mesmo Referencial Técnico, as UAIG são responsáveis pela avaliação dos controles internos da gestão das Unidades Auditadas, devendo verificar, entre outros elementos, a conformidade com leis, regulamentos, políticas e procedimentos vinculantes.

Por meio da Instrução Normativa nº 8/2017, a CGU aprovou o Manual de Orientações Técnicas da Atividade de Auditoria Interna Governamental do Poder

Executivo Federal, que elenca os tipos de serviço de avaliação a serem conduzidos pelas UAIG. De acordo com o Manual, a Avaliação de Conformidade ou *Compliance* é aquela que se ocupa em verificar se certas atividades financeiras ou operacionais de um objeto de auditoria obedecem às condições, às regras e aos regulamentos aplicáveis.

Depreende-se, portanto, que cabe ao CCIMAR, CCIEEx e CENCIAR, como UAIG componentes do SCI do Poder Executivo Federal, a identificação de possíveis arranjos, envolvendo empresas participantes de licitações, que venham a caracterizar práticas de cartel, haja vista sua falta de aderência à legislação. Cabe ressaltar que essa não é uma tarefa fácil em virtude da alta demanda por atividades de auditoria. Nessa seara, o advento da MD surge como ferramenta a fim de auxiliar e, por vezes, agilizar tais atividades.

## 2.4 Descoberta de Conhecimento em Bases de Dados

De Battista *et al.* (2016) salientam que é de notório interesse, para as organizações que detêm dados massivos, poder extrair conhecimento deles e convertê-los em informações úteis para a tomada de decisão. Assim, surge a necessidade de se aplicarem técnicas pertencentes a um ramo da Ciência

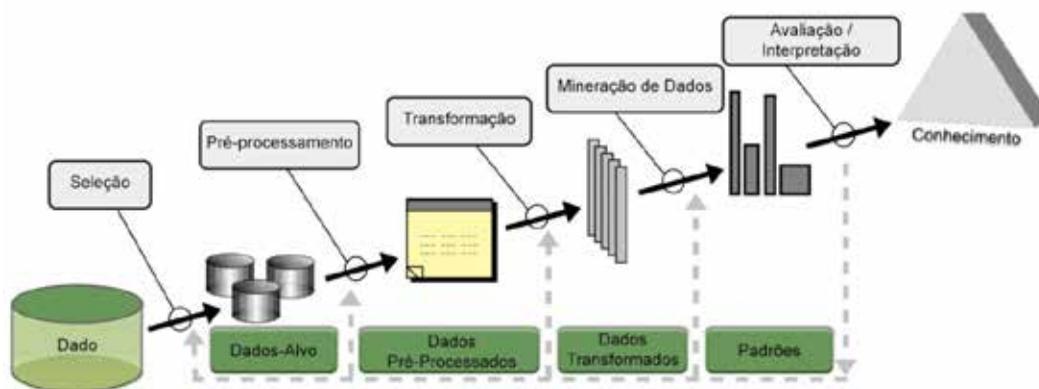
da Computação conhecido como Descoberta de Conhecimento em Bases de Dados (DCBD) ou *Knowledge Discovery in Databases* (KDD).

Segundo definiram Fayyad, Piatetsky-Shapiro e Smyth (1996), o KDD é um processo iterativo e iterativo, não trivial, composto por várias etapas, de extração de informações implícitas, previamente desconhecidas e potencialmente úteis, a partir dos dados armazenados em um banco de dados.

Na Figura 1, podem ser observadas as etapas do KDD, descritas por Han, Pei e Kamber (2012) como:

- Seleção: escolha de um conjunto de dados ou uma amostra dos dados com que se pretende aplicar o processo de descoberta;
- Pré-processamento: etapa em que serão tratados eventuais problemas como ruídos e dados incompletos;
- Transformação: os dados são transformados e consolidados nas formas adequadas para a mineração, de acordo com a técnica a ser utilizada;
- Mineração de dados: um processo essencial no qual métodos inteligentes são aplicados no intuito de extrair padrões a partir dos dados; e
- Avaliação/Interpretação: consiste na identificação dos padrões obtidos que foram realmente relevantes, dadas as medidas

Figura 1: Etapas do KDD



Fonte: Adaptado de Fayyad, Piatetsky-Shapiro e Smyth (1996).

ou funções de avaliação. Além disso, técnicas de representação e visualização são utilizadas para apresentar o conhecimento minerado ao usuário.

A etapa de MD é a mais característica do KDD, por isso, com frequência, utiliza-se essa fase para denominar todo o processo. O objetivo principal do KDD é a busca por novo conhecimento que possa ser útil ao usuário (DE BATTISTA *et al.*, 2016; GOLDSCHMIDT; PASSOS; BEZERRA, 2015).

## 2.5 Mineração de Dados

A mineração de dados objetiva a análise de grandes conjuntos de dados a fim de encontrar relacionamentos, padrões ou tendências de forma a resumí-los para que sejam úteis e possam auxiliar as tomadas de decisão dos mais diversos setores (LOKBANI, 2017).

Goldschmidt, Passos e Bezerra (2015) ressaltam que os avanços na área de Tecnologia da Informação (TI) resultam no armazenamento de grandes volumes de dados, que podem apresentar tendências, correlações ou padrões com potencial para melhorar os processos de uma organização. Tais avanços têm despertado, cada vez mais, o interesse de organizações detentoras de *big data* com vistas a extrair conhecimentos atinentes aos mais diversos assuntos.

Dessa forma, por meio do advento da MD, tornou-se possível a utilização de dados importantes que antes deixavam de ser explorados por estarem agrupados a muitos outros de menor relevância. Isto é, as tomadas de decisão, bem como muitas diretrizes organizacionais, são amparadas, por vezes, em dados que, anteriormente, encontravam-se armazenados e sem nenhuma utilidade aparente (DE BATTISTA *et al.*, 2016).

A fim de se obterem respostas ou se extrair conhecimento interessante, existem diversas técnicas de MD disponíveis, tais como: classificação, modelos de relacionamento entre

variáveis, regressão, clusterização, sumarização, modelo de dependência, regras de associação e análise de séries temporais (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). Dentre as técnicas de MD apresentadas, serão discutidas a clusterização e as regras de associação por apresentarem relevância para o presente estudo.

### 2.5.1 Clusterização

Clusterização é uma tarefa descritiva em que se procura identificar um conjunto finito de categorias ou *clusters* para descrever uma informação (JAIN; DUBES, 1988). Conforme definem Baensens, Vlasselaer e Verbeke (2015), o objetivo da clusterização (agrupamento) é dividir um conjunto de observações em segmentos de forma que a homogeneidade dentro de cada segmento seja maximizada e a heterogeneidade entre os segmentos também seja maximizada.

Portanto, um *cluster* pode ser entendido como uma coleção de registros que apresentam semelhanças entre si e que, ao mesmo tempo, são diferentes dos demais *clusters*. Por meio desse tipo de análise, é possível encontrar diferentes grupos dentro de um conjunto de dados.

### 2.5.2 Regras de Associação

Segundo Araújo e Maciel (2018), a técnica de regras de associação, de origem descritiva, tem como princípio descobrir possíveis associações entre itens em uma mesma transação, encontrando padrões ou tendências frequentes relevantes para análise, em forma de regras de associação. Cabe ressaltar que essa técnica é uma das mais conhecidas devido aos bons resultados obtidos (CAMILO; SILVA, 2009 *apud* SILVA; SANTOS; HORA, 2018).

A regra de associação tem o formato  $A \rightarrow B$ , na qual  $A$  e  $B$  são conjuntos de itens, também chamados de *itemsets*, e  $A \cap B = \emptyset$ .  $A$  é chamado de antecedente;  $B$ , de consequente; e o

conjunto de A mais B é chamado de conjunto de itens ou transações. A regra pode ser lida como: o atributo A frequentemente implica o atributo B (ARAÚJO; MACIEL, 2018; PEREIRA *et al.*, 2019).

Conforme exemplificado por Melo e Silva (2016, p. 9), “uma regra é uma implicação tipo  $\{\text{licitante } A, \text{ licitante } B\} \rightarrow \{\text{licitante } C\}$  significando que se os licitantes A e B participaram de uma licitação é provável que o licitante C também tenha participado”. Assim, ao se aplicar tal técnica, no presente estudo, objetiva-se gerar regras que indiquem os licitantes que, de forma recorrente, participam juntos em licitações.

Melo e Silva (2016) salientam que, quando se vale de ferramentas como a descrita acima, faz-se necessária a utilização de algumas medidas de qualidade propostas na literatura, sendo as mais relevantes o *support* (suporte), a *confidence* (confiança) e o *lift* (elevação).

Consoante definem Boguski e Cury (2018), o suporte de uma regra é representado pela proporção em que os itens presentes na regra ocorrem simultaneamente em toda a base de dados. Então, o suporte da regra  $A \rightarrow B$  é obtido por:

$$\text{supp}(A \rightarrow B) = \frac{\sigma(A \cup B)}{N}, \text{ sendo } \sigma(A \cup B)$$

representado pela quantidade de vezes em que os *itemsets* em A e B ocorreram juntos no conjunto de todas as transações, representadas por N (MELO; SILVA, 2016). PEREIRA *et al.* (2019) definem que a regra  $A \rightarrow B$  tem confiança representada por

$$\text{conf}(A \rightarrow B) = \frac{\text{supp}(A \cup B)}{\text{supp}(A)}$$

no conjunto de N transações, isto é, ela fornece a probabilidade de se obter o consequente B, dado o antecedente A.

O *lift* pode ser definido como uma medida de associação entre os *itemsets* A e B de uma regra, que indica o afastamento da hipótese de independência entre o antecedente A e o

consequente B de uma regra. Essa medida é dada por

$$\text{lift}(A, B) = \frac{\text{supp}(A \rightarrow B)}{\text{supp}(A) \times \text{supp}(B)}$$

Faz-se importante mencionar que, para um *lift* = 1, significa dizer que A e B são independentes (MELO; SILVA, 2016).

Ralha e Silva (2012 *apud* BALDOMIR, 2017) salientam que considerar valores altos para o suporte mínimo não garante boas regras, pois uma regra que associa alguns fornecedores e que possui suporte elevado possivelmente evidencia a presença de grandes fornecedores participando de muitas licitações. Dessa forma, estabelecer um suporte mínimo alto pode suprimir a aparição de boas regras, com reais características de cartéis. Em relação à confiança, valores altos garantem que a frequência de ocorrência dos fornecedores de uma regra seja aproximadamente igual e, por isso, os participantes dela como um todo podem ser considerados um grupo para fins de identificação de cartéis.

No âmbito do presente estudo, em face das práticas de cartel organizadas por algumas empresas licitantes, aumentam-se os desafios das UAIG subordinadas aos Comandos Militares. A identificação de práticas como essa, por vezes, requer um esforço significativo dos auditores e, por conseguinte, demanda tempo, componente escasso nos dias atuais. Logo, valendo-se primordialmente das duas técnicas supracitadas, busca-se demonstrar como a MD auxilia na verificação de possíveis relações que podem evidenciar conluios entre empresas participantes de pregões.

### 3 METODOLOGIA

#### 3.1 Classificação da pesquisa

Com fundamento em Prodanov e Freitas (2013), esta pesquisa apresenta uma

abordagem qualitativa e quantitativa, sendo classificada como: i) quanto à natureza, uma pesquisa aplicada, pois envolveu a aplicação de técnicas de MD em base de dados de licitações realizadas; e ii) quanto aos objetivos, uma pesquisa descritiva, pois buscou descrever o fenômeno do cartel em licitações e como a MD pode auxiliar na identificação de indícios relacionados a esse fenômeno.

Em relação aos procedimentos realizados, o estudo se enquadra nas seguintes tipologias:

- a) Pesquisa documental, analisando documentos das duas classificações de Gil (2008):
  - i) fontes de primeira mão, principalmente documentos oficiais normativos e orientativos do Governo Federal; e
  - ii) fontes de segunda mão, notadamente relatórios de pesquisas da OCDE;
- b) Pesquisa bibliográfica, por meio de consultas a livros, dissertações, periódicos e artigos científicos publicados em relação aos temas: i) mineração de dados; ii) auditoria governamental; iii) fraudes em licitações públicas; e iv) combate a cartéis.

### 3.2 Coleta e processamento de dados

A coleta de dados e o processamento deu-se em 4 fases, descritas a seguir.

- a) 1ª fase - Inicialmente, procedeu-se a uma pesquisa bibliográfica e documental com a finalidade de aprofundar os conhecimentos atinentes à utilização da MD em auditoria de processos licitatórios governamentais. Ademais, buscou-se estudar, na literatura, as formas mais comuns de atuação de práticas anticoncorrenciais.

Em seguida, foram delimitados os parâmetros a serem utilizados no presente estudo. Isto é, a fonte para coleta dos dados pertinentes; o período a ser considerado para a extração dos dados; a modalidade de licitação de maior relevância para a pesquisa; e as unidades governamentais a serem estudadas.

Para obtenção dos dados, foi selecionado o DW do SIASG, plataforma onde se operacionalizam as compras do Governo Federal e cujos dados encontram-se disponíveis no Portal Brasileiro de Dados Abertos. Além disso, estabeleceu-se o período de estudo de 2015 a 2018. Quanto à escolha da modalidade de licitação, foi selecionado o pregão, pois, de acordo com o Portal de Brasileiro de Dados Abertos, ele possui expressiva utilização em relação às demais modalidades. Durante o período selecionado para o estudo, o pregão representou 91,17 % das contratações do Governo Federal.

No que tange às UASG condutoras de certames, optou-se por analisar todas as que compõem os comandos militares, integrados ao Ministério da Defesa. Tal escolha deveu-se ao fato de essas UASG, no período selecionado para análise, terem perfeito uma média de gastos de R\$ 13,5 bilhões, conforme evidenciado pelo mesmo Portal.

- b) 2ª fase – Nesta fase, foram realizadas as três primeiras etapas da Descoberta de Conhecimento em Bases de Dados, ou seja, a seleção, o pré-processamento e a transformação, descritas por Han, Pei e Kamber (2012).

O Portal Brasileiro de Dados Abertos disponibiliza seu conteúdo por meio de *Application Programming Interface* (API), isto é, um conjunto definido de mensagens de requisição e resposta *Hypertext Transfer Protocol* (HTTP). A API de Compras Públicas do Governo Federal<sup>2</sup> acessa informações das aquisições realizadas.

Na extração dos dados disponíveis no Portal, observou-se que o sistema possuía um limite de resposta por consulta realizada. Sendo assim, diante da magnitude do volume dos dados, desenvolveu-se um *script* em *Python* no qual foram utilizados os pacotes *request*, *json* e *pandas* a fim de automatizar a requisição de dados, que se limitava à resposta de 500 *outputs* por vez.

Dentre os módulos disponíveis no sistema para a coleta dos dados, o *script* em *Python* foi utilizado naqueles de interesse para o presente estudo, sendo eles: Módulo de Licitações, de Pregões, de Fornecedores, de Materiais, de Serviços e de Itens. Com cada um deles, operou-se de forma iterativa, limitando-se sempre a 500 *outputs* por vez, até que todos os dados fossem obtidos. Assim que acabavam as iterações de um determinado módulo, o respectivo conteúdo era salvo em um arquivo de formato CSV. Nesse caso, a opção por esse formato deveu-se ao fato de ele ser satisfatoriamente suportado pelo *Software Microsoft Office Excel®*.

Uma vez extraídos todos os dados pertinentes, fez-se necessário agrupá-los em um único arquivo a fim de proceder à mineração. Para tal, foram identificados os atributos comuns em cada conjunto de dados até que se tornasse possível a unificação de todos os dados, obtendo-se, por conseguinte, um grande conjunto de dados pronto para ser minerado. Esse procedimento de unificação foi realizado por meio de um *script* contendo o pacote *pandas* do *Python*.

De posse do banco de dados apropriado, por medida de precaução, executou-se um *script* contendo o pacote *pandas* do *Python* a fim de verificar se todos os dados correspondiam aos parâmetros delimitados inicialmente, quais sejam: o período escolhido, a modalidade de licitação e as UASG pertinentes.

A Tabela 2 mostra algumas informações da base de dados utilizada neste estudo. Cabe ressaltar que cada registro da referida base representa uma participação de uma empresa em determinada licitação. As duplicidades foram desconsideradas, já que uma mesma empresa, por vezes, aparecia mais de uma vez em uma mesma licitação por apresentar propostas de preço para mais de um item.

**Tabela 2: Conteúdo da base de dados utilizado na mineração de dados**

Informações	Total contabilizado
Registros	161.501
Pregões	6.302
Empresas	22.523
Empresas vencedoras de pelo menos 20 pregões	1.010
Empresas vencedoras de pelo menos 50 pregões	241

Fonte: Elaborado pelos autores (2019).

c) 3ª fase – Fase em que se executou a quarta etapa da Descoberta de Conhecimento em Bases de Dados, a MD. A fim de facilitar a MD, utilizou-se um *script* contendo o pacote *pandas* do *Python* para associar os CNPJ das empresas a cada um dos pregões em que elas haviam participado, gerando um arquivo<sup>3</sup> CSV com esses dados. De forma análoga aos estudos conduzidos por Baldomir (2017) e Melo e Silva (2016), esse arquivo foi utilizado como *input* no *Software R (open source)*.

No *software* estatístico R, para obtenção das regras de associação, fez-se uso da função *apriori* do pacote *arules*. Esse pacote oferece um ambiente completo para a exploração da técnica de regras de associação em R, destacando-se por disponibilizar mais de 30 medidas de avaliação do grau de interesse referente a uma regra. A respeito dessas medidas, foram utilizados o suporte (*support*), a confiança (*confidence*) e a elevação (*lift*). O *software* requereu os valores de suporte e confiança a serem considerados para o experimento, sendo assim, optou-se por atribuir os valores mínimos de 70% para a confiança e 0,1% para o suporte, analogamente ao estudo de Melo e Silva (2016).

Como resultado, obtiveram-se 20 regras de associação. Com o intuito de se proceder a uma boa avaliação das regras, foi utilizado o pacote *aruleViz* do R, que forneceu

visualizações gráficas e, por conseguinte, facilitou a interpretação.

d) 4ª fase – Fase durante a qual se buscou, a partir das empresas elencadas nas regras, realizar uma busca quanto às localidades em que ocorreram as associações, bem como os itens de fornecimento envolvidos. Para tal, foi executado um *script* que contém o pacote *pandas* do *Python*. Em complemento, de posse dos CNPJ dos licitantes presentes nas associações, procedeu-se a consultas visando à busca por informações atinentes ao logradouro, correio eletrônico, telefones e sócios. Esse procedimento teve por objetivo encontrar características e informações comuns entre as empresas.

A última etapa da Descoberta de Conhecimento em Bases de Dados, em que ocorrem a avaliação e a interpretação dos dados minerados, será discutida na próxima seção deste artigo.

## 4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

### 4.1 Regras de associação geradas

Ao se aplicar a técnica de regras de associação, por meio do *software R*, na base de dados da pesquisa, 20 regras foram obtidas, com a configuração mínima para o suporte de 0,1% e para a confiança de 70%. Na Tabela 3, são

Tabela 3: Regras de associação

	Regras	Suporte	Confiança	Lift
1	{xx.xxx.918/0001-06} => {xx.xxx.818/0001-17}	0.01142494	1	68.5
2	{xx.xxx.622/0001-13, xx.xxx.048/0001-67} => {xx.xxx.984/0001-06}	0.01063155	1	56.77477
3	{xx.xxx.467/0001-96} => {xx.xxx.510/0001-38}	0.02221517	1	45.01429
4	{xx.xxx.510/0001-38} => {xx.xxx.467/0001-96}	0.02221517	1	45.01429
5	{xx.xxx.368/0001-99} => {xx.xxx.934/0001-08}	0.01523326	1	27.04721
6	{xx.xxx.048/0001-67} => {xx.xxx.984/0001-06}	0.0172961	0.9909091	56.25864
7	{xx.xxx.984/0001-06} => {xx.xxx.048/0001-67}	0.0172961	0.981982	56.25864
8	{xx.xxx.327/0001-70, xx.xxx.075/0001-17} => {xx.xxx.622/0001-13}	0.01047287	0.9295775	10.76874
9	{xx.xxx.308/0001-68, xx.xxx.934/0001-08} => {xx.xxx.622/0001-13}	0.01142494	0.9113924	10.55808
10	{xx.xxx.075/0001-17, xx.xxx.934/0001-08} => {xx.xxx.622/0001-13}	0.01126626	0.9102564	10.54492
11	{xx.xxx.313/0001-95, xx.xxx.308/0001-68} => {xx.xxx.622/0001-13}	0.0125357	0.908046	10.51931
12	{xx.xxx.534/0001-07, xx.xxx.137/0001-37} => {xx.xxx.622/0001-13}	0.01015551	0.9014085	10.44242
13	{xx.xxx.518/0001-07} => {xx.xxx.487/0001-71}	0.01348778	0.8854167	13.34903
14	{xx.xxx.534/0001-07, xx.xxx.308/0001-68} => {xx.xxx.622/0001-13}	0.01221834	0.8850575	10.253
15	{xx.xxx.136/0001-67, xx.xxx.704/0001-54} => {xx.xxx.622/0001-13}	0.01063155	0.8815789	10.2127
16	{xx.xxx.681/0001-44} => {xx.xxx.622/0001-13}	0.01396382	0.88	10.19441
17	{xx.xxx.137/0001-37, xx.xxx.313/0001-95} => {xx.xxx.622/0001-13}	0.01015551	0.8767123	10.15633
18	{xx.xxx.301/0001-45, xx.xxx.308/0001-68} => {xx.xxx.622/0001-13}	0.01110758	0.875	10.13649
19	{xx.xxx.534/0001-07} => {xx.xxx.622/0001-13}	0.01634402	0.8728814	10.11195
20	{xx.xxx.303/0001-60} => {xx.xxx.622/0001-13}	0.01063155	0.8701299	10.08007

Fonte: Elaborado pelos autores com base nas regras identificadas pelo *Software R* (2019).

apresentadas as 20 regras geradas, das quais 7 serão analisadas. Essa escolha deveu-se ao fato de elas possuírem, dado o suporte mínimo, valores maiores que 0,95 para confiança, o que garante boas regras (RALHA; SILVA, 2012 *apud* BALDOMIR, 2017). A fim de se preservarem as identidades dos licitantes, foram omitidos os cinco primeiros dígitos constantes do CNPJ de cada um deles.

A primeira regra pode ser interpretada da seguinte forma: quanto ao suporte, em 1,14% de todos os pregões de UASG vinculadas aos Comandos Militares conduzidos entre 2015 e 2018, o CNPJ xx.xxx.918/0001-06 (antecedente) e o CNPJ xx.xxx.818/0001-17 (consequente) participaram juntos. Nesse contexto, é importante evidenciar que, de acordo com o estudo de Ralha e Silva (2012 *apud* BALDOMIR, 2017), valores menores para o suporte tornam possível que boas regras não sejam suprimidas durante a MD.

Em relação à confiança, pode-se afirmar que, na base dados, para 100% das licitações que contêm o CNPJ xx.xxx.918/0001-06 e CNPJ xx.xxx.818/0001-17, a regra 1 está correta e, portanto, é válida. Quanto ao *lift*, observa-se que a participação da empresa consequente tem 68,5 mais chances de ocorrer quando a empresa antecedente participa do pregão. Sendo assim, elas podem ser consideradas interdependentes. De forma análoga, todas as demais regras da Tabela 3 podem ser interpretadas da mesma maneira, seguindo suas respectivas medidas.

Além disso, nota-se que, à confiança maior que 0,95, todas as regras selecionadas descrevem as associações entre duas empresas, com exceção da regra 2. No caso específico dessa regra, é possível verificar que os licitantes de CNPJ xx.xxx.622/0001-13 e xx.xxx.048/0001-67, atuando juntos na condição de antecedentes, possuem como consequente a empresa de CNPJ xx.xxx.984/0001-06. Dentre as regras

apresentadas no presente estudo, este é o único caso, à confiança mínima de 0,95, identificado por dois antecedentes possuindo um consequente.

As regras 3 e 4 são as mesmas em termos práticos, apenas havendo a alternância entre o antecedente e o consequente da regra. Isso pode indicar ligação íntima entre as empresas, fator de alto risco para o concerto de ações em certames públicos. O mesmo ocorre com as regras 6 e 7, virtualmente idênticas e que envolvem empresas que também constam na regra 2.

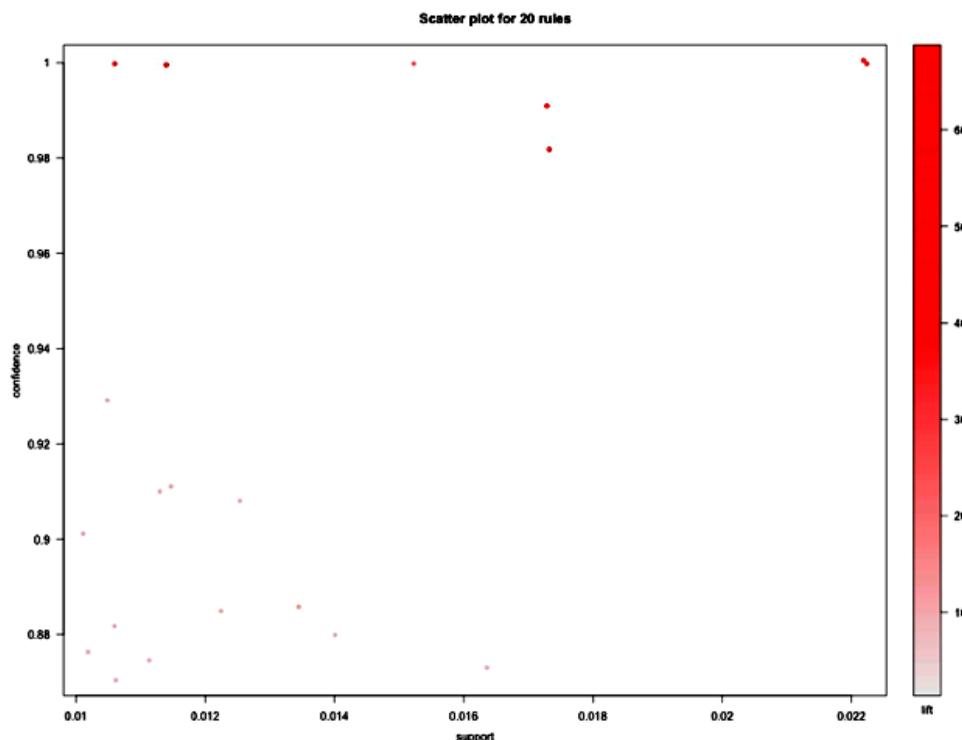
Para facilitar a interpretação das regras da Tabela 3, foram utilizadas representações gráficas das associações, conforme apresentado no Gráfico 1. Nesse gráfico, em que as 20 regras podem ser observadas, cada ponto representa uma regra. A posição do ponto é estabelecida pelo suporte e confiança da regra, e sua cor varia conforme o *lift*. O eixo vertical corresponde à confiança, variando da confiança mínima definida até o máximo encontrado nas regras geradas. Já o eixo horizontal corresponde ao suporte, também variando do suporte mínimo definido até o máximo encontrado nas regras. A saturação da cor numa escala de vermelho corresponde ao *lift*. Quanto maior for o valor do *lift*, mais intenso será o vermelho.

O Gráfico 1 é especialmente útil no presente estudo por facilitar a visualização das 20 regras de associação resultantes da MD, identificando, numa mesma visualização, três das principais medidas de significância e interesse utilizadas na seleção das regras. Nesse contexto, nota-se que as 7 primeiras regras da Tabela 3, selecionadas como as de maior interesse, encontram-se representadas na parte superior do Gráfico 1.

## 4.2 Análise de associações

O Gráfico 2 ilustra as regras geradas pelo software R, presentes na Tabela 3. Nesse gráfico, cada regra é representada por um

Gráfico 1: Diagrama de dispersão das regras de associação



Fonte: Adaptado do Software R (2019).

círculo, cujo tamanho é diretamente proporcional ao valor do suporte da regra representada. Assim, quanto maior o suporte, maior o diâmetro do círculo.

Em cada círculo, a saturação da cor numa escala de vermelho corresponde ao *lift*, sendo mais intenso o vermelho quanto maior for o *lift*. Em relação às setas, elas ligam os CNPJ das empresas que representam os antecedentes das regras geradas aos CNPJ das empresas que atuam como consequentes das regras, passando, antes disso, pelo círculo indicador da própria regra que os une.

Constatou-se que as empresas apresentadas no Gráfico 2 pertencem majoritariamente ao ramo de materiais de construção. À primeira vista, com foco nesse segmento, os dados podem não parecer muito significativos para indícios de cartel, pois há baixa concentração no mercado de varejo de materiais de construção<sup>4</sup>. No entanto,

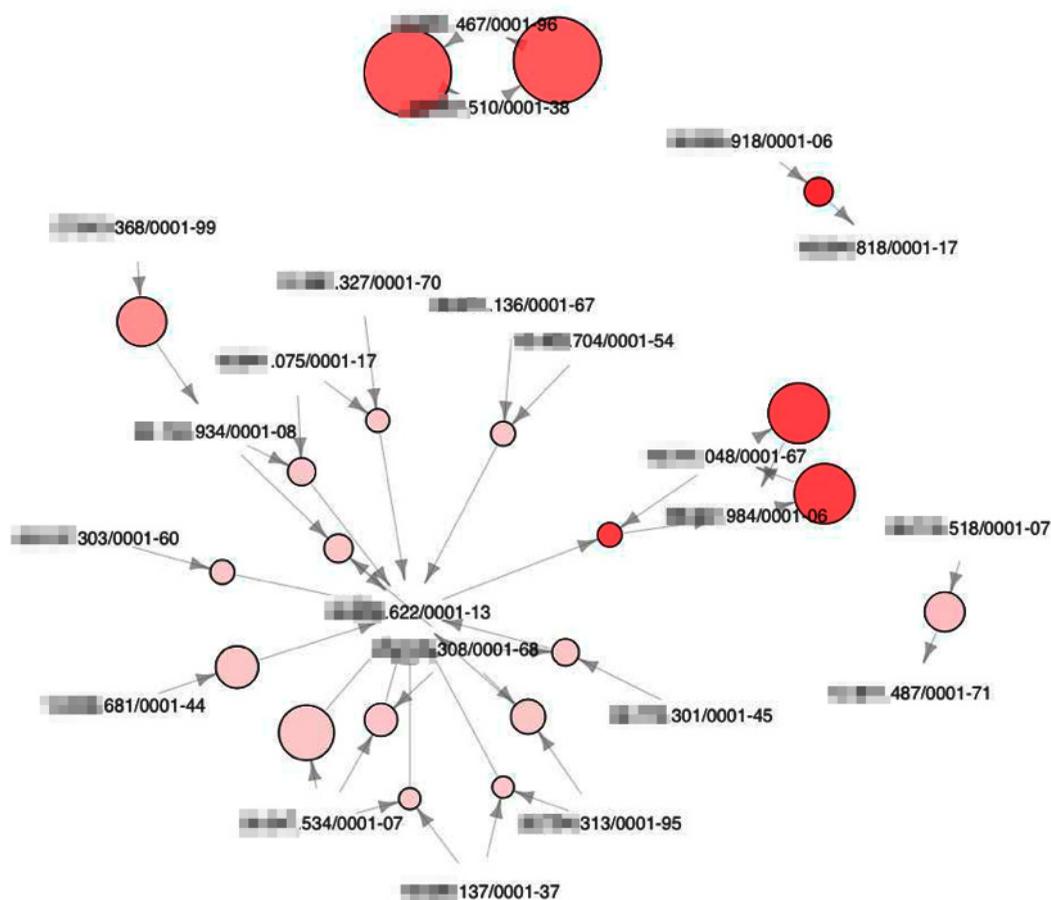
empresas desse ramo já sofreram condenação no Tribunal Administrativo do CADE por práticas de cartel em licitações de materiais de construção em um município de Santa Catarina<sup>5</sup>.

Além disso, outro dado importante é que as empresas se concentram, expressivamente, nas Regiões Sul e Sudeste do país. Essas localidades também aparecem no estudo de Albuquerque e Leal (2015), tendo sido identificadas como as regiões com mais casos de cartéis condenados pelo CADE.

Dentre as empresas representadas no Gráfico 2, foram selecionadas para análise somente as que apresentaram regra com confiança acima de 0,95. A fim de facilitar o entendimento da discussão dos licitantes escolhidos, dividiu-se a análise em duas partes, quais sejam: i) empresas integrantes do *cluster*; e ii) empresas não integrantes do *cluster*.

Gráfico 2: Grafo das regras de associação

Graph for 20 rules



Fonte: Adaptado do Software R (2019).

#### 4.2.1 Empresas integrantes do cluster

No Gráfico 2, há somente um *cluster* de relacionamentos bem definido, composto de 16 empresas, com o CNPJ xx.xxx.622/0001-13 destacando-se no ponto central.

Em relação ao *cluster*, o estado com maior representação é o Rio Grande do Sul, respondendo por 25% das empresas. Das empresas gaúchas no *cluster*, três estão localizadas na mesma cidade e a quarta tem sede em município limítrofe, o que é um interessante sinal do fator de regionalidade presente nos cartões, conforme também apontam Albuquerque e Leal (2015).

No *cluster*, observa-se um relacionamento envolvendo a empresa do ponto central — CNPJ xx.xxx.622/0001-13 — e duas na posição nordeste do aglomerado — CNPJ xx.xxx.048/0001-67 e xx.xxx.984/0001-06. No período observado, tais empresas participaram do mesmo pregão em 67 ocasiões, na maioria em licitações de Organizações Militares localizadas no Rio Grande do Sul. Aqui, destaca-se que esse Estado é sede de 187 Organizações Militares, sendo 16 da Marinha, 28 da Força Aérea e 143 do Exército.

Além disso, salienta-se que o Estado do Rio Grande do Sul possui um programa consolidado de Redes de Cooperação, conduzido pela Secretaria Estadual de Desenvolvimento Econômico e Turismo<sup>6</sup>. Esse programa possibilita uma atuação conjunta em cada segmento específico para a superação das fragilidades inerentes às empresas de médio e pequeno porte. O mercado varejista de materiais de construção do Rio Grande do Sul, área de atuação das empresas analisadas, é um dos setores mais bem-sucedidos em relação a iniciativas de cooperação, tendo o aumento de sua participação na economia gaúcha seguido o próprio avanço da urbanização do estado, conforme verificado na pesquisa de Flores (2018).

Apesar de se constituírem para outros fins, as associações de qualquer tipo, como as redes de cooperação e sindicatos, em razão de intensa troca de informações acerca do segmento de interesse, podem suscitar a formação de cartéis, conforme alerta o estudo de Ferreira (2015). Sob a aparência de organização formal e regular, essas associações podem encobrir reuniões realizadas com o intuito de combinarem práticas, preços e, inclusive, conluíes em licitações públicas.

As regras de associação entre as empresas com os CNPJ xx.xxx.048/0001-67 e xx.xxx.984/0001-06 possuem características muito parecidas, o que pode ser indicativo de conexão mais próxima entre as empresas, como o caso de empresas da mesma família, as quais estão localizadas no mesmo bairro, em logradouros a apenas duas quadras de distância.

A situação de participação de empresas da mesma família ou, de alguma outra forma, conectadas intimamente é potencialmente danosa ao caráter competitivo da licitação, uma

vez que facilita a formação de conluíes e não garante o sigilo das propostas, além de ferir o princípio da moralidade. Uma prática que naturalmente surge com os relacionamentos entre empresas dessa natureza é o rodízio de vencedores dos certames, conforme explicado por Weishaar (2013).

Além disso, na extremidade noroeste do *cluster*, verificou-se forte associação entre as empresas com os CNPJ xx.xxx.368/0001-99 e xx.xxx.934/0001-08. Essas empresas apareceram juntas em 233 pregões no período da pesquisa. Apesar de estarem sediadas em diferentes cidades no Estado de São Paulo, compartilham o mesmo endereço de correio eletrônico. De alguma forma, esse dado aponta para uma simbiose entre as empresas, seja pela constituição de dois cadastros para a mesma estrutura organizacional, seja pela vinculação entre os sócios da empresa.

A participação de empresas com sócios em comum nas licitações públicas pode ensejar a atuação de uma das empresas do conluio como “coelho”, causando prejuízos para a Administração, além dos que já foram discutidos. Para evitar a ocorrência desses relacionamentos indevidos, inclusive mediante a atuação de “coelhos”, o TCU, por meio do Acórdão nº 1793/2011, recomendou que os agentes públicos, ao conduzirem procedimentos licitatórios, tomem ciência do quadro societário das empresas participantes, o que já vem ocorrendo por meio de alertas automáticos no Comprasnet.

#### 4.2.2 Empresas não integrantes do *cluster*

Fora do *cluster*, seis empresas apresentam associações aos pares: xx.xxx.467/0001-96 com xx.xxx.510/0001-38; xx.xxx.918/0001-06 com

xx.xxx.818/0001-17; e xx.xxx.518/0001-07 com xx.xxx.487/0001-71. Em virtude da qualidade das regras de associação, somente dois desses pares serão discutidos.

O primeiro par é formado pelas empresas de CNPJ xx.xxx.918/0001-06 e xx.xxx.818/0001-17. A regra que as associa aponta para um alto grau de interdependência entre as licitantes. Essa observação pode ser constatada pelo *lift*, cujo valor é de 68,5.

Essas empresas possuem extensa lista de atividades econômicas secundárias, o que lhes permite participar de licitações de diferentes categorias de materiais e serviços. Isso, por si só, não explica a grande quantidade de 92 pregões em comum às duas empresas. Além disso, estão localizadas no mesmo município do interior do Estado de Santa Catarina. Esse é relevante sinalizador da característica de localidade, comumente encontrada em cartéis, conforme ressaltam Albuquerque e Leal (2015).

O segundo par analisado envolve duas empresas posicionadas ao norte do *cluster*, cujos CNPJ são xx.xxx.467/0001-96 e xx.xxx.510/0001-38. Trata-se de empresas mineiras que possuem em comum a participação em 140 pregões. Suas relações demonstram uma situação peculiar, em que as duas regras que as relacionam são relevantes e de características idênticas, o que pode configurar uma ligação íntima e interativa entre as empresas.

## 5 CONSIDERAÇÕES FINAIS

Este estudo teve por objetivo demonstrar como técnicas de MD podem auxiliar as Unidades de Auditoria Interna das Forças Armadas na verificação, entre empresas participantes de licitações públicas, de possíveis relações que caracterizem práticas de cartel.

Esse objetivo foi atingido após a realização de pesquisa documental e bibliográfica e a aplicação de técnicas de MD numa base de dados comportando os pregões realizados pelas UASG vinculadas aos Comandos Militares no período de 2015 a 2018.

Como resultado, foram geradas relevantes regras de associação entre as empresas, concomitantes à aparição de um *cluster*. As regras geradas na pesquisa descreveram o comportamento das coparticipações das empresas nos pregões constantes na base de dados. Pela interpretação das regras, houve indicação de fortes ligações entre algumas licitantes, o que pode facilitar a ocorrência de acordos anticoncorrenciais nos certames.

Além disso, durante a análise dos resultados, procurou-se, de alguma forma, reforçar e aprofundar as descobertas obtidas com as regras, buscando informações comuns entre as empresas sinalizadas. Aqui, simulou-se o que o sistema de regras permitiria ao auditor, indicando-lhe caminhos para ampliar o escopo de análise em relação a associações entre licitantes, seguindo, em última instância, para a verificação de evidências de relacionamentos concretos.

Já o *cluster* surgente, relacionado ao setor de varejo de materiais de construção, apontou para uma área de atuação comercial com propensão à formação de relações inidôneas em licitações públicas. Essa é uma preocupação que pode ser reforçada com o avanço de redes de cooperação na Região Sul, o que configura facilidade à articulação de empresas já tradicionalmente ligadas por laços familiares e/ou societários.

De posse de conhecimentos como esses, os auditores que atuam no CCIMAR, no CCIEx ou no CENCIAR podem melhor planejar suas atividades de avaliação, adicionando o uso

da MD às técnicas de análise de riscos para seleção dos objetos de auditoria. Além disso, podem valer-se da metodologia aplicada para, já na fase de execução das atividades, debruçar-se sobre indícios que comporão as evidências de auditoria na área de licitações.

Quanto a limitações, ressalta-se que, nesta pesquisa, foi utilizado apenas um *software* para a verificação das associações frequentes das empresas participantes dos pregões. O uso de outras ferramentas em paralelo poderia resultar em validação das regras de associação geradas no *software* R.

Assim, sugere-se a realização de estudos futuros, com vistas a: i) proceder, na base de dados desta pesquisa, à aplicação de outras técnicas de MD, comparando os resultados; ii) analisar, separadamente, possíveis situações de cartel, por meio de técnicas de MD, em licitações realizadas pelas UASG de cada uma das Forças; e iii) aplicar técnicas de MD em outras áreas de interesse da auditoria interna nas Forças Armadas.

## NOTAS

1. Dois dos casos mais importantes que já passaram pelo Conselho Administrativo de Defesa Econômica envolvem licitações públicas: a apuração sobre o cartel de empreiteiras nas contratações da estatal Petrobras (Processo Administrativo nº 08700.007777/2016-95), parte do esquema de corrupção investigado também pela Operação Lava Jato, e o cartel em licitações para construção de linhas de trens e metrô na cidade de São Paulo (PA nº 08700.004617/2013-41), com condenação recente.
2. <http://compras.dados.gov.br/docs/home.html>.
3. Acesso ao arquivo CSV utilizado no *software* R: <http://tinyurl.com.br/k3>.
4. Segundo a Associação Nacional dos Comerciantes de Material de Construção (Anamac), o segmento de comércio varejista de material de construção é constituído por cerca de 150 mil lojas em todo o país, sendo a maioria pequenas e médias empresas. Fonte: <http://www.grandesconstrucoes.com.br/Noticias/Exibir/mais-de-50-das-lojas-de-material-de-construcao-no-brasil-nao-tem-gestao-profissionalizada>.

5. Processo Administrativo nº 08012.006199/2009-07.

6. <https://sedetur.rs.gov.br/redes-de-cooperacao>.

## REFERÊNCIAS

- ALBUQUERQUE, Felipe Braga; LEAL, Leonardo José Peixoto. Prática de cartel no Brasil: um estudo sobre as decisões do CADE e o perfil das condenações por cartel. *CONPEDI Law Review*, Madri, v.1, n. 8, p. 66-87, 2015.
- ANDRADE, Rodolpho Pinto de. **Reparação de danos decorrentes da prática de cartel**. 2019. Dissertação (Mestrado em Direito Político e Econômico) — Faculdade de Direito, Universidade Presbiteriana Mackenzie, São Paulo. 2017. Disponível em: <http://tede.mackenzie.br/jspui/bitstream/tede/3929/5/Rodolpho%20Pinto%20de%20Andrade.pdf>. Acesso em: 18 set. 2019.
- ARAÚJO, Bettina Cavalcanti; MACIEL, Alexandre Magno Andrade. Aplicação de regras de associação em dados da criminalidade da cidade do Recife. *Revista de Engenharia e Pesquisa Aplicada*, Recife, vol. 3, n. 3. p. 158-168. ago. 2018. Disponível em: <https://doi.org/10.25286/rep.v3i3.974>. Acesso em: 14 set. 2019.
- BAENSENS, Bart; VLASSELAER, Véronique Van; VERBEKE, Wouter. **Fraud analytics: using descriptive, predictive [...] techniques: a guide [...]**. 1 ed. Cary: Wiley, 2015.
- BALDOMIR, Rebeca Andrade. **Aplicação do algoritmo Apriori para detectar relacionamentos entre empresas nos processos licitatórios do Governo Federal**. 2017. Monografia (Bacharelado em Ciências da Computação) — Universidade de Brasília, Brasília, 2017. Disponível em: [http://bdm.unb.br/bitstream/10483/19987/1/2017\\_RebecaAndradeBaldomir\\_tcc.pdf](http://bdm.unb.br/bitstream/10483/19987/1/2017_RebecaAndradeBaldomir_tcc.pdf). Acesso em: 19 set. 2019.
- BOGUSKI, Rodrigo Ruy; CURY, Davidson. Usando regras de associação para a identificação de falhas conceituais. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 29., 2018, Fortaleza. *Anais [...]*. Fortaleza: SBC, 2018. Disponível em: <https://br-ie.org/pub/index.php/sbie/article/view/8105>. Acesso em: 16 set. 2019.
- BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília: Presidência da República, 1988. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 21 set. 2019.

- BRASIL. Controladoria-Geral da União. **Instrução Normativa nº 3, de 9 de junho de 2017**. Aprova o Referencial Técnico da Atividade de Auditoria Interna Governamental do Poder Executivo Federal. Brasília: Controladoria-Geral da União, 2017a.
- BRASIL. Controladoria-Geral da União. **Instrução Normativa nº 8, de 6 de dezembro de 2017**. Aprova o Manual de Orientações Técnicas da Atividade de Auditoria Interna Governamental do Poder Executivo Federal. Brasília: Controladoria-Geral da União, 2017b.
- BRASIL. **Decreto nº 5.450, de 31 de maio de 2005**. Regulamenta o pregão, na forma eletrônica, [...] e dá outras providências. Brasília: Presidência da República, 2005. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2004-2006/2005/Decreto/D5450.htm](http://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2005/Decreto/D5450.htm). Acesso em: 26 set. 2019.
- BRASIL. **Lei nº 10.520, de 17 de julho de 2002**. Institui [...] modalidade de licitação denominada pregão [...] e dá outras providências. Brasília: Presidência da República, 2002. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/LEIS/2002/L10520.htm](http://www.planalto.gov.br/ccivil_03/LEIS/2002/L10520.htm). Acesso em: 21 set. 2019.
- BRASIL. **Lei nº 8.666, de 21 de junho de 1993**. [...] institui normas para licitações e contratos da Administração Pública e dá outras providências. Brasília: Presidência da República, 1993. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/l8666cons.htm](http://www.planalto.gov.br/ccivil_03/leis/l8666cons.htm). Acesso em: 20 set. 2019.
- BRASIL. Tribunal de Contas da União. [...] Verificação da consistência e da confiabilidade dos dados dos sistemas SIASG e COMPRASNET [...]. **Acórdão nº 1.793/2011**. Brasília, 06 de julho de 2011. Disponível em: <https://www.lexml.gov.br/urn/urn:lex:br:tribunal.contas.uniao:plenario:acordao:2011-07-06;1793>. Acesso em: 25 set. 2019.
- DE BATTISTA *et al.* Minería de Datos aplicada a datos masivos. In: WORKSHOP DE INVESTIGADORES EN CIENCIAS DE LACOMPUTACIÓN, 18., 2016, Entre Ríos. **Libro de actas** [...]. Entre Ríos: UNER, 2016. p. 288-292. Disponível em: <http://sedici.unlp.edu.ar/handle/10915/52901>. Acesso em: 3 set. 2019.
- FAYYAD, Usama; PIATETSKY-SAPHIRO, Gregory; SMYTH, Padhraic. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, [s.l.], vol. 17, n. 3, p. 37-54, 1996. Disponível em: <https://doi.org/10.1609/aimag.v17i3.1230>. Acesso em: 16 set. 2019.
- FERREIRA, Maurício Alencar Alves. **O cartel e suas práticas nas licitações públicas**. 2015. Monografia (Bacharelado em Direito) — Universidade de Brasília, Brasília, 2015. Disponível em: <http://bdm.unb.br/handle/10483/9777>. Acesso em: 20 set. 2019.
- FLORES, A. M. As grandes redes de varejo no mercado de casa e construção e sua atuação no espaço regional gaúcho através do setor terciário. **Boletim Geográfico do Rio Grande do Sul**, Porto Alegre, n. 31, p. 101-119, jun. 2018.
- GIL, Antonio. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.
- GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel; BEZERRA, Eduardo. **Data Mining: conceitos, técnicas, [...]**. 2. ed. Rio de Janeiro: Elsevier, 2015.
- HAN, Jiawei; KAMBER, Micheline; PEI, Jian. **Data Mining: concepts and techniques**. 3. ed. Waltham: Morgan Kaufmann, 2012.
- JAIN, Anil K.; DUBES, Richard C. **Algorithms for clustering data**. Englewood Cliffs: Prentice-Hall, 1988.
- LOKBANI, Ahmed Chaouki. **Le problème de sécurité par le Data Mining**. 2017. Tese (Doutorado em Informática) — Université Djillali Liabès, Sidi Bel Abbès, 2017. Disponível em: <http://rdoc.univ-sba.dz/handle/123456789/2087>. Acesso em: 20 set. 2019.
- MELO, Bruno Mattos Souza de Souza; SILVA, Marcos Ferreira da. Análise de dados no planejamento e na execução de auditorias governamentais. In: NAZARETH, Paula Alexandra C. P.; CHAISE, Rosa Maria; MIRANDA, Maria Bernadette Dayrell de (org.). **Programa de Pesquisa: coletânea de artigos** [...]. Rio de Janeiro: ECG/TCE-RJ, 2016.
- NUNES, Daniele de Oliveira. Cartéis em licitações: fronteiras entre a infração à ordem econômica e o ato de improbidade administrativa. **Revista de Defesa da Concorrência**, Brasília, vol. 4, n. 1, p. 185-210, mai. 2016.
- OCDE. **Preventing corruption in public procurement**. [S.l.]: OCDE, 2016. Disponível em: <http://www.oecd.org/gov/ethics/Corruption-Public-Procurement-Brochure.pdf>. Acesso em: 07 ago. 2019.
- OCDE. **Recommendation of the Council concerning effective action against hard core cartels**. OCDE, 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0452>. Acesso em: 12 set. 2019.
- PEREIRA, Júlio César Brasil *et al.* Patterns identification in the consumption of Oracle database resources using Weka tool in Data Mining. **Journal of Engineering and Technology for Industrial Applications**, Manaus, vol. 5, n. 17, mar. 2019. Disponível em: <http://itegam-jetia.org/journal/index.php/jetia/article/download/296/238>. Acesso em: 23 set. 2019.

PRODANOV, Cleber C.; FREITAS, Ernani C. **Trabalho científico: métodos e técnicas** [...]. 2. ed. Novo Hamburgo: Feevale, 2013.

RIBEIRO, Cássio Garcia; INÁCIO JÚNIOR, Edmundo. **O mercado de compras governamentais brasileiro (2006-2017): mensuração e análise**. Brasília: Ipea, 2019. Disponível em: [http://www.ipea.gov.br/portal/index.php?option=com\\_content&view=article&id=34850](http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=34850). Acesso em: 28 set. 2019.

SILVA, Antonio Hevertton Martins; SANTOS, Wallan Azevedo dos; HORA, Henrique Rego Monteiro da. Identificação de padrões de falhas em válvulas submarinas utilizando Mineração de Dados. **Revista Cereus**, Gurupi, v. 10, n. 2, 2018. Disponível em: <http://ojs.unirg.edu.br/index.php/1/article/view/2215>. Acesso em: 24 set 2019.

WEISHAAR, Stefan E. **Cartels, competition and public procurement: law and economics approaches to bid rigging**. Northampton: Edgar Elgar, 2013.