XXI ENTAC – Encontro Nacional de Tecnologia do Ambiente Construído
14 a 16 de outubro de 2026 | Goiânia-Brasil
DOI: https://doi.org/10.46421/entac.v21i1.8342

Mineração de dados aplicada à gestão de contratos de obras públicas: descoberta de padrões por regras de associação

Data Mining Applied to the Management of Public Construction Contracts: Pattern Discovery through Association Rules

Thiago Araújo Fernandes do Nascimento
Instituto Militar de Engenharia | Rio de Janeiro | Brasil |nascimento.thiago@ime.eb.br
Diana Schrok Bezerra
Instituto Militar de Engenharia | Rio de Janeiro | Brasil | diana.schrok@ime.eb.br
Paulo César Pellanda
Instituto Militar de Engenharia | Rio de Janeiro | Brasil |pellanda@ime.eb.br
Giuseppe Miceli Junior
Instituto Militar de Engenharia | Rio de Janeiro | Brasil |giusepe.miceli@ime.eb.br

Resumo

Este trabalho utiliza técnicas de mineração de dados para identificar padrões em contratos de obras públicas, com foco nas relações entre prazo, valor, aditivos e perfil do órgão executor. A partir de uma base real de contratos públicos, aplicou-se o algoritmo FP-Growth, da família das regras de associação, no ambiente RapidMiner, após pré-processamento e discretização das variáveis contínuas. As regras obtidas, com elevado suporte e confiança, indicaram maior regularidade em obras de pequeno porte executadas por órgãos especializados, maior incidência de aditivos em contratos de alto valor ou longa duração, correlação entre aditivos de prazo e valor e concentração desses eventos em determinados órgãos executores. O estudo demonstra a viabilidade da ciência de dados como apoio à administração pública e contribui com subsídios objetivos para aprimorar a gestão contratual e as ações de controle.

Palavras-chave: Mineração de dados. Regras de associação. Gestão Contratual. Apoio à decisão. Obras públicas

Abstract

This study uses data mining techniques to identify patterns in public construction contracts, focusing on the relationships among duration, value, amendments, and the profile of the executing agency. Based on a real database of public contracts, the FP-Growth algorithm, from the association rule mining family, was applied in the RapidMiner environment after preprocessing and discretization of continuous variables. The rules obtained, with high support and confidence, indicated greater regularity in small-scale projects executed by specialized agencies, higher incidence of amendments in high-value or long-duration contracts, correlation between time and value amendments, and concentration of these events in specific executing agencies. The study demonstrates the feasibility of data science as support for public administration and contributes objective evidence to improve contract management and control actions.

Keywords: *Data mining. Association rules. Contract management. Decision support. Public works.*

NASCIMENTO, T. A. F.; BEZERRA, D. S.; PELLANDA, P. C.; MICELI JUNIOR, G. Mineração de dados aplicada à gestão de contratos de obras públicas: descoberta de padrões por regras de associação. In: ENCONTRO NACIONAL DE TECNOLOGIA DO AMBIENTE CONSTRUÍDO, 21., 2026, Goiânia. Anais... Goiânia: ANTAC, 2026.

DOI: https://doi.org/10.46421/entac.v21i1.8342

INTRODUÇÃO

A execução de obras públicas representa atividade relevante para a administração pública, tanto pelo volume de recursos envolvidos quanto por seu impacto sobre a infraestrutura [1]. Nesse contexto, desafios ligados a prazos, aditivos e eficiência de execução tornam pertinente o uso de técnicas de ciência de dados para extração de conhecimento a partir de bases administrativas e apoio à gestão contratual [2].

Embora auditorias operacionais, análises estatísticas descritivas e estudos comparativos sejam amplamente utilizados, tais abordagens nem sempre capturam a complexidade e a interdependência das variáveis envolvidas [3][4]. Em resposta a essa limitação, técnicas de mineração de dados e aprendizado de máquina têm sido empregadas para identificar padrões ocultos e relações significativas em grandes bases contratuais [5]. Entre essas técnicas, a descoberta de regras de associação se destaca por revelar correlações relevantes entre variáveis categóricas. Neste trabalho, adotou-se o algoritmo FP-Growth, reconhecido por sua eficiência na mineração de padrões frequentes e por gerar regras interpretáveis para apoio à decisão [6][7].

Diante disso, este artigo tem como objetivo aplicar técnicas de mineração de dados, por meio do algoritmo FP-Growth, para identificar padrões frequentes e relações relevantes em contratos de obras públicas. A proposta consiste em transformar uma base operacional em regras de associação compreensíveis, capazes de apoiar atividades de planejamento, fiscalização e controle.

O estudo foi desenvolvido com base nas etapas do processo de descoberta de conhecimento em bases de dados, envolvendo pré-processamento, discretização de variáveis e extração de regras de associação no ambiente RapidMiner. A análise permitiu identificar padrões consistentes entre atributos contratuais, especialmente quanto às relações entre prazo, valor e ocorrência de aditivos, evidenciando o potencial da abordagem para qualificar diagnósticos e aprimorar a gestão de contratos públicos.

METODOLOGIA

CARACTERIZAÇÃO DO MÉTODO DE PESQUISA

Do ponto de vista metodológico, este estudo caracteriza-se como pesquisa documental, de natureza descritiva e abordagem quantitativa. A pesquisa documental utiliza registros e documentos como fonte principal de evidência, sem intervenção direta do pesquisador na produção dos dados, sendo adequada à análise de bases administrativas previamente constituídas [8]. Neste trabalho, a fonte documental corresponde a registros contratuais de obras públicas, dos quais foram extraídas variáveis relacionadas a prazo, valor, aditivos, situação da obra e perfil do órgão executor.

O caráter descritivo decorre do objetivo de identificar, organizar e interpretar padrões recorrentes na base analisada, sem testar relações causais ou realizar intervenção experimental sobre o fenômeno. A análise busca descrever associações entre atributos

contratuais e evidenciar regularidades relevantes para a gestão pública, permitindo compreender como determinadas características dos contratos se relacionam à ocorrência de aditivos e ao desempenho da execução.

A abordagem quantitativa decorre do tratamento estruturado dos dados e da aplicação de técnica computacional de mineração de dados. Após o pré-processamento e a transformação das variáveis, utilizou-se o algoritmo FP-Growth para extrair regras de associação, com base em métricas como suporte, confiança e lift. Assim, o delineamento combina pesquisa documental e descritiva com mineração quantitativa de dados, permitindo converter registros administrativos em padrões interpretáveis de apoio à decisão.

BASE DE DADOS

A base de dados utilizada neste estudo foi extraída de um sistema administrativo de acompanhamento de obras públicas, que reúne informações operacionais e gerenciais sobre contratos de obras e serviços de engenharia executados em diferentes unidades gestoras. Foram considerados apenas contratos ativos ou encerrados com registros consolidados até o momento do processamento, totalizando 4.914 contratos distintos, com datas de início entre junho de 2011 e dezembro de 2024. A base contempla atributos administrativos, técnicos e contratuais, incluindo identificação do contratante, valores inicial e final, prazos pactuados e executados, aditivos de prazo e valor, percentual físico de execução e situação da obra. Esse conjunto de variáveis permitiu examinar padrões estruturais da execução contratual sob diferentes perspectivas analíticas.

PRE-PROCESSAMENTO DE DADOS

A preparação da base seguiu um fluxo estruturado de pré-processamento, com o objetivo de garantir consistência, relevância analítica e adequação das variáveis à aplicação do algoritmo FP-Growth. Inicialmente, foram removidos registros duplicados, dados ausentes críticos e valores inconsistentes, especialmente aqueles que poderiam comprometer a discretização ou distorcer os padrões extraídos. Ao final dessa etapa, a base foi reduzida para 3.603 registros válidos.

Em seguida, procedeu-se à discretização das variáveis contínuas, etapa necessária para a aplicação de técnicas de regras de associação, que operam com variáveis categóricas. As variáveis valor inicial do contrato e prazo contratado foram categorizadas em faixas definidas com base na distribuição estatística da própria base, de modo a refletir portes de obra e níveis de complexidade temporal. Já os percentuais de aditivo de prazo e de valor foram discretizados em intervalos capazes de distinguir diferentes níveis de alteração contratual. No caso do aditivo de valor, a categorização também considerou os limites legais aplicáveis aos contratos de obras e serviços de engenharia.

Além da discretização, foram construídas variáveis derivadas com base em critérios lógicos e qualitativos, como a fase da obra, definida a partir do percentual executado, e a classificação do órgão executor quanto à especialização em atividades de engenharia. Por fim, as variáveis categóricas foram transformadas em formato binário

(*one-hot encoding*), conforme requerido para a extração de padrões frequentes no ambiente de mineração de dados adotado.

TÉCNICA APLICADA E CONFIGURAÇÃO DO EXPERIMENTO

Para a mineração de padrões frequentes, foi adotado o algoritmo FP-Growth (*Frequent Pattern Growth*), amplamente utilizado na descoberta de regras de associação em bases categorizadas. Sua escolha se justifica pela eficiência computacional na identificação de itemsets frequentes, sobretudo em bases com grande número de atributos binarizados, evitando a geração exaustiva de conjuntos candidatos característica de algoritmos como o Apriori. Neste estudo, o algoritmo foi executado no RapidMiner Studio, plataforma utilizada para organização do fluxo analítico, transformação dos dados e geração das regras de associação.

O experimento foi estruturado segundo a lógica do processo de descoberta de conhecimento em bases de dados, envolvendo preparação da base, extração de padrões e avaliação das regras obtidas. As execuções foram realizadas tanto com a base completa quanto com subconjuntos de atributos selecionados, a fim de capturar padrões gerais e relações específicas entre variáveis como valor inicial, prazo, aditivos, perfil do órgão executor e situação da obra. O algoritmo foi configurado com suporte mínimo de 5% e confiança mínima de 80%, parâmetros definidos para equilibrar volume de regras, robustez estatística e utilidade interpretativa.

As regras geradas foram avaliadas com base nas métricas de suporte, confiança e *lift*, priorizando-se associações positivas e descartando-se regras redundantes ou triviais. A seleção final considerou, de forma combinada, a robustez estatística das regras, sua clareza interpretativa e seu potencial de aplicação na gestão contratual, permitindo identificar padrões relevantes para planejamento, fiscalização e controle de contratos de obras públicas.

RESULTADOS E DISCUSSÃO

Após a aplicação do algoritmo FP-Growth sobre a base discretizada, foram geradas centenas de regras de associação. Para esta seção, selecionaram-se cinco regras não triviais e não redundantes, com base em três critérios: elevado suporte e confiança, relevância prática para a gestão de contratos de obras públicas e capacidade de revelar associações não evidentes por simples inspeção dos dados. A análise buscou interpretar esses padrões à luz da gestão contratual, evidenciando seu potencial de aplicação em planejamento, fiscalização e controle.

EFICIÊNCIA SUPERIOR DE ÓRGÃOS EXECUTORES ESPECIALIZADOS EM OBRAS DE MENOR PORTE

Um dos padrões mais expressivos identificados mostra que contratos executados por unidades especializadas, com valor inicial inferior a R$ 218 mil, resultaram em obras concluídas sem aditivos em 96% dos casos. Para contratos da mesma faixa de valor, unidades não especializadas apresentaram apenas 63% de confiança para conclusão sem aditivos, o que representa diferença superior a 30 pontos percentuais. O resultado

sugere que, em obras de pequeno porte, a especialização do órgão executor influencia diretamente a regularidade contratual e a aderência ao escopo inicialmente pactuado.

Esse comportamento pode ser interpretado à luz da maior maturidade técnico-operacional das unidades especializadas, que tendem a reunir experiência acumulada, rotinas mais consolidadas de planejamento e acompanhamento, além de maior familiaridade com a gestão de riscos típicos da execução de obras. O próprio fato de unidades não especializadas raramente executarem contratos de maior porte reforça a concentração da expertise em determinados executores. Em termos gerenciais, o achado sugere que a especialização institucional não deve ser vista apenas como característica organizacional, mas como variável estratégica para alocação de contratos, especialmente em contextos em que a administração busca maior previsibilidade de prazo e custo.

CONTRATOS DE MAIOR VALOR INICIAL APRESENTAM MAIOR PROPENSÃO A ADITIVOS

Outra regra de destaque mostrou que contratos com valor inicial superior a R$ 690 mil apresentaram 92% de confiança para ocorrência de aditivos de valor, com suporte de 48,7%, o que significa que praticamente metade dos contratos dessa faixa integra esse padrão. O resultado indica que o porte financeiro do contrato funciona como importante marcador de complexidade contratual, associando-se a maior probabilidade de modificações orçamentárias durante a execução.

Essa associação é consistente com a lógica de execução de obras mais complexas, que tendem a envolver maior número de frentes de serviço, dependência de insumos e serviços especializados, exposição mais prolongada a oscilações de mercado e maior suscetibilidade a ajustes de escopo. Além disso, contratos de maior valor normalmente também apresentam maior duração, o que amplia a chance de ocorrência de eventos supervenientes com impacto financeiro. Sob a ótica da gestão pública, esse padrão é relevante porque permite tratar o valor inicial como sinal de atenção já na fase de planejamento, justificando maior rigor na elaboração de projetos, no detalhamento técnico, na avaliação de riscos e no desenho de mecanismos de acompanhamento contratual. Em vez de ser lido apenas como dado financeiro, o valor inicial do contrato passa a operar como variável com capacidade preditiva sobre a necessidade de aditivos.

CONCENTRAÇÃO DE ADITIVOS EM UNIDADES EXECUTORAS ESPECIALIZADAS ESPECÍFICAS

Também foram identificadas regras que associam elevada incidência de aditivos de valor e prazo a unidades executoras especializadas específicas, com confiança superior a 95%. O ponto central aqui não é apenas a recorrência de aditivos, mas sua concentração em parte do universo de executores especializados, o que indica que a ocorrência de alterações contratuais não se distribui de forma homogênea nem mesmo entre órgãos vocacionados à execução de obras.

Esse padrão admite diferentes interpretações. Pode refletir características regionais que afetam diretamente a execução contratual, como condições logísticas,

disponibilidade de fornecedores, sazonalidade climática ou restrições orçamentárias. Pode também decorrer do perfil do portfólio de obras dessas unidades, eventualmente mais complexo, emergencial ou sujeito a maior variabilidade técnica. Há ainda a possibilidade de que diferenças nos modelos de gestão e fiscalização influenciem a frequência com que alterações são formalizadas contratualmente. Por isso, o achado não deve ser interpretado como evidência automática de pior desempenho, mas como indicador de concentração de complexidade contratual. Em termos práticos, esse tipo de padrão pode orientar monitoramentos diferenciados por unidade executora, com foco em compreender se a recorrência de aditivos decorre de fatores estruturais, operacionais ou gerenciais.

ADITIVOS DE VALOR E PRAZO TENDEM A OCORRER CONJUNTAMENTE

Um dos achados mais consistentes da análise foi a forte associação entre aditivos de valor e aditivos de prazo. A regra extraída mostrou que, quando um contrato apresentou aditivo financeiro, houve alta probabilidade de ocorrência simultânea de aditivo temporal, com confiança superior a 80%. Esse resultado evidencia que as dimensões financeira e temporal da execução contratual não devem ser tratadas como fenômenos isolados, mas como dimensões interdependentes de um mesmo processo de ajuste contratual.

Do ponto de vista substantivo, a associação faz sentido: acréscimos de escopo, reequilíbrios econômico-financeiros, necessidade de serviços adicionais ou imprevistos técnicos frequentemente repercutem no cronograma, exigindo reprogramação de etapas, extensão de prazos e readequação do planejamento. O valor analítico da regra está em formalizar quantitativamente essa relação, oferecendo base para uma abordagem de fiscalização mais integrada. Em lugar de monitorar cada tipo de aditivo separadamente, a administração pode utilizar a identificação de um como indício objetivo da probabilidade de ocorrência do outro, construindo alertas gerenciais, indicadores de risco e rotinas de acompanhamento orientadas por evidências. Esse é um dos pontos em que a mineração de dados se mostra mais útil: transformar relações conhecidas empiricamente em padrões explicitados e passíveis de incorporação a processos decisórios.

OBRAS MUITO LONGAS APRESENTAM MAIOR RISCO DE PRESSÃO ORÇAMENTÁRIA

A última regra selecionada mostra que contratos com prazo superior a 720 dias apresentaram alta incidência de aditivos de valor, com confiança de 86,3% e suporte de 6,7%. Ainda que esse grupo seja numericamente menor, o padrão é estatisticamente relevante e particularmente importante do ponto de vista gerencial, porque demonstra que a duração contratual, por si só, pode funcionar como fator de risco para desequilíbrios financeiros.

A explicação para esse comportamento é plausível: quanto maior o prazo de execução, maior a exposição do contrato a variações de preços, alterações de projeto, entraves administrativos, fatores climáticos, descontinuidades orçamentárias e riscos operacionais acumulados. O aspecto mais relevante é que esse padrão se manifestou mesmo em contratos cujo valor inicial não era necessariamente elevado, o que mostra

que a duração não atua apenas como variável acessória do porte financeiro, mas como dimensão autônoma de risco contratual. Sob a ótica da gestão, isso sugere que contratos muito longos devem receber tratamento diferenciado desde a fase de planejamento, com mecanismos mais robustos de contingência, acompanhamento mais próximo e maior atenção à sustentabilidade financeira da execução ao longo do tempo.

SÍNTESE ANALÍTICA DOS RESULTADOS

Em conjunto, os resultados mostram que a mineração de regras de associação permitiu identificar padrões relevantes e operacionalmente úteis em bases contratuais da administração pública. Os achados convergem para uma leitura estruturada da gestão contratual: a especialização do executor tende a elevar a regularidade em contratos menores; o aumento do valor inicial e da duração amplia a probabilidade de aditivos; e alterações de prazo e valor tendem a ocorrer de forma articulada. Além disso, a concentração de aditivos em unidades específicas sugere que a heterogeneidade organizacional e contextual também precisa ser considerada nas estratégias de monitoramento.

Esses resultados não apenas confirmam tendências observadas na prática administrativa, mas acrescentam um ganho importante de formalização analítica. Ao converter regularidades empíricas em regras quantitativas, o estudo oferece base para a criação de indicadores, filtros de priorização, critérios de alocação de recursos e rotinas de fiscalização mais seletivas. Assim, a principal contribuição desta seção não está apenas na descrição dos padrões encontrados, mas na demonstração de que técnicas de mineração de dados podem apoiar uma gestão contratual mais proativa, comparável e orientada por evidências.

CONCLUSÃO

Este trabalho investigou a aplicabilidade da mineração de regras de associação, por meio do algoritmo FP-Growth, para identificar padrões relevantes em contratos de obras públicas a partir de dados administrativos. A análise de uma base real permitiu revelar associações consistentes entre variáveis contratuais, demonstrando que técnicas de ciência de dados podem transformar registros operacionais em informações úteis à gestão e ao controle. Entre os principais achados, destacam-se a maior regularidade de contratos de menor porte executados por unidades especializadas, a elevada incidência de aditivos em contratos de maior valor ou longa duração, a concentração de aditivos em unidades executoras específicas e a forte correlação entre aditivos de prazo e de valor.

As contribuições do estudo concentram-se em dois planos. No plano metodológico, demonstra-se a viabilidade do uso de técnicas de mineração de dados em bases administrativas do setor público com apoio de ferramentas acessíveis. No plano aplicado, os padrões identificados oferecem subsídios objetivos para o aprimoramento do planejamento contratual, da alocação de recursos, da priorização de unidades especializadas e do direcionamento da fiscalização. Ao formalizar essas relações em

regras quantitativas e interpretáveis, o estudo reforça o potencial de abordagens orientadas por dados para qualificar a gestão contratual na administração pública.

DECLARAÇÃO

Durante a preparação deste artigo, o primeiro autor utilizou o ChatGPT, modelo GPT-5.5 Thinking, da OpenAI, exclusivamente para revisão final, aprimoramento da redação científica, melhoria de coesão e ajustes de clareza textual. Após a utilização da ferramenta, os autores revisaram e editaram o conteúdo conforme necessário e assumem total responsabilidade pelo conteúdo do artigo publicado.

REFERÊNCIAS

[1] SANTOS, N. J.; LIMA, L. O. **Obras públicas: aspectos legais, planejamento e execução**. *Revista Foco – Public Works*, v. 16, n. 12, 2023.

[2] REVISTA DA CGU. Editorial: **Ciência de dados na administração pública: desafios e oportunidades**. *Revista da CGU*, v. 15, n. 27, jan.-jun. 2023. DOI: 10.36428/revistadacgu.v14i27.665.

[3] ALBUQUERQUE, F. F. Tenório. ***A auditoria operacional e seus desafios: um estudo a partir da experiência do Tribunal de Contas da União***. 2006. Dissertação (Mestrado em Administração) – Universidade Federal da Bahia, Salvador.

[4] BRASIL. Ministério do Planejamento, Orçamento e Gestão (MPOG). ***Metodologia de benchmarking colaborativo: guia prático para uso na gestão pública orientada a resultados***. Brasília: MPOG, 2013.

[5] MODRUŠAN, N.; RABUZIN, K. **Improving public sector efficiency using advanced text mining in the procurement process**. In: *Conferência Internacional*, 2020.

[6] HAN, J.; PEI, J.; YIN, Y. **Mining frequent patterns without candidate generation: a performance study of FP-Growth vs Apriori**. In: *Proceedings of the ACM SIGMOD International Conference on Management of Data*, 2000.

[7] LAWAL, M. M.; OGBEDENGBE, T. M. **FP-Growth Algorithm: Mining Association Rules without Candidate Sets Generation**. *Kasu Journal of Computer Science*, v. 1, n. 2, p. 392–411, jun. 2024. DOI: 10.47514/kjcs/2024.1.2.0016.

[8] GIL, A. C. **Como elaborar projetos de pesquisa**. 6. ed. São Paulo: Atlas, 2017.