

Contabilidade Pública – uma Ferramenta de Suporte à Auditoria de Dados em Municípios do Estado de Pernambuco

George G. Cabral¹, George Valença¹, Camila N. P. Souza¹, Yves Emmanuel²

¹Departamento de Computação – UFRPE, Recife/PE – Brazil

²Centro de Informática – UFPE, Recife/PE – Brazil

{george.gcabral, george.valenca}@ufrpe.br, camilanpsouza@gmail.com

yefo@cin.ufpe.br

Abstract. *Social welfare is undoubtedly related to equality and social development. These in turn are characteristics of societies with a low level of corruption. A low level of corruption, theoretically, is something only achievable by developed nations with a high degree of transparency in public accounts. On the other hand, underdeveloped nations struggle with corruption while the bodies responsible for overseeing accounts try to cope with the high demand for audits. In this work, three methods are proposed that can extract information from raw public data in order to assist the human auditor in deciding which public entities need more attention. For this, it was necessary to develop a preliminary step of data extraction and transformation in order to create a database that contains information necessary to solve the three methods above mentioned.*

Resumo. *O bem-estar social é sem dúvida relacionado à igualdade e desenvolvimento social. Por sua vez, esses são características de sociedades com baixo nível de corrupção. Um baixo nível de corrupção, teoricamente, é algo só alcançável por nações desenvolvidas e com um alto grau de transparência nas contas públicas. Por sua vez, nações subdesenvolvidas sofrem com a corrupção enquanto os órgãos responsáveis por fiscalizar contas tentam dar conta da alta demanda por auditorias. Nesse trabalho são propostos três métodos capazes de extrair informações de dados públicos brutos de forma a auxiliar o auditor humano na decisão de quais entidades públicas necessitam de uma maior atenção. Para isso, foi necessário o desenvolvimento de uma etapa preliminar de extração e transformação dos dados de maneira a se criar uma base de dados que contenha informações necessárias à solução dos três métodos acima mencionadas.*

1. Introdução

A corrupção pode ser vista como um tema de interesse global diretamente ligado ao bem estar e igualdade social. A organização Transparency International¹ investiga cerca de 180 países e mede um índice de percepção de corrupção nesses países. Em 2022, a Dinamarca foi classificada como o país menos afetado pela corrupção ao passo que a Somália foi tida como a nação mais prejudicada. Nesse mesmo ranking, o Brasil se situa na posição 94

¹<https://www.transparency.org/>

junto com países como a Etiópia, Argentina, Marrocos e Tanzânia. É importante frisar que a percepção de corrupção no Brasil mostra uma piora de 2012 (início da série histórica) a 2022, mesmo com a disponibilização de instrumentos transparência de dados públicos como, por exemplo, portais de transparência [de Oliveira Rego 2021] de governos municipais, estaduais e federal. Por outro lado, países com economias fortes e baixa percepção de corrupção também têm um papel importante na corrupção global ao passo que esses aceitam riquezas advindas como consequência da corrupção em países menos desenvolvidos.

Na máquina pública as formas de praticar a corrupção [de Araujo 2014] são inúmeras e, de certa forma, contextualizadas de acordo com o país em questão. Alguns exemplos comuns no cenário brasileiro são: (i) apropriação ou desvio de dinheiro público (peculato) - quando um agente público utiliza de recursos financeiros públicos para fins de natureza pessoal; (ii) propina - quando um agente público exige um pagamento para beneficiar terceiros dentro do setor público; (iii) superfaturamento - quando em comum acordo setores público e privado aumentam valores de produtos e (ou) serviços no intuito de tomarem vantagem parcial no valor total do contrato; (iv) conluio em licitações - quando uma ou mais empresas combinam valores de propostas em um certame de forma a decidir antecipadamente o vencedor desse; etc.

A receita pública municipal (ou orçamento) é formada por um grande número de tipos de fontes de recurso. Três principais categorias englobam essas fontes: recursos próprios, transferências constitucionais e transferências voluntárias. Os recursos próprios são majoritariamente obtidos através de impostos e taxas como Imposto sobre a Propriedade Predial e Territorial Urbana (IPTU), Imposto sobre Transmissão 'Inter Vivos' de Bens Imóveis e de Direitos Reais sobre Imóveis (ITBI), Taxa de Coleta de Lixo, etc. As transferências constitucionais são recursos obtidos da esfera federal e estadual. As transferências voluntárias são recursos repassados aos municípios através de programas, emendas parlamentares e compensação financeiras, por exemplo. É importante salientar que as fontes de recursos podem ser vinculadas a uma finalidade (despesa) específica (exemplo: emenda parlamentar para construção de creches) ou não (exemplo: valor arrecadado de IPTU).

Na contabilidade pública, toda despesa gerada por compra de bens ou serviços é realizada através de empenhos. Um empenho consiste em um contrato entre uma entidade pública e um fornecedor (empresa ou pessoa física) de forma que um recurso financeiro é alocado para a garantia de pagamento do empenho. Por sua vez, o pagamento pode ser efetuado em uma ou diversas parcelas, dependendo do contrato. Nesse contexto, um ato bastante importante é a liquidação parcial ou total do empenho. Uma liquidação consiste na formalização da entrega parcial ou total do produto ou serviço, objeto daquele empenho. Uma vez que a liquidação ocorre, ela se torna elegível para pagamento.

O atraso nos pagamentos de liquidações, assim como o pagamento antecipado de liquidações posteriores a ela, pode configurar indícios de irregularidades em uma unidade gestora (i.e., entidade pública responsável pela contratação e pagamento de bens ou serviços, como a secretaria de saúde do município). Adicionalmente, a contratação de empenhos pode também ser usada de forma irregular para outras finalidades (e.g., pagamento de salários). Diante destes problemas, este trabalho apresenta métodos para auxiliar à resposta das seguintes questões de pesquisa (QP): *como auxiliar profissionais*

de auditoria de contas públicas na esfera municipal na identificação de: (i) potenciais casos de pagamentos de salário através de empenhos?; (ii) municípios com alto grau de atrasos de pagamentos?; e (iii) municípios com alto grau de desrespeito à ordem cronológica de pagamentos? Para isso, usamos os municípios do estado de Pernambuco como ambiente de experimentação da ferramenta.

Como principais resultados, temos:

- Criação de um método para extração, carregamento e transformação dos dados (ELT) dos municípios culminando com a criação de um banco de dados contendo informações sobre a conexão de cada pagamento a sua respectiva liquidação;
- Criação de um método capaz de destacar casos potenciais de pagamentos de salários através de empenhos de modo a trazer essa informação de maneira facilitada a um auditor humano;
- Desenvolvimento de um método para a avaliação do grau de ocorrência de atrasos nos pagamentos de empenhos para diferentes municípios; e
- Desenvolvimento de um método para a avaliação do grau de desrespeito à ordem cronológica no pagamento de empenhos para diferentes municípios.

As próximas seções estão organizadas da seguinte forma. A Seção 2 apresenta trabalhos relacionados ao trabalho em questão. A Seção 3 apresenta detalhes sobre os métodos desenvolvidos nesse trabalho para auxílio a auditores humanos na condução de auditorias de contas públicas municipais. Por fim, a Seção 4 apresenta as conclusões, propõe trabalhos futuros e apresenta ameaças à validade do trabalho.

2. Trabalhos Relacionados

Existe uma relação muito forte entre a transparência (i.e., publicidade dos dados de maneira clara) de informações e a corrupção. No Brasil, a transparência nas contas públicas é obrigatória e regida pela Lei de Acesso à Informação (LEI Nº 12.527, DE 18 DE NOVEMBRO DE 2011 - LAI). Em [de Mattos Guimarães 2019], Guimarães apresenta um estudo sobre a difusão de políticas públicas assim também como aborda o caso da adoção de mecanismos que possibilitem a prática da transparência nas esferas administrativas públicas. Já em [Filho 2019], o autor realizou uma pesquisa qualitativa para descobrir qual forma de transparência a LAI visa produzir, a fim de compreender os impactos que a lei produz no regime democrático brasileiro.

Dada a alta disponibilidade pública de dados, a busca por sinais de corrupção se torna mais fácil e vira um importante objeto de estudo. Em [Vasconcelos et al. 2021] os autores cruzaram informações de 8 fontes de dados diferentes para gerar uma base de dados capaz de prover informações para criação de um modelo capaz de identificar eventos de corrupção. Um problema nessa abordagem foi o alto grau de desbalanceamento na base de dados gerada (i.e., muitos exemplos de dados da classe lícito em detrimento da classe ilícito). Na mesma linha, [Carvalho et al. 2013] Carvalho et. al se utilizaram de ontologias probabilísticas para projetar e testar um modelo que realiza fusão de informações para detectar possíveis fraudes em compras envolvendo dinheiro público federal no Brasil. Uma ideia de uso da ferramenta gerada seria a realização de uma triagem dos casos mais prováveis de delitos de forma a direcionar melhor o trabalho do auditor humano.

A auditoria de dados é um processo que requer muito conhecimento específico dos dados e, dentre outras, por essa razão, se encontra ainda distante de uma solução

através de técnicas de aprendizagem de máquina. Em [Marghescu et al. 2010], os autores criaram um modelo para detecção de empresas com débitos fiscais. A ideia, novamente, é evidenciar casos mais passíveis de uma auditoria humana. Kiefer e Pesch [Kiefer and Pesch 2021], por sua vez, propõem uma arquitetura de aprendizagem de máquina capaz de ajudar auditores financeiros a detectar de forma transparente casos anômalos na ausência dos rótulos corretos das classes (i.e., de maneira não supervisionada). Além disso, eles aplicam também técnicas de pós-processamento no modelo de forma a tornar a decisão mais transparente. Para isso eles usam técnicas de *Explainable AI* [Islam et al. 2021].

3. Métodos Propostos

Este trabalho foi desenvolvido analisando-se as contas públicas dos municípios do estado de Pernambuco nos anos de 2019 e 2020. Esses dados podem ser obtidos através da API de acesso aos dados públicos desenvolvida pelo Tribunal de Contas do Estado de Pernambuco (TCE-PE)². O estado de Pernambuco conta com 184 municípios mais a ilha de Fernando de Noronha. Esse trabalho contempla apenas 183 municípios, deixando de fora a capital Recife e a ilha de Fernando de Noronha.

Sistemas de informação para a gestão de contas públicas naturalmente respeitam as regras da contabilidade pública. É de se esperar que essa classe de sistemas não permita o registro de um pagamento sem sua devida conexão com uma liquidação. Infelizmente, no contexto do estado de Pernambuco, os dados fornecidos pelos municípios ao portal do TCE-PE não fornecem essa ligação entre um pagamento e sua respectiva liquidação. Dessa forma, é impossível se analisar o nível de atraso nos pagamentos de empenhos assim como o nível de desrespeito à ordem cronológica dos pagamentos. Nessa seção, (i) relatamos o tratamento dos dados para conectar pagamentos e liquidações, (ii) apresentamos um método para ordenar casos com maiores chances de serem pagamentos de folha salarial através de empenhos, (iii) apresentamos um método para cálculo de índices de atrasos de pagamentos de modo a destacar municípios que, relativamente, mais violam esse critério, e (iv) apresentamos um método para cálculo de índices de desrespeito à ordem cronológica dos pagamentos destacando municípios que, relativamente, mais violam essa regra.

Via de regra, ocorrida uma liquidação, a entidade contratante tem um prazo definido para o pagamento do respectivo valor. O atraso nesse pagamento (i.e., a distância entre as datas de liquidação e pagamento) pode acontecer, dentre outros motivos, pela cobrança de propina por agentes públicos. Outra forma de analisar o problema dos atrasos de pagamentos é pela validação da ordem cronológica dos pagamentos em relação às suas respectivas liquidações. Sejam as liquidações l_1 e l_2 efetuadas nos tempos t_1 e t_2 , respectivamente, tal que $t_1 < t_2$. Um cenário onde a liquidação l_2 seja paga antes da liquidação l_1 pode indicar uma vantagem indevida do fornecedor responsável pela l_2 em detrimento do fornecedor responsável pela l_1 . Dessa forma, o desrespeito à ordem cronológica de pagamentos é uma falta grave, que pode ser decorrente de outros ilícitos.

3.1. Extract Load Transform - ELT

Para calcular o atraso de pagamento de empenhos é preciso inicialmente vincular cada pagamento à sua respectiva liquidação. Para essa finalidade, foi criado um *script* de pro-

²<https://www.tce.pe.gov.br/internet/index.php/dados-abertos/definicoes-da-api>

cessamento de dados que, para um conjunto de liquidações L e outro de pagamentos P , relativos a um mesmo empenho, trata os dados de forma a remover informações desnecessárias e que possam prejudicar esta vinculação.

O primeiro desafio consiste na existência de pagamentos do tipo ‘retenção’ que se trata de taxas advindas de um pagamento específico. Uma ou mais retenções existem para um pagamento, porém não para uma liquidação. Então para cada pagamento é verificado se existem retenções e essas são somadas ao valor do pagamento de forma que essa soma é então comparada com o valor de uma dada liquidação. Esse mesmo processamento é feito para pagamentos do tipo ‘restos a pagar’ que consistem em pagamentos referentes a liquidações ocorridas no ano anterior.

Outro desafio consiste no tratamento de estornos de pagamentos e liquidações. Caso haja a ocorrência de estornos de pagamentos, inicialmente o valor do estorno é comparado com pagamentos anteriores de forma que o pagamento mais antigo com mesmo valor do estorno seja excluído. Caso não haja pagamentos no mesmo valor do estorno, o valor do estorno é subtraído do pagamento mais antigo com o valor maior que o estorno. Esse pré-processamento é realizado tanto para pagamentos quanto para liquidações. Essa é uma análise bastante comum pois o erro na inserção de dados em sistemas contábeis por agentes públicos ocorre com frequência.

Uma vez realizados os pré-processamentos descritos acima, os dados de pagamentos e liquidações estão prontos para serem vinculados. Lembrando que um pagamento só pode ser vinculado a uma liquidação ocorrida em uma data igual ou anterior à data do pagamento.

O Algoritmo 1 se aplica a todos os casos em que os valores de pagamentos são menores que os valores das liquidações e o valor da liquidação é múltiplo dos valores dos seus respectivos pagamentos (e.g., liquidação = 500 e pagamentos = {230, 270}). Para o Algoritmo 1 é dado como entrada a lista de valores de pagamentos até então sem liquidação associada ($pags$), o índice atual do pagamento a ser analisado (idx), o valor da liquidação atual ($valLiq$), uma lista contendo os índices dos pagamentos cuja soma resulta no valor da liquidação em questão ($idxsPagsFinal$) e os índices de todos os pagamentos a serem investigados como possíveis pagamentos da liquidação em questão ($idxsPags$).

A ideia do Algoritmo 1 é que todas as possibilidades de soma de pagamentos sejam testadas de modo que se uma combinação de pagamentos coincidir com o valor da liquidação (linha 10), os valores desses pagamentos e seus respectivos índices serão retornados possibilitando assim a vinculação desses pagamentos à liquidação em questão. Como exemplo, sendo $pags = \{23, 42, 67, 12, 92, 43\}$ e $valLiq = 152$, ao final da execução do Algoritmo 1 para esses parâmetros os valores das listas solução e $pagsIdxs$ seriam respectivamente $\{42,67,43\}$ e $\{1,2,5\}$.

O Algoritmo 1 resolve a maioria dos casos de vinculação de pagamentos às suas respectivas liquidações, porém, nem sempre os valores somados de pagamentos combinam exatamente com o valor de uma liquidação (condição necessária para execução do algoritmo). Como ilustração de caso onde o Algoritmo 1 não se aplica, vamos usar as sequências de liquidações $\{(15-01; R\$ 1000,00), (15-02; R\$ 1000,00), (15-03; R\$ 1000,00)\}$ e pagamentos $\{(20-02; R\$ 1500,00), (20-03; R\$ 300,00), (20-04; R\$ 1200,00)\}$, como apresentado na Figura 1. Nesse caso, o fornecedor tinha um saldo a receber de R\$

Algorithm 1 Soma de Subconjuntos (SubsetSum)

```
1: solucao = {}
2: pagsIdxs = {}
3: sair = False
4: procedure SUBSETSUM(pags, idx, valLiq, pagsFinal, idxsPagsFinal, idxsPags)
5:   res = copia(pagsFinal)
6:   ind = copia(idxsPagsFinal)
7:   if sair then
8:     return False
9:   end if
10:  if valLiq == 0 then
11:    solucao = pagsFinal
12:    pagsIdxs = idxsPagsFinal
13:    sair = True
14:    return True
15:  end if
16:  if (idx >= tamanho(pags)) ou (valLiq < 0) then
17:    return False
18:  end if
19:  res = res + pags[idx]
20:  ind = ind + idxsPags[idx]
21:  inc = subsetsum(pags, idx + 1, valLiq - pags[idx], res, ind, idxsPags)
22:  exc = subsetsum(pags, idx + 1, valLiq, pagsFinal, idxsPagsFinal, idxsPags)
23:  return inc or exc
24: end procedure
```

2000,00 na data 20-02, porém, R\$ 1500,00 foi pago quitando totalmente a liquidação 1 e parcialmente a liquidação 2. Dessa forma o atraso do pagamento 1 foi de 46 dias. O pagamento 2 (R\$ 300,00) somado a R\$ 500,00 que sobraram do pagamento 1 se referem à liquidação 2, dessa forma o pagamento 2 tem atraso de 33 dias (considerando um fevereiro de 28 dias). Por sua vez, o atraso do pagamento 3 será de 64 dias visto que ele quita duas liquidações, porém, temos que escolher uma para computar o atraso e nesse caso opta-se pela mais distante a ser liquidada (i.e., 15-02). Essa regra se aplica em um baixo percentual de pagamentos de forma que a imprecisão inerente a esse método gera um baixo impacto na solução final.

3.2. Caso de Estudo 1 - Extração de Pagamentos Regulares

A ocorrência de pagamentos regulares (valores e datas semelhantes ao longo do tempo) através de empenhos não é uma prática que configura um ato ilícito, porém, há casos em que entidades públicas podem se utilizar da contratação de empenhos para o pagamento de funcionários, ou seja, utilizar empenhos de forma auxiliar à folha de pagamento. Esse é um ato ilícito que requer um esforço relativamente alto para ser identificado pela auditoria de dados.

A maneira mais comum de se conduzir essa investigação é através da análise individual de empenhos e seus respectivos valores e datas de pagamentos. Ou seja, o analista

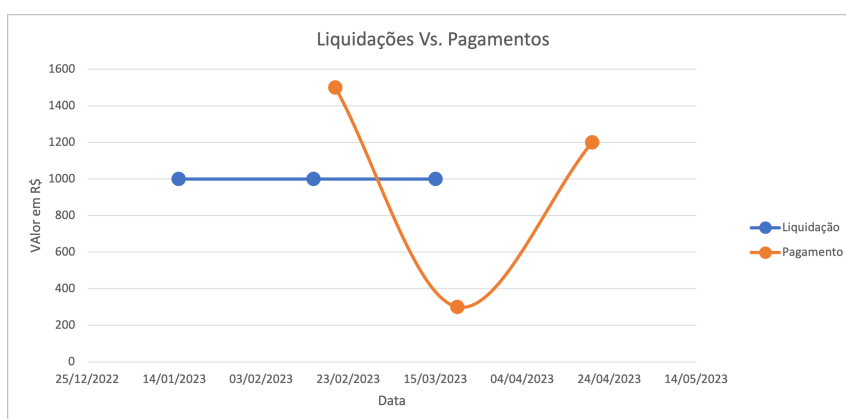


Figura 1. Exemplo de valores e datas de pagamentos e liquidações cujos valores de pagamentos não coincidem com os valores de liquidações (Fonte: autoria própria).

deve buscar pela existência de regularidade nas datas de pagamentos (e.g., pagamentos recorrentes aproximadamente no dia 05 de cada mês) assim como a regularidade no valor. Essa é uma busca muitas vezes realizada através da criação de códigos SQL (*Structured Query Language*) a serem aplicados diretamente no banco de dados. Dentre os problemas dessa abordagem está a impossibilidade de se ordenar os casos mais prováveis desse delito assim como o grande número de empenhos a serem analisados. Dessa forma, para um auditor humano é, virtualmente, impossível se analisar todos os empenhos de um município no período de um ano, por exemplo.

O presente trabalho introduz um método automatizado capaz de organizar empenhos como potenciais pagamentos de salários (folha de pagamento) de forma a apresentar os casos mais prováveis de ocorrência desse ilícito de maneira ordenada ao auditor humano.

O índice que possibilita a ordenação é formado por dois componentes. O primeiro componente (c_1) consiste na média da lista de diferenças entre datas consecutivas de pagamentos menos 30 dias. Dessa forma, quanto maior for a recorrência no valor das datas de pagamentos mais próxima de zero será a média do vetor \vec{D} . A equação 1 ilustra como é computado o vetor \vec{D} . Na equação 1 a função *abs* retorna o valor absoluto de um dado argumento.

$$\vec{D} = \{abs((dt_1 - dt_0) - 30)\}, abs((dt_2 - dt_1) - 30), \dots, abs((dt_n - dt_{n-1}) - 30)\} \quad (1)$$

O segundo componente do índice (c_2) consiste no desvio padrão dos valores de pagamentos normalizados entre 0 e 1. Intuitivamente, quanto maior o desvio padrão no valor de pagamento menor é a regularidade (repetição) desses valores (i.e., valores repetidos são mais facilmente associados a pagamento de folha). Dessa forma, o índice então consiste na soma $i = c_1 + c_2$. A Figura 2 apresenta as informações de um empenho com um valor baixo para o índice i (i.e., valores baixos para esse índice podem configurar pagamentos recorrentes/regulares). Nesse caso as datas notavelmente tendem a ser em um período aparentemente mensal enquanto o valor permanece o mesmo. Adicionalmente, a

descrição do empenho é fornecida ao auditor para que ele possa decidir pelo aprofundamento ou não na investigação.

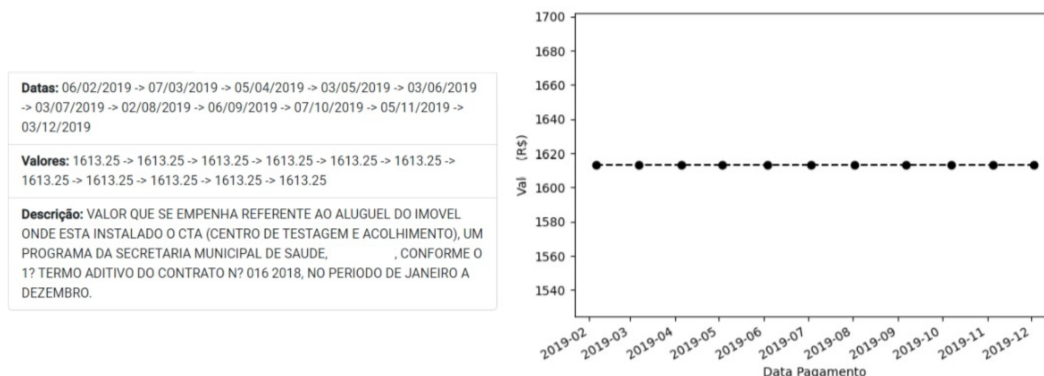


Figura 2. Exemplo de empenho com índice alto para início de pagamentos regulares (Fonte: autoria própria).

3.3. Caso de Estudo 2 - Análise do nível de Atrasos de Pagamentos

Como mencionado anteriormente, o mau uso da máquina pública pode acontecer de diversas formas. Uma forma muito comum é através da cobrança de vantagens indevidas por agentes públicos com a finalidade de acelerar processos ou pagamentos. No caso de pagamentos de empenhos, a análise do período entre a liquidação (registro de entrega do produto ou serviço pelo fornecedor) e o pagamento dessa liquidação podem indicar a necessidade de uma maior atenção a uma ou outra entidade pública.

Para que a comparação entre municípios de diferentes portes econômicos seja possível é necessário a criação de um método para comparação não sensível ao volume da movimentação financeira do município. Adicionalmente, o cálculo do nível de atraso de pagamentos não deve ser realizado considerando-se a granularidade de município, ao invés disso, deve ser considerada a granularidade de unidade gestora (i.e., secretaria de saúde, secretaria de educação, etc.) e fonte de recurso (i.e., recursos ordinários, fundeb, etc.). Então, dada essa granularidade, e dado também que já está disponível a data de liquidação para cada pagamento (cujo cálculo foi apresentado na Seção 3.1), podemos calcular o total de dias de atraso (i.e., número de dias entre a liquidação e o pagamento) para cada pagamento.

Dado uma unidade gestora e uma fonte de recursos específicas, o índice de atraso de pagamentos (δ) para essa combinação é gerado de acordo com a equação 2. Nessa equação, n é o total de pagamentos com atraso maior que um máximo de dias tolerados (tol) e t é o total de pagamentos realizados por aquela unidade gestora e fonte de recurso para um dado período. O valor 30 é necessário para retornar em meses a diferença entre datas. Ao final, esse cálculo é realizado para todas as combinações de unidades gestoras e fontes de recurso de um município gerando-se um vetor de índices de forma que o índice de atraso de pagamentos do município será a média dos valores desse vetor.

$$\delta = \frac{\sum_{i=1}^n 1 + \left(\frac{(dtPag - dtLiq) - tol}{30} \right)}{t} \quad (2)$$

A Figura 3 ilustra um mapa de calor gerado pela ferramenta. Neste caso, (i) consideramos apenas pagamentos do ano de 2019 com tolerância de 30 dias; (ii) empenhos do tipo licitação e (iii) os fornecedores são empresas. Os municípios com maior índice de atraso se encontram na cor mais avermelhada e esses devem ser avaliados com maior atenção pelo auditor humano pois concentram um maior nível médio de atrasos nos pagamentos.

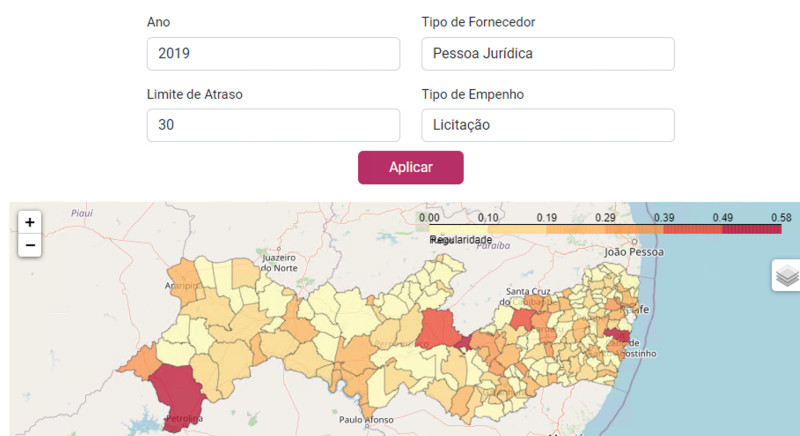


Figura 3. Exemplo de mapa gerado onde os municípios com relativamente mais atrasos de pagamentos se encontram na cor mais avermelhada (Fonte: de autoria própria)

É importante salientar que uma vez que o auditor humano decida investigar uma cidade em particular, a ferramenta gerada é capaz de apresentar informações específicas de cada unidade gestora e fonte de recurso de forma que o auditor terá acesso a dados de cada pagamento atrasado de forma facilitada.

3.4. Caso de Estudo 3 - Análise de Desrespeito à Ordem Cronológica dos Pagamentos

A observação da ordem cronológica para pagamento de empenhos já é uma realidade no governo federal através da Instrução Normativa (SEGES/ME N° 77, DE 4 DE NOVEMBRO DE 2022). Essa IN rege o pagamento de produtos e serviços de forma que a ordem de pagamento respeite uma fila de prioridades. Ou seja, uma liquidação realizada na data d pode apenas ser paga após todas as liquidações realizadas em datas anteriores a d terem sido pagas. Infelizmente, essa regra ainda não é adotada no âmbito de governos estaduais e municipais.

Para a avaliação do nível de desrespeito à ordem cronológica, foi desenvolvida um método que, assim como na Seção 3.3, considera os pagamentos realizados em uma unidade gestora e fonte de recurso específicas. Para cada pagamento p com atraso maior que um período de tolerância tol , é computado o número de pagamentos cujas liquidações ocorreram após a liquidação do pagamento p . Esse valor é então somado de forma a obter um $score$ para a unidade gestora e a fonte de recurso. Esse $score$ varia de zero a valores bastante altos. Por essa razão, o índice de desrespeito é dado então por $\log_{10}(score)$ o que possibilita uma comparação mais justa entre municípios maiores (com bastantes empenhos e pagamentos) e municípios menores - dado que a partir de um certo grau de desrespeito à fila de pagamentos a variação no índice se torna menos acentuada. O índice

de cada município será o índice médio de todas as combinações de suas unidades gestoras e fontes de recurso.

Assim como na Seção 3.3, um mapa será gerado de forma a mostrar os municípios com maior índice de desrespeito à ordem cronológica para pagamentos de empenhos. Selecionando-se um município específico é possível ter informações detalhadas sobre suas unidades gestoras e seus respectivos pagamentos como mostra a Figura 4. Na Figura 4 é possível ver que para a cidade de Abreu e Lima, no ano de 2019, a Secretaria de Educação com recursos ordinários beneficiou 308 e 293 pagamentos em detrimento do fornecedor responsável pelo empenho 452 desta unidade gestora.

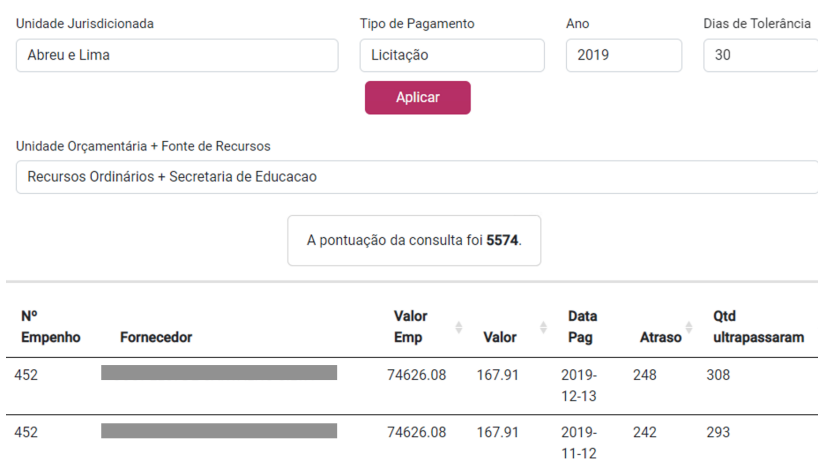


Figura 4. Exemplo de análise gerada para um município de forma explicitar pagamentos prejudicados na fila de pagamentos (Fonte: de autoria própria).

4. Conclusão

Como principal **contribuição**, este artigo apresentou os principais aspectos para a criação de uma ferramenta capaz de auxiliar profissionais de auditoria em três diferentes tipos de auditoria: (i) identificação de pagamentos de salário através da contratação de empenhos; (ii) identificação de municípios com maior nível de atrasos de pagamentos de empenhos; e (iii) identificação de municípios com maior nível de desrespeito à ordem cronológica de pagamentos de empenhos. Outra contribuição desse trabalho foi descrever o processo de extração e transformação de dados de empenhos a partir de uma API de acesso aos dados do TCE-PE - o que permitiu as três análises mencionadas acima.

A proposta é fundamental para uma maior cobertura na fiscalização de entidades públicas de forma a coibir delitos envolvendo o erário. Convidamos profissionais de pesquisa a também explorar os desafios do cenário de auditoria de dados, que demandam conhecimentos específicos tanto da área fim (e.g. controle externo) quanto da área de ciência da computação. Contribuições futuras podem aumentar a execução em larga escala dessa atividade bem como reduzir a sobrecarga dos poucos profissionais que atuam neste contexto no governo, permitindo que explorem qualificações adicionais, como de interpretação dos resultados de uma fiscalização à luz de conceitos de políticas públicas.

Como **trabalhos futuros**, temos: (i) extensão da pesquisa atual para outras unidades federativas de modo a aumentar a granularidade de comparação dos aspectos investigados de municípios para estados da federação; (ii) identificação de fornecedores

beneficiados/prejudicados de forma mais recorrente; (iii) comparação dos aspectos investigados para diferentes gestões administrativas de forma a identificar gestores potencialmente inidôneos. Não obstante, as análises automatizadas apresentadas nesse trabalho precisam também ser validadas por especialistas humanos.

4.1. Ameaças à Validade

O presente estudo não envolveu experimentos dos quais derivam-se hipóteses que necessitem ser confirmadas através de testes estatísticos, por exemplo. Dessa forma, apenas os tipos de validade interna e externa se aplicam ao objeto de estudo.

- **Interna** - O presente trabalho foi desenvolvido a partir de dados disponíveis na API de dados abertos do TCE-PE. Dessa forma, a correteza dos métodos e resultados apresentados depende da precisão com que esses dados são disponibilizados nessa API.
- **Externa** - Teoricamente, não há fatores que impeçam a adoção dos métodos apresentados nesse trabalho por órgãos fiscalizadores da mesma categoria, ou seja, tribunais de contas de outros estados. Considerando que a dinâmica de contas públicas municipais é semelhante para todos os municípios da federação, os estudos conduzidos no presente trabalho se aplicam a todos os estados da federação.

Referências

- Carvalho, R. N., Matsumoto, S., Laskey, K. B., Costa, P. C. G., Ladeira, M., and Santos, L. L. (2013). Probabilistic ontology and knowledge fusion for procurement fraud detection in brazil. In *Uncertainty Reasoning for the Semantic Web II*, pages 19–40, Berlin, Heidelberg. Springer Berlin Heidelberg.
- de Araujo, R. F. (2014). A percepção de diferentes tipos de corrupção na ótica dos profissionais da contabilidade.
- de Mattos Guimarães, E. (2019). *A difusão da política de transparência pública no Brasil: a atuação da Controladoria Geral da União no estado do Espírito Santo*. PhD thesis.
- de Oliveira Rego, H. (2021). *Transparência e accountability em portais brasileiros*. PhD thesis.
- Filho, M. C. C. (2019). *A Construção da Transparência Pública no Brasil: Análise da Elaboração e Implementação da Lei de Acesso à Informação no Executivo Federal (2003-2019)*. PhD thesis.
- Islam, S. R., Eberle, W., Ghafoor, S. K., and Ahmed, M. (2021). Explainable artificial intelligence approaches: A survey.
- Kiefer, S. and Pesch, G. (2021). Unsupervised anomaly detection for financial auditing with model-agnostic explanations. In *KI 2021: Advances in Artificial Intelligence: 44th German Conference on AI, Virtual Event, September 27 – October 1, 2021, Proceedings*, page 291–308. Springer-Verlag.
- Marghescu, D., Kallio, M., and Back, B. (2010). Using financial ratios to select companies for tax auditing: A preliminary study. In *World Summit on the Knowledge Society*.
- Vasconcelos, M. O., Chaim, R. M., and Cavique, L. (2021). Imbalanced learning in assessing the risk of corruption in public administration. In *Progress in Artificial Intelligence*, pages 510–523. Springer International Publishing.