

Web Analysis: Análise de Dados de Transparência Municipal

Web Analysis: Analysis of Municipal Transparency Data

Gustavo Henrique de Moraes Almeida
Instituto de Ciências Exatas e Tecnológicas
Universidade Federal de Viçosa
Florestal, Brasil
gustavo.h.almeida@ufv.br

Fabrcio Aguiar Silva
Instituto de Ciências Exatas e Tecnológicas
Universidade Federal de Viçosa
Florestal, Brasil
fabricio.asilva@ufv.br

RESUMO

Será descrito neste artigo um projeto para análise dos dados disponibilizados em um Portal da Transparência, para fins de fiscalização dos gastos públicos por intermédio do cidadão.

Utilizando conceitos de Data Science e outros recursos computacionais, o sistema desenvolvido é capaz de capturar os dados do portal da transparência, formatá-los de maneira mais interessante, gerar gráficos, relatórios e estrutura similares, que auxiliam na identificação de possíveis problemas relacionados aos gastos públicos.

Assim, o projeto adquire características sociais, ao auxiliar o cidadão a reivindicar seus direitos. Também, aborda mecanismos para fácil entendimento dos dados mencionados, sendo um diferencial no cenário atual. Dessa forma, é atingido o objetivo do real entendimento dos dados disponibilizados por parte de entidades públicas municipais.

Palavras-Chave

Análise de Dados. Portal da Transparência. Município. Web Crawler. Selenium.

ABSTRACT

This article will describe a project to analyze data from the Transparency Portal, for the purposes of public expenditure control through the citizen.

Using Data Science concepts and other computational resources, the developed system is capable of capturing transparency portal data, formatting it in a more interesting way, generating graphs, reports and similar structures, which helps to identify possible problems related to public expenses.

Thus, the project acquires social characteristics, helping the citizen to claim their rights. Also, it addresses mechanisms for easy understanding of the mentioned data, being a differential in the current scenario. In this way, the objective of the real understanding of the data made available by municipal public entities is reached.

Keywords

Data Analysis. Transparency Portal. City. Web Crawler. Selenium.

1. INTRODUÇÃO

A Lei nº 12.527/2011 - Lei de Acesso à Informação (LAI) (1) define normas referentes aos gastos públicos da União, dos Estados, dos Municípios e do Distrito Federal. As informações devem ser disponibilizadas em meios eletrônicos (Portais da Transparência), de forma detalhada para simples conferência do tipo, natureza e valor total, destinadas a toda e qualquer pessoa da população que deseje acesso. Cada unidade política citada é responsável pela divulgação e validação de seus dados, sendo necessário muitas vezes, principalmente nos municípios, a contratação de empresas de tecnologia da informação para desenvolver um sistema que se adéque às necessidades.

Nesses sistemas, foram encontrados diversos dados com formatação própria de cada entidade relacionada, pois cada uma define suas metodologias de disposição dos gastos. Assim, muitas divergências foram encontradas entre os mais diferentes tipos de portais, principalmente nos municipais, já que cada um é responsável pela contratação de empresas para desenvolver o sistema, além de alimentar o banco de dados respectivo.

Cada cidadão poderia, por conta própria, analisar essas informações disponíveis, e relatar possíveis não conformidades nos gastos públicos. No entanto, existe uma grande quantidade de dados dispostos em classificações, muitas vezes confusas, que deixa a análise complexa. Além disso, seria necessário a verificação individual de cada gasto, realizando algum cálculo sobre os valores quando necessário. Depois de realizar todo esse esforço, seria necessário registrar tudo o que fora encontrado e, só assim, verificar possíveis inconformidades.

O sistema desenvolvido descrito neste artigo auxilia em tal processo, além de outros recursos. O sistema então auxilia toda a população de um município que o utiliza, deixando os gastos públicos mais fáceis de serem compreendidos. Assim, toda população seria beneficiada com o sistema, pois poderiam reivindicar seus direitos, além de sempre cobrar melhorias no investimento público junto à prefeitura. Dessa forma, a própria prefeitura se beneficiaria, pois possuiria um sistema claro de disponibilização de seus gastos, tornando-a bem conceituada no serviço público. E também, dentre outros benefícios, otimizar o uso de seus recursos através do auxílio de sua própria população, que são os principais envolvidos.

Este artigo está organizado da seguinte forma: na Seção 2

será tratado os trabalhos relacionados, no qual foi utilizado embasamento teórico semelhante, na Seção 3 será tratado o sistema desenvolvido, uma explicação sobre sua arquitetura, na Seção 4 será abordado o caso de estudo prático (Portal de Pará de Minas - MG (2)), na Seção 5 a conclusão e trabalhos futuros, e por fim, na Seção 6 as referências utilizadas.

2. TRABALHOS RELACIONADOS

Recentemente tem ocorrido um crescente interesse por parte dos cidadãos quanto aos gastos públicos, frente a quantidade de escândalos referentes a desvios e outras formas ilícitas de enriquecimento próprio praticado por alguns governantes. Mesmo sabendo disso, existem poucos trabalhos relacionados no Brasil, que utilizam recursos computacionais para análise automática e sistemática dos dados.

No Brasil, temos o OPEN KNOWLEDGE BRASIL (3), que visa avaliar a estruturação dos dados abertos de cada cidade ou país, e suas características, como o tipo do dado, formato, entre outros. Através dos dados disponibilizados pelo governo, geralmente em forma de planilhas, são realizadas diversas manipulações desses dados até atingir o objetivo esperado.

Uma iniciativa interessante a se destacar é a chamada “Cuidando do Meu Bairro”, que, utilizando os dados mencionados anteriormente, juntamente com um serviço de geolocalização, posiciona cada informação em seu devido local de atuação em São Paulo, seja em uma rua, praça, e assim por diante. Com isso, cada pessoa pode verificar determinados serviços em tempo real, e reivindicar possíveis atrasos ou superfaturamentos. Fica evidente a sensibilidade do projeto frente ao nível de detalhamento necessário para um bom entendimento dos dados.

Podemos citar como o mais conhecido trabalho brasileiro a Operação Serenata de Amor (4) que, por meio dos reembolsos efetuados por parlamentares (verbas destinadas a transporte, alimentação, etc), constrói uma grande base de dados, que então é analisada utilizando conceitos de Ciência de Dados. Os dados são disponibilizados pela Câmara dos Deputados, Receita Federal, entre outros, lembrando que o acesso a essas informações são garantidas por lei, salva proteção jurídica. O trabalho é realizado por um grupo de pessoas espalhados pelo Brasil e em outros países, com acesso remoto ao código. A operação disponibiliza seu código fonte para quem quiser entender ou atribuir seus conhecimentos e, além dessa forma de ajuda, aceitam auxílio financeiro, o que é chamado de “campanha de financiamento coletivo”.

Em Fevereiro de 2018, foi anunciada a união da equipe da Serenata de Amor e a OPEN KNOWLEDGE BRASIL, pela própria OKBR, que trará maior abrangência dos dados, maior flexibilidade da análise, além de torná-los sempre atuais. A utilização de inteligência artificial possibilita retirar informações antes não levantadas.

Diante disso, vemos que esse projeto existente é bem robusto e organizado, mas abrange somente entes federativos da União, do Distrito Federal e, algumas vezes, dos Estados. Devido à falta de iniciativa para gastos menores, como os existentes nos municípios, esse trabalho visa elucidar uma forma prática e usual de captura e análise dos dados mencionados, com uma abordagem muito semelhante à já existente (como linguagem de programação e técnicas de Ciência de Dados).

3. O SISTEMA

3.1 Arquitetura

Na Figura 1 podemos analisar o funcionamento geral do sistema desenvolvido. Cada elemento representado é responsável por um componente individual, que pode ou não utilizar recursos de outros.

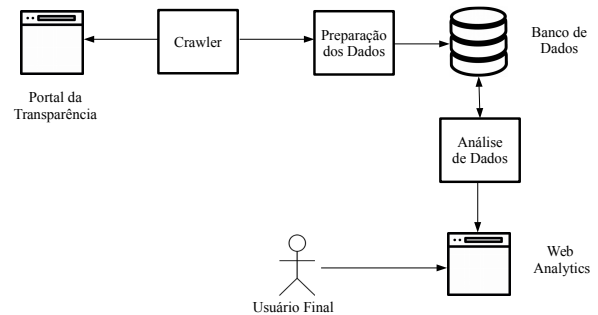


Figura 1: Arquitetura do sistema

3.1.1 Portal da Transparência

Contém todo gasto público corrente até a data de acesso ao sistema. Os gastos são dispostos, geralmente, em uma ou mais tabelas, com suas especificações conforme a LAI: Classificação mais abrangente (saúde, educação, etc), Classificação Intermediária (Hospital, Gabinete do Prefeito, etc) e Elementos de Despesas (com valores, fornecedores ou favorecidos, etc).

O Portal possui diversas nomenclaturas próprias, tornando-o confuso para um usuário leigo. Podemos citar, por exemplo: Valor Empenhado (designa o valor acordado no contrato), Valor Liquidado (valor referente a quando o serviço for executado ou o bem for recebido), Valor Pago (quando o valor é efetivamente pago), dentre outras existentes, que serão detalhadas posteriormente dentro de cada componente respectivo.

3.1.2 Web Crawler

É um programa de computador que navega de forma sistemática em páginas da web, podendo capturar dados relevantes para um objetivo e, em seguida, manipular da melhor forma possível. É utilizado para manutenção de sites, validação de HTML, em motores de busca para otimizar indexação de páginas, entre outras utilidades. Para o sistema desenvolvido, é responsável pela varredura e captura dos dados, que em seguida são formatados e persistidos. De forma mais abstrata, podemos subdividir a captura dos dados da forma descrita na Figura 2.

Para esse componente, foi utilizada como base a ferramenta Selenium (5), que possui um conjunto de recursos possibilitando o desenvolvimento de um crawler. Possui um ambiente de desenvolvimento próprio e extensões para outras IDEs e linguagens de programação. Utiliza navegadores externos para realizar suas tarefas, construindo uma “ponte” entre o código desenvolvido e um navegador web externo ao programa. Bastante utilizado no Teste Funcional de softwares, através da

criação de rotinas voltadas às funcionalidades de um sistema. Utilizando seus recursos, é possível clicar em botões (e assim podendo executar códigos em JavaScript existentes nas páginas), clicar em links, retornar páginas, pesquisar por componentes web, enviar chaves textuais para pesquisas, capturar textos, dentre outras funcionalidades.

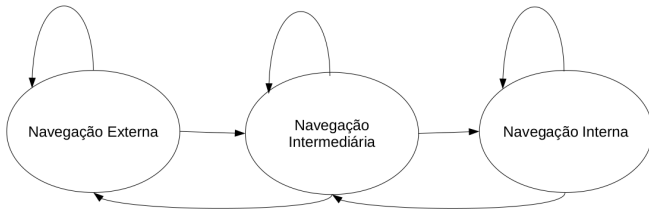


Figura 2: Representação por estados

1. Navegação Externa: processo em que o crawler navega por diversos tipos de gastos mais concentrados, com uma ampla subdivisão. Por exemplo: gasto com saúde e educação.
2. Navegação Intermediária: processo em que o crawler navega onde exista dados mais específicos que o anterior, podendo ser por exemplo: gasto com hospitais e pagamento a servidores da educação.
3. Navegação Interna: processo em que o crawler navega pelos dados com nível de detalhamento maior possível, que é o foco principal do trabalho. Contém nome do favorecido ou fornecedor em questão, valores pagos, órgão responsável, etc. O crawler interrompe sua navegação momentaneamente para capturar e formatar os dados, para posteriormente armazená-los e retornar à navegação.

Primeiramente o Web Crawler inicia a Navegação Externa, onde pode existir uma ou mais opções. Após clicar na primeira opção disponível, serão listadas despesas mais específicas (geralmente em maior quantidade que a etapa anterior), referente à Navegação Intermediária. Ao clicar em um deles, inicia-se a Navegação Interna, onde consta os dados mais específicos possíveis. Nesse momento, utilizando ferramentas do Selenium e da própria linguagem de programação, são capturados todos os dados lá disponíveis, que são temporariamente armazenados em memória primária e em seguida, ao capturar todas despesas disponíveis nessa etapa, é armazenado em memória secundária.

Após isso ocorrer, temos o primeiro gasto público armazenado, referente a um favorecido ou fornecedor específico. O Web Crawler retorna para a etapa anterior (Navegação Intermediária) e busca por mais opções disponíveis, caso exista, essa opção é selecionada e o processo é refeito.

Ao verificar que não existem mais opções na Navegação Intermediária, o Web Crawler retorna à Navegação Externa para selecionar a opção seguinte. Com o término de opções na Navegação Externa o programa encerra.

3.1.3 Preparação dos Dados

O Portal da Transparência possui diversas informações adicionais que devem ser ignoradas ou substituídas. Especificações dos atributos dos Elementos de Despesas (“Nome: ”,

“Órgão: ”, “Subfunção: ”, e todos os demais atributos disponíveis) não são necessários, então são eliminados dos dados capturados, restando somente o atributo em questão. Essas partes removidas são utilizadas como nomes dos atributos no banco de dados, que será melhor detalhado posteriormente, garantindo maior confiabilidade nos dados armazenados.

Quanto aos pagamentos, existem os Valores Liquidados e Valores Pagos. Valores Liquidados são os valores que devem ser pagos ao respectivo favorecido ou fornecedor, mediante a realização do que está descrito no contrato. Valores Pagos referem-se aos valores que, depois de liquidados, foram efetivamente pagos pela prefeitura. Sendo assim, foram desconsiderados os dados referentes aos valores liquidados e capturados somente os valores pagos, que são o principal interesse.

Outro ponto importante é utilizar recursos da própria linguagem para formatação de valores para o padrão americano (relacionado ao uso de ponto e vírgula), além da formatação de datas (utilizando o padrão ano/mês/dia). Dessa forma, conseguimos otimizar a utilização do banco de dados, deixando-o flexível para qualquer outra prefeitura desejada.

3.1.4 Banco de Dados

Se tratando de persistência dos dados, a Figura 3 representa uma modelagem simples capaz de armazenar a maioria dos tipo de despesa pública municipal. O modelo é composto por três tabelas que representam as principais entidades envolvidas:

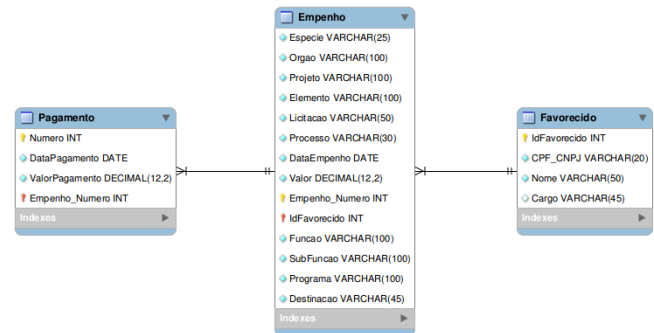


Figura 3: Modelo Relacional do Banco de Dados utilizado.

1. Favorecido: pessoa física ou jurídica que fornece algum tipo de bem ou serviço, com quem a prefeitura possui um débito. Temos o nome, CPF parcial e algumas vezes o cargo, se tratando de pessoas físicas, e nome fantasia e CNPJ caso pessoa jurídica.
2. Empenho: é o primeiro estágio de uma despesa pública. Refere-se a um compromisso assumido entre a prefeitura e um favorecido ou fornecedor. Representa a oficialização de um contrato vigente, garantindo um débito que deve ser pago, mediante o cumprimento de um acordo válido. Cada empenho possui um identificador único que o representa (conforme a LAI), e existem alguns tipos importantes de empenhos:
 - Espécie Ordinária: tipo de empenho que representa despesas de valor fixo, previamente especificado em comum acordo entre as partes interessadas. O pagamento deve ocorrer de uma só vez.

- Espécie Estimativa: tipo de empenho que representa despesas cujo montante não pode ser determinado previamente, como por exemplo: gastos de energia elétrica, água, etc.
 - Espécie Anulação: representa quebra de contrato entre as partes interessadas. Pode ocorrer anulação parcial ou total de um Empenho da Espécie Ordinária.
 - Espécie Complementar: representa a extensão ou assinatura de um novo contrato, embasado em um Empenho Ordinário já existente.
 - Espécie Global: semelhante à Espécie Ordinária, mas seu valor decorre de parcelamentos contratuais, como alugueis.
 - Não Informado: não informado no momento de inserir os dados no sistema.
3. Pagamento: refere-se ao pagamento efetivado por parte da prefeitura (oficialização de um gasto público). Pode ocorrer um pagamento integral ou em parcelas que devem totalizar o valor acordado, estes devem conter um identificador único que o representa.

3.1.5 Análise de Dados

Este componente é responsável por analisar os dados armazenados e, conseqüentemente, possibilitar um rápido e fácil entendimento do conteúdo abordado. Atualmente, essa etapa se baseia em consultas SQL com um foco específico de análise (como, por exemplo, gasto mensal da prefeitura), que podem ser detalhados na forma gráfica ou na forma de Relatórios Específicos.

O principal foco desse componente é selecionar parte dos dados já preparados e armazenados, os quais possuem características ou atributos em comum. Com base nisso, é possível gerar gráficos e relatórios específicos para o próximo componente.

3.1.6 Web Analytics

Já com os dados armazenados e com todas as informações geradas a partir desses dados, essas informações são disponibilizadas de forma mais intuitiva e de fácil entendimento, trazendo gráficos atrativos, constituídos a partir de agrupamentos relevantes. Existem também, Relatórios Específicos, com dados mais pontuais, dispensando o uso gráfico.

Após a solicitação de captura dos dados e concluído todo processo de geração dos gráficos e relatórios, qualquer pessoa do município alvo, ou mesmo outras pessoas que se interessem pelo assunto, podem ter acesso às informações disponíveis. Então, essas informações devem ficar disponíveis em um site e basta ter o link de acesso para acessá-las.

O principal objetivo desse componente é que qualquer pessoa, vinculada ou não ao projeto, possa analisar os gráficos e os relatórios disponíveis, e assim levantar argumentos referentes ao mau uso dos recursos públicos da prefeitura em questão.

3.1.7 Usuário Final

É dever de todo cidadão fiscalizar o gasto público que envolva a sociedade em que vive e a si próprio, de forma mais direta. Assim, após realizar todo processo de captura, preparação e exibição dos dados, qualquer outra pessoa teria acesso a essas informações.

A principal ideia é que o usuário final, ao desconfiar de alguma inconformidade, entre em contato com os desenvolvedores do Web Analytics, seja por e-mail ou formulários disponíveis. Com esse feedback, é possível verificar a veracidade da observação relatada e, caso seja uma possível inconformidade, redirecionar para órgãos públicos ou pessoas responsáveis.

Em alguns municípios, existem pessoas proativas na fiscalização dos gastos públicos. Sendo assim, todo o trabalho é manual, se tornando complexo descobrir algum tipo de inconformidade. O sistema se torna uma ferramenta importante para esse processo, tanto para pessoas proativas quanto para qualquer outro cidadão que, de certa forma, é afetado por esses gastos.

4. CASO DE ESTUDO

4.1 Portal Da Transparência

O Portal de Transparência utilizado como caso de estudo desse trabalho foi o de Pará de Minas – MG. A cidade possui uma boa estrutura dos dados disponibilizados, além de possuir um bom nível de detalhamento (conformidade com a LAI). Foram capturados os dados disponíveis no ano de 2017 como objetivo inicial do estudo.

Importante ressaltar que foram encontrados alguns erros no portal da transparência, como: a descrição dos gastos públicos possuem erros gramaticais graves (falta de concordância verbal, separação silábica no meio de frases, erros de grafia, entre outros), repetição do número identificador dos pagamentos realizados (que deveria ser único) e tipo de gasto não condizente com o relatório oficial final disponível no mesmo site.

4.2 Estruturação dos Dados

Com base na Figura 2, é possível especificar a estruturação dos dados pelo qual o crawler deve navegar até alcançar os dados mais relevantes.

1. Navegação Externa: primeiramente, existe uma tabela composta por gastos mais concentrados, classificados como: Despesas Correntes (3.0.00.00.00) e Despesas de Capital (4.0.00.00.00).
2. Navegação Intermediária: composta por diversas tabelas contendo gastos mais específicos, que compõem as anteriores. As Despesas Correntes possuem 137 registros diferentes, com características próprias (como Subsídio Do Prefeito, Medicamentos, etc). As Despesas de Capital são compostas por 22 registros (como Equipamentos Para Áudio, Video E Foto, Máquinas E Equipamentos Energéticos, etc).
3. Navegação Interna: Contém os dados mais relevantes para o projeto em questão. Possui informações de órgão referente, número único que identifica o elemento de despesa, respectivo projeto, dentre outras. Nessa etapa, o crawler interrompe a navegação nas páginas e captura os dados disponíveis, para em seguida formatá-los e inserir no banco de dados, nos respectivos componentes.

Como preparação, primeiramente exclui-se os nomes dos atributos (como “órgão: ”, que é capturado juntamente ao atributo), que não é interessante para o processo. Em seguida, é trocado o formato da data e do valor monetário, para facilitar a inserção no banco de dados e torna-lo mais flexível para outras aplicações futuras.

4.3 Gráficos Gerados

Os gráficos gerados servem como base para analisar, por exemplo: onde estão os maiores gastos da prefeitura; onde poderia existir uma otimização nos gastos; constar quais gastos possuem um nome suspeito como “outros” ou “ordinários”, que podem esconder algum gasto indevido; etc.

Para gerar esses gráficos no componente Web Analytics, foi realizado agrupamento de cada atributo relevante para o estudo, por exemplo: Órgão, Função, Subfunção, Programa, Projeto, Elemento Despesa, Destinação, Licitação, Gasto mensal, Cargo, Total de diárias Pagas.

A Figura 4 exemplifica um dos agrupamentos citados acima, e consta na primeira versão do Web Analytics.

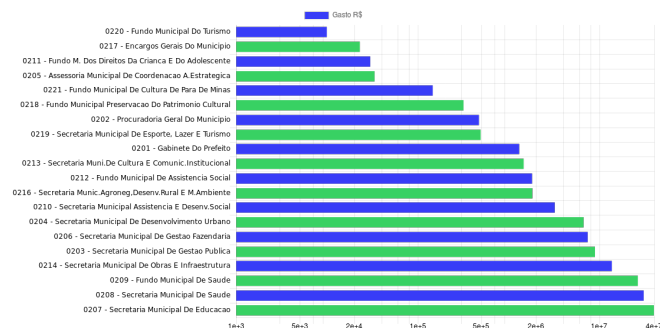


Figura 4: Gráfico de agrupamento por Órgão.

Com base na Figura 4 podemos concluir algumas informações:

1. O maior gasto da prefeitura, no ano de 2017, foi na área da Educação (0207 - Secretaria Municipal de Educação), totalizando aproximadamente R\$40.000.000,00.
2. Segundo maior gasto da prefeitura foi destinado a saúde (0208 - Secretaria Municipal de Saúde), que totaliza aproximadamente, R\$30.800.000,00.
3. Existe o gasto com o gabinete do prefeito (0201 - Gabinete do Prefeito), que totaliza, aproximadamente, R\$ 1.300.000,00. Esse valor supera alguns outros gastos, como por exemplo, gasto com a cultura municipal, que totalizou um gasto de R\$146.000,00 no ano de 2017.

4.4 Relatórios Específicos Gerados

Os gráficos gerados, como mostrado anteriormente, são muito abrangentes, com uma grande densidade de dados em uma única barra. Como os gráficos são geralmente grandes, expandir essas barras pode levar a uma confusão na leitura do usuário final, podendo ficar perdido diante tantos dados.

Diante disso, uma alternativa encontrada foi a utilização de Relatórios Específicos que complementam os gráficos listados. Possuem, individualmente, dados mais pontuais, com um objetivo específico. Dessa forma, ao analisar um dos gráficos, pode-se recorrer aos relatórios para sanar alguma dúvida ou curiosidade.

Por exemplo, na Figura 5 podemos verificar a enumeração dos Favorecidos e seus respectivos cargos, abordando as despesas de diárias. Assim, ao analisar o gasto por favorecido nos gráficos, caso seja necessário um maior entendimento, é possível recorrer ao relatório que contém o nome do favorecido, o

cargo que ocupa em ordem alfabética. De forma semelhante, é possível realizar diversas outras consultas.

Cargos de Diárias

Nome	Cargo
Adailton Antonio Moreira	Motorista Carteira D
Adaise Barroso Sampaio	Nutricionista
Adelmo Batista Lemos	Biologo
Adriana Aparecida De Sousa	Fiscal Sanitario
Adriana Aparecida Moreira	Auxiliar De Enfermagem
Adriana Silva Naime Moura	Gerente
Agliã Campolina Leitao Mendonca	Assessor V
Ailton Rodrigues Maia	Auditor Controle Interno
Alaercio Antonio Delfino	Secretaria
Alessandra Marcelino Rodrigues Lemos	Auxiliar De Enfermagem
Alexandre Carlos De Faria	Fiscal De Transito E Transport
Aline Da Silva Soares	Gerente
Amanda De Miranda E Silva	Gerente
Amaury Jose Da Silva	Motorista Carteira D
Ana Carolina Campolina Santos	Gerente
Ana Carolina Carvalho Melo Franco	Fiscal Ambiental

Figura 5: Relatório Específico parcial dos cargos das Despesas Diárias

4.5 Possíveis Inconformidades

Estão listadas abaixo algumas possíveis inconformidades encontradas nos dados capturados.

1. As Despesas de Diárias mostram o pagamento a pessoas físicas que oferecem serviços à prefeitura, como por exemplo, motoristas. Na Figura 6 podemos observar o gráfico referente à quantidade de diárias pagas aos favorecidos. O último favorecido no Gráfico recebeu um total de 41 diárias durante todo o ano. Em sua grande maioria, referente a alimentação enquanto esteve a serviço da prefeitura.

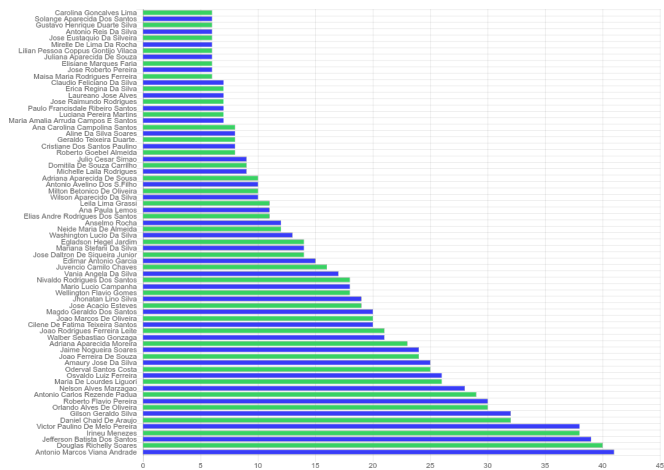


Figura 6: Gráfico parcial de quantidades de diárias pagas.

Ao analisar o Relatório Específico na Figura 7, referente às Despesas Diárias (Pessoas Físicas), podemos concluir um gasto de R\$8.650,00 em 2017 no nome do favorecido em questão. Ele possui o cargo de Gerente, e recebeu mais diárias do que qualquer outro funcionário, até mesmo motoristas, que devem transportar os representantes da prefeitura. Como pode ser observado, os motoristas ocupam as primeiras posições dos maiores gastos por diárias.

Despesas por Diárias (Pessoas Físicas)

Nome	Valor Total	Cargo
Douglas Richelly Soares	12510.00	Motorista Carteira D
Oswaldo Luiz Ferreira	12385.00	Motorista Carteira D
Irineu Menezes	11940.00	Motorista Carteira D
Juvencio Camilo Chaves	11365.00	Motorista Carteira D
Jaime Nogueira Soares	10775.00	Motorista Carteira D
Victor Paulino De Melo Pereira	10165.00	Motorista Carteira D
Walber Sebastiao Gonzaga	10030.00	Motorista Carteira D
Adriana Aparecida Moreira	9940.00	Auxiliar De Enfermagem
Cilene De Fatima Teixeira Santos	9910.00	Auxiliar De Enfermagem
Jose Acacio Esteves	9895.00	Motorista Carteira D
Roberto Flavio Pereira	9505.00	Motorista Carteira D
Orlando Alves De Oliveira	8820.00	Motorista Carteira D
Antonio Marcos Viana Andrade	8650.00	Gerente
Gilson Geraldo Silva	8400.00	Motorista Carteira D
Washington Lucio Da Silva	8325.00	Motorista Carteira D

Figura 7: Relatório Específico parcial de Pessoas Físicas

É importante ressaltar que essa informação não identifica um gasto indevido propriamente dito, e sim levanta a hipótese de que possa ser um. A principal ideia então, é repassar esse tipo de informação retirada a partir do WEB Analytics para pessoas, ou órgãos responsáveis, para que estas verifiquem a veracidade e deem prosseguimento à denúncia caso necessário.

2. As Multas Indedutíveis mostram as multas geradas por funcionários da prefeitura, em maioria motoristas. Na Figura 8 podemos observar os gastos com multas em diferentes locais e órgãos específicos. Foram gastos, em 2017, R\$7.599,27 com multas. As multas referem-se a municípios da redondeza da cidade, como Belo Horizonte e Contagem.

Um exemplo de multa citada é pelo veículo de placa PUN-0356, por usar equipamentos de som com volume não permitido na cidade de Belo Horizonte. O condutor se chama Júlio César Simão, o pagamento da multa foi retirado dos recursos da saúde da prefeitura (processo informado: 4087/16).

Outro exemplo interessante a se ressaltar, ainda na mesma classificação, temos uma multa em nome de Juvêncio Camilo Chaves, veículo placa HMM-9624, por avançar sinal vermelho no município de Belo Horizonte. O pagamento dessa multa foi, assim como a anterior, através dos recursos de saúde na prefeitura, com o número do processo 09442/12.

Assim, com base nos dados encontrados, fica evidente a necessidade de maior responsabilidade por parte dos

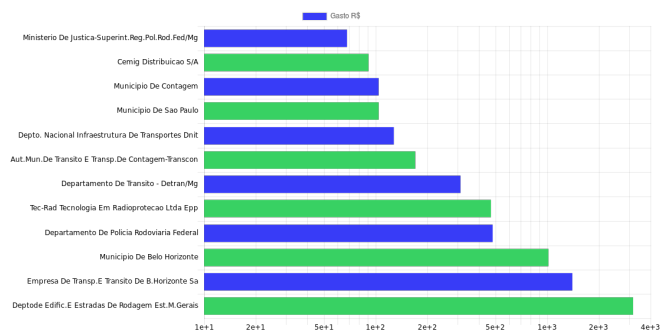


Figura 8: Gráfico de multas indedutíveis.

funcionários a serviço da prefeitura, para evitar gastos desnecessários, além de providenciar medidas para que não seja recorrente esse tipo de despesa.

4.6 Sensibilidade à Rede

Muitas vezes, a execução do programa pode ser mais rápida que a resposta do servidor onde se encontra os dados do portal. Devem existir pontos específicos no código do sistema para tratar a latência da rede, garantindo maior confiabilidade dos dados capturados.

A própria ferramenta utilizada possui recursos de espera de resposta do servidor, que foram utilizadas para atingir esse objetivo. Esse recurso aguarda o carregamento total da página ou algum atributo específico, antes de capturar efetivamente.

5. CONCLUSÕES E TRABALHOS FUTUROS

Podemos concluir que, atualmente, existe uma necessidade de análise dos dados no âmbito municipal, por falta de iniciativa dessa prática. Juntamente com o trabalho existente que abrange entes federativos superiores, podemos chegar em um momento em que todos entes sejam atingidos.

Utilizando recursos gráficos com agrupamentos de atributos semelhantes, pode-se obter um bom resultado no entendimento de grande quantidade de dados. Havendo uma alta densidade nos gráficos, deve-se disponibilizar formas mais simples de buscas que possam auxiliar em seu entendimento e nas conclusões esperadas.

Como trabalhos futuros, podemos incluir um modelo automático na análise dos dados, abordando por exemplo aprendizado de máquina. Assim, pode-se criar um ambiente mais interativo e sempre atualizado, de acordo com as configurações inseridas. A inteligência de máquina poderia identificar algum dado indevido, algum valor incomum ou inconsistente encontrado e assim por diante.

Além dos dados já disponibilizados no Web Analytics, construir um sistema web com diversos recursos com suas opções personalizáveis. Sendo assim, o usuário final faria seu login no sistema (para fins de controles de acesso), poderia solicitar a captura dos dados de sua cidade apenas informando a URL do Portal da Transparência de sua prefeitura. Depois disso, todos os dados ficariam lá disponíveis, e sempre atualizados. Qualquer outro usuário que deseje visualizar os dados poderia realizar o login no sistema e seria redirecionado para as informações já existentes.

Tornar o sistema mais completo, em que os dados de uma cidade serão permanentemente fiscalizados, salvo alguma so-

licitação de interrupção. Iniciar o processo com outro portal, diferente do que está em andamento, não deve acarretar sobrecarga ou interação entre as informações. Cada município deve ter seus próprios gráficos e relatórios, podendo se alternar entre os existentes e disponíveis.

Um detalhe importante é que a Web Analytics em sua primeira versão foi toda desenvolvida utilizando o framework Django (6). Sabendo que a linguagem de programação Python possui diversos recursos referentes a Data Science, além de um crescente estudo em desenvolvimento, pode-se utilizar esses recursos no servidor em que se encontra o sistema web, abrindo a possibilidade de uma constante otimização.

6. REFERÊNCIAS

- [1]BRASIL. *LEI Nº 12.527*. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm>. Acessado em 31/05/2018.
- [2]PORTAL DA TRANSPARÊNCIA DE PARÁ DE MINAS. Disponível em: <<http://transparencia.parademinas.mg.gov.br/principal>>. Acessado em 31/05/2018.
- [3]OPEN KNOWLEDGE BRASIL. Disponível em: <<https://br.okfn.org>>. Acessado em 31/05/2018.
- [4]OPERAÇÃO SERENATA DE AMOR. Disponível em: <<https://serenata.ai/>>. Acessado em 31/05/2018.
- [5]SELENIUM BROWSER AUTOMATION. Disponível em: <<https://www.seleniumhq.org/>>. Acessado em 31/05/2018.
- [6]DJANGO WEB FRAMEWORK. Disponível em: <<https://www.djangoproject.com/>>. Acessado em 31/05/2018.