

# Solução de BI para Monitoramento de Gastos e Perfil de Deputados Federais Eleitos

*BI Solution to Monitoring Expenses and Profile of the Elected Federal Deputies*

**Diogo Farkatt Kabbaz<sup>1</sup>**

 [orcid.org/0000-0001-5772-6177](https://orcid.org/0000-0001-5772-6177)

**Elizabeth Lopes Silva<sup>2</sup>**

 [orcid.org/0000-0003-3325-3058](https://orcid.org/0000-0003-3325-3058)

**Paulo Henrique Oliveira<sup>3</sup>**

 [orcid.org/0000-0003-0590-4220](https://orcid.org/0000-0003-0590-4220)

<sup>1</sup>Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil.  
E-mail: [kabbaz@gmail.com](mailto:kabbaz@gmail.com)

**DOI: 10.25286/rep.v6i5.1749**

Esta obra apresenta Licença [Creative Commons Atribuição-Não Comercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).

Como citar este artigo pela NBR 6023/2018: KABBAZ, D. F.; SILVA, E. L.; OLIVEIRA, P. H. Solução de BI para Monitoramento de Gastos e Perfil de Deputados Federais Eleitos. Revista de Engenharia e Pesquisa Aplicada, Recife, v.6, n. 5, p. 11-17, Novembro, 2021.

## RESUMO

Apesar dos avanços alcançados com as leis de transparência, a maneira como esses dados são publicados não estimulam o interesse e não contribuem com a fiscalização por parte da sociedade. Nesse artigo, é proposto um BI dos gastos dos deputados federais com as cotas parlamentares e ainda o perfil dos parlamentares que ocuparam e ocupam a Câmara dos Deputados. As visualizações desenvolvidas permitem conhecer como se dá a distribuição de gênero, faixa etária e nível de escolaridade dos deputados, além de compreender como as cotas parlamentares estão sendo utilizadas. Oferecendo, assim, uma ferramenta de fiscalização útil à população. O banco de dados utilizado no desenvolvimento do BI é público e está disponível em [1], em seguida, o ETL foi realizado com o auxílio do software Pentaho e, finalmente, as visualizações foram desenvolvidas com o software Tableau.

**PALAVRAS-CHAVE:** Business Intelligence; BI; Leis de transparência; ETL; Pentaho; Tableau; Análise de dados; Visualização de dados;

## ABSTACT

Despite the advances achieved by the transparency laws, the way these data are published does not stimulate interest and does not contribute to the oversight by society. This paper proposes a BI solution to monitoring the expenses of the federal deputies and also the profile of the parliamentarians who have occupied and who occupy the the House of Representatives. The visualizations designed allow to know how the distribution of ge The visualizations developed allow us to know how the distribution of gender, age group and level of education of the deputies is, besides understanding how the parliamentary quotas are being used. Thus, offering a useful monitoring tool to the population. The database used in the development of the BI is public and is available at [1], then the ETL was performed with the help of Pentaho software and, finally, the visualizations were developed with Tableau software.

**KEY-WORDS:** Business Intelligence; BI; Transparency laws; ETL; Pentaho; Tableau; Data analysis; Data visualization;

## 1 INTRODUÇÃO

Realizar um acompanhamento eficiente sobre os dados públicos não é uma atividade simples pois, apesar das leis nacionais de transparência exigirem que os órgãos públicos disponibilizem esses dados, não significa que esses estarão estruturados para fácil análise e interpretação da população.

É comum encontrar esses dados em formato de planilhas ou, para informações mais recentes, por meio de APIs (*Application Programming Interface*), dispostos visando unicamente cumprir com a regulamentação.

Assim, visando solucionar tal problemática, esse projeto teve como objetivo criar uma ferramenta que organiza e dispõe uma parte dos dados disponibilizados pela Câmara dos Deputados Federais [1], de forma a possibilitar que a população consiga acompanhar tanto as mudanças de paradigmas ao longo das legislaturas, como também os gastos realizados por partidos políticos, Estado e/ou Deputado eleito.

Dito isso, o objetivo fundamental desse projeto foi, por meio do processo de *Business Intelligence* (BI), transformar esses dados brutos em informação útil para o público em geral, permitindo, assim, uma tomada de decisão consciente para as próximas eleições, cobrança aos deputados atualmente empossados e/ou meio de comprovação de notícias e combate às *fake news*.

Dessa forma, primeiramente, esse artigo se propõe a analisar as três últimas legislaturas, Quadro 1.

**Quadro 1** - Legislaturas

Legislatura	Início	Fim
54	2011	2015
55	2015	2019
56	2019	2021*

\*Ano corrente a data de desenvolvimento desse projeto

Por conseguinte, vale salientar que, ainda que tal projeto tenha se delimitado à análise apenas de três legislaturas, nada impede que esse escopo seja expandido para ocorrências prévias e futuras.

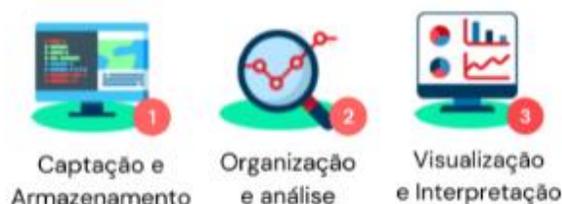
Ademais, no decorrer do artigo, serão explanadas as etapas e ferramentas utilizadas para extração, processamento e análise dos dados obtidos e as técnicas de BI aplicadas.

## 2 FONTE DE DADOS

O primeiro item a ser estruturado na proposta do projeto foi a escolha da base de dados. Após a avaliação de bancos existentes, identificou-se que os dados disponibilizados pela Câmara dos Deputados Federais [1] dariam análises coerentes e, com isso, seria possível construir uma ferramenta útil para acompanhamento dos deputados.

Logo, após a avaliação da necessidade de negócio, prosseguiu-se o processo descrito na Figura 1 para construção da solução.

**Figura 1** - Processo de BI



Fonte: [2].

No website dos dados abertos [1], observou-se que os dados dos deputados são disponibilizados em APIs para integração usando o formato *json* ou *xml* e, para períodos mais antigos, em arquivos *xml*. As informações disponíveis variam entre dados dos deputados, despesas, fornecedores, discursos, eventos, entre outros.

Após avaliação dos dados, definiu-se qual seria o foco do desenvolvimento, sendo ele, os gastos dos deputados e, também, as características dos deputados, no que tange: sexo, idade, escolaridade e partido político.

Com essas informações dispostas de forma apropriada, algumas questões foram avaliadas, como:

- A escolaridade dos deputados eleitos por legislatura;
- A proporção entre homens e mulheres;
- Os deputados e/ou partidos que possuem mais gastos (quais são esses gastos? Quais os principais fornecedores?);
- A média de gastos por Estado.

## 3 COLETA DE DADOS

Para executar a coleta dos dados, utilizou-se o *Pentaho Data Integration*, ferramenta que permite realizar as etapas do ETL (*Extraction, Transformation and Loading*), que são:

- **Extração** – os dados são coletados dos bancos originais, que podem ser de diferentes fontes. Conforme Costa [3], os dados devem ser extraídos diante da necessidade do projeto, para uma economia não só do espaço em disco, como também do tempo de extração.
- **Transformação** – A transformação dos dados é o segundo e principal passo do processo ETL [5]. Nessa etapa, será realizada toda a limpeza de dados necessária e modificação dos dados.
- **Carga dos dados** – Os dados são enviados para o banco de destino.

Antes de iniciar o ETL, é necessário realizar a modelagem multidimensional dos dados, que possui dois tipos de tabelas:

- **Tabela fato** - valores mensuráveis do negócio.

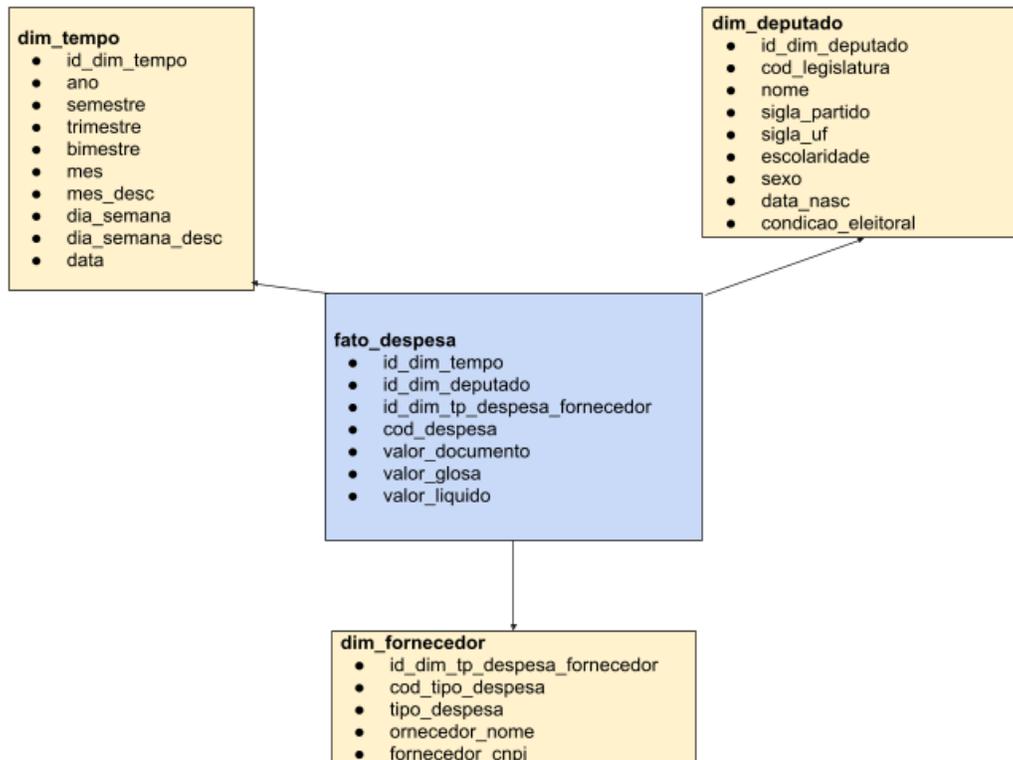
- **Tabela dimensão** - atributos das dimensões.

Assim, considerando o objetivo do BI e as questões levantadas, o esquema protagoniza com as tabelas de fato e dimensão, Figura 2, em que foi o escolhido para a modelagem dos dados em análise.

Após a definição do modelo, foram criadas as transformações necessárias para extração e tratamento dos dados, como pode ser visto, para ilustração, da dimensão deputado (Figura 4) e de fato despesa (Figura 5).

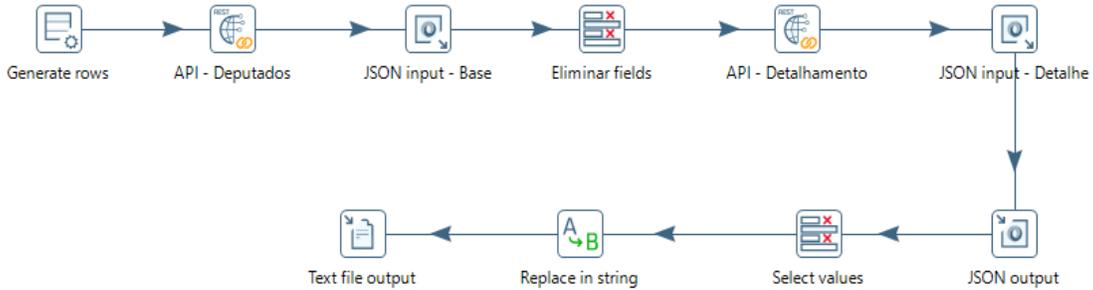
Enfim, o modelo pôde ser criado com o auxílio do *Pentaho*. Para ilustrar, as Figuras 4 e 5 apresentam a criação da dimensão deputado. Na Figura 4, é mostrada a criação dos arquivos tipo *json* que são parte da base de dados da dimensão deputado. Na Figura 5, tem-se o processo de criação da dimensão deputado baseada na integração dos dados, junção, limpeza dos dados e, finalmente, um *lookup*.

Figura 2 - Esquema estrela do BA



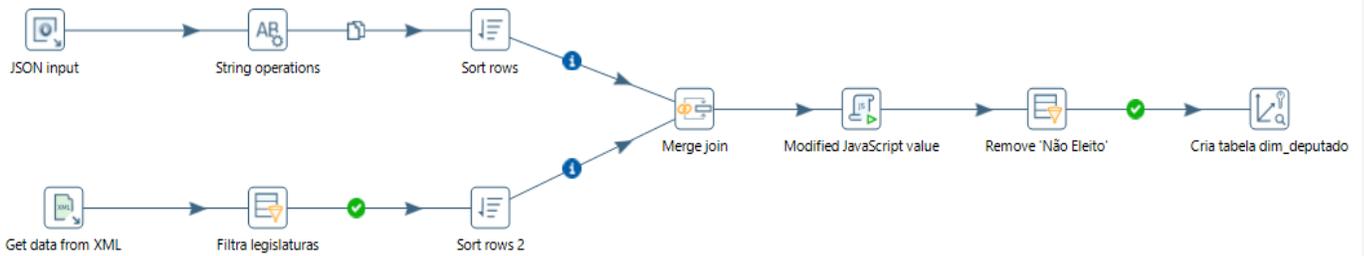
Fonte: Os Autores.

Figura 3 – Criação do arquivo Deputados



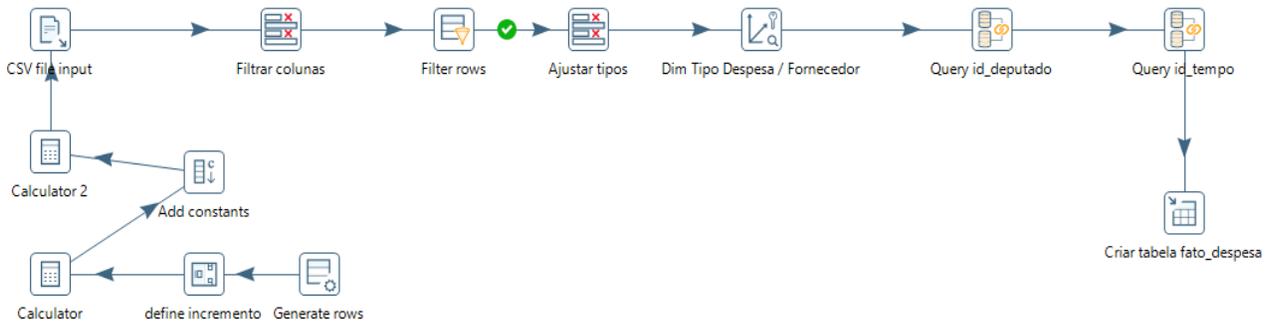
Fonte: Os Autores.

Figura 4 – Dimensão Deputado



Fonte: Os Autores.

Figura 5 – Tabela Fato Despesa



Fonte: Os Autores.

## 4 ANÁLISE DE DADOS

Como o principal objetivo do artigo é permitir um acompanhamento dos gastos públicos dos deputados e um maior conhecimento das pessoas que ocupam a Câmara do Deputados, a forma como as informações são apresentadas é a etapa fundamental do projeto e para realizar foi escolhido o *software* de visualização de dados

Tableau, que possui uma série de produtos (*Tableau Public, Tableau Desktop, Tableau Mobile, Tableau Online*) e se propõe a facilitar ao usuário a exploração e gerenciamento dos dados e agiliza a descoberta e o compartilhamento de informações [6].

Para construir os dashboards, foi utilizado o *Tableau Desktop - Public Edition* que possui uma versão teste que permitiu a criação das

visualizações necessárias. E, em seguida, a própria ferramenta disponibiliza uma página [7] para acesso às visualizações construídas.

## 5 METODOLOGIA

A interface desenvolvida [7] foi separada em quatro *dashboards*: Gastos por Deputado; Gastos por Região; Perfil dos Deputados e Comparativo de Gêneros. É importante destacar, que algumas comparações devem ser realizadas entre pares. Por exemplo, analisar os deputados de um estado ou região para validade da comparação, pois alguns gastos são diretamente relacionados com a localização, como Passagens Aéreas.

### 5.1 GASTOS POR DEPUTADO

No primeiro *dashboard*, tem-se informações dos gastos de um único deputado (selecionado pelo usuário). A informação principal (Figura 6), os gastos mensais, são apresentados em forma de gráfico de barra com esquema de cores para permitir análise visual mais simplificada (verde para gastos mensais abaixo de vinte mil reais, amarelo para gastos entre vinte e quarenta mil reais, vermelho para gastos acima de quarenta mil reais mensais). O esquema de cores foi definido de forma arbitrária.

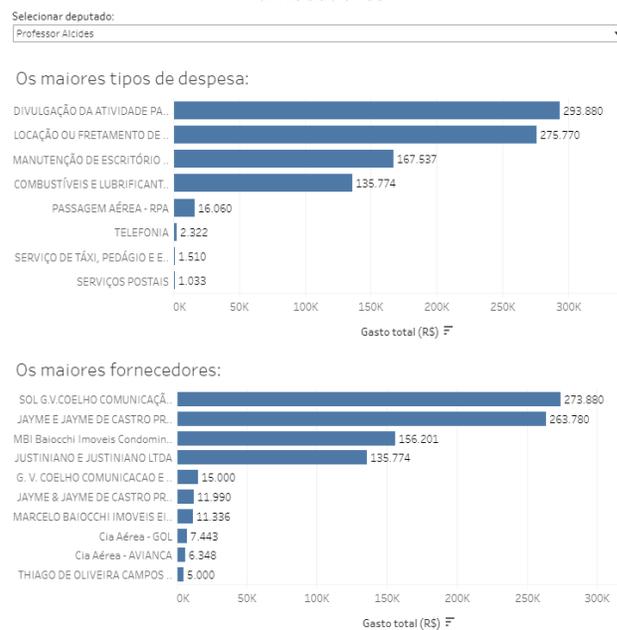
Figura 6 - Gastos mensais.



Fonte: Os Autores.

Além disso, existem duas outras informações nesse primeiro *dashboard* (Figura 7), os dez maiores tipos de despesa e os dez maiores fornecedores (razão social dos dez fornecedores cadastrados que mais faturaram).

Figura 7 - Os maiores tipos de despesa e os maiores fornecedores

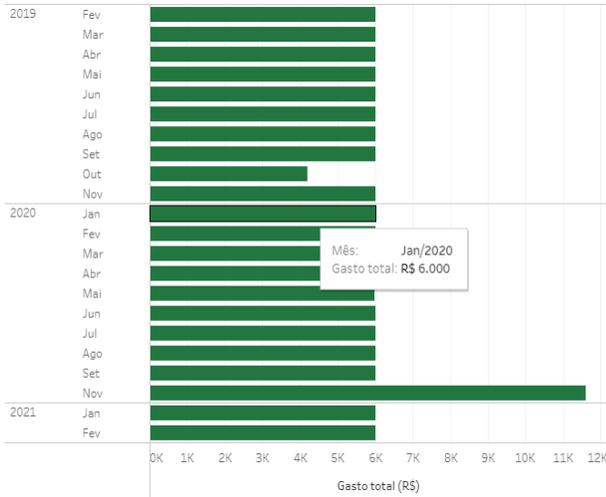


Fonte: Os Autores.

O *dashboard* é interativo, podendo-se filtrar os gráficos a partir de cliques nos outros gráficos. Por exemplo, se clicar no ano "2018", os outros dois gráficos passam a mostrar os dez maiores tipos de despesa de 2018 e os dez maiores fornecedores de 2018. Como exemplo de uso, selecionamos o deputado "Professor Alcides" e o tipo de gastos "COMBUSTÍVEIS E LUBRIFICANTES". Pode-se ver algo suspeito: o deputado selecionado gasta, em um único posto de gasolina, exatamente seis mil reais quase todos os meses (Figura 8).

Figura 8 – Exemplo de gastos por deputado

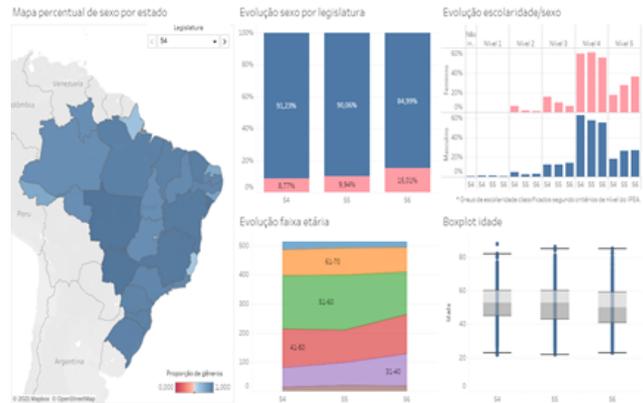
Gastos mensais de Professor Alcides



Fonte: Os Autores.

eleitos. Foram consideradas características como sexo dos deputados, idade e escolaridade.

Figura 10 - Perfil dos Deputados.



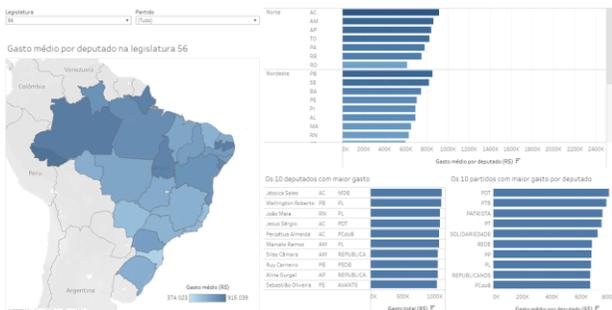
Fonte: Os Autores.

### 5.2 GASTOS POR REGIÃO

As informações contidas neste próximo dashboard (Figura 9) podem ser filtradas para um único partido ou para todos os partidos, e uma legislatura deve ser selecionada entre as disponíveis.

Foi colocado um mapa com um esquema de cores onde pode-se ver a média dos gastos totais dos deputados por estado (quanto mais escuro, maiores os gastos). Além disso, há a informação dos dez deputados com maiores gastos e os dez partidos com maiores gastos por deputado (média do partido). De forma semelhante ao anterior, pode-se atualizar as visualizações a partir de interações com os gráficos (filtrar por estados, partidos, região do Brasil).

Figura 9 - Dashboard Gastos por Região



Fonte: Os Autores.

### 5.3 PERFIL DOS DEPUTADOS

O terceiro dashboard (Figura 10) foi criado para se visualizar a evolução do perfil dos deputados

Um mapa foi criado para mostrar em cada estado a proporção entre homens e mulheres eleitos. Alguns destaques na última legislatura são o Acre (50% homens/50% mulheres), o Distrito Federal (37,5% homens/62,5% mulheres) e o Piauí (60% homens/40% mulheres). Além do mapa, há um gráfico da evolução da proporção total de homens e mulheres. Pode-se ver que, apesar da quantidade de mulheres eleitas estar aumentando (8,77% na legislatura 54, 9,94% na legislatura 55 e 15,01% na última legislatura), a grande maioria ainda é masculina.

Quando se compara o nível de escolaridade, classificado segundo critério de nível do IPEA - Instituto de Pesquisa Econômica Aplicada (Quadro 2), identifica-se o aumento do nível de escolaridade dos deputados eleitos.

Quadro 2 - Classificação dos níveis de escolaridade

Descrição	Nível de Escolaridade
Analfabeto	Nível 1
Até 5º Ano Incompleto	
5º Ano Completo	
6º ao 9º Ano do Fundamental	Nível 2
Fundamental Completo	
Médio Completo	Nível 3
Médio Completo	
Superior Incompleto	Nível 4
Superior Completo	
Mestrado	Nível 5
Doutorado	



Fonte: IPEA.

Na visualização do perfil do deputado eleito por faixa etária, observa-se que a parcela mais jovem vem sendo mais representativa com o passar do tempo.

## 5.4 COMPARATIVO GÊNEROS

Um último *dashboard* foi criado Figura 11 para representar um pouco melhor a diferença entre os dois gêneros, considerando faixa etária e escolaridade.

Figura 11 - Dashboard Comparativo entre Gêneros.



Fonte: Os Autores.

## 6 CONCLUSÕES

No contexto do projeto, identificamos que é possível criar uma solução para que a população possa interpretar os dados de forma eficiente. Com isso, o cidadão pode acompanhar os gastos dos deputados do seu estado e realizar cobranças, validar notícias veiculadas pela mídia e, também, combater a divulgação de *fake news*. Tudo isso junto, serve como uma ferramenta útil para tomada de decisão nas próximas eleições.

Vale salientar, que por ausência de dados suficientes, as trocas entre partidos não foram consideradas, portanto, sendo considerado escopo negativo do trabalho desenvolvido.

Como forma de evoluir o trabalho atual, propomos a inclusão de mais legislaturas na base de dados, para que tendências possam ser identificadas. Também, podemos incluir um

mecanismo para sincronização automática da base de dados, ou seja, com a disponibilização de novos dados, os *dashboards* seriam atualizados automaticamente. Outra melhoria possível, é reduzir a granularidade das categorias dos gastos e normalizar os fornecedores duplicados na base disponibilizada pela Câmara.

Além disso, a criação de um modelo de aprendizado de máquina para identificação de fraudes seria um passo importante para a análise do volume disponível de gastos.

## REFERÊNCIAS

- [1] DADOS ABERTOS DA CÂMARA DOS DEPUTADOS. Disponível em: <https://dadosabertos.camara.leg.br/swagger/api.html>. Acesso em 16 de fev. de 2021.
- [2] UCOMMERCE. **O Que É Business Intelligence E Para Que Serve O BI Em Uma Empresa?** 2020 Disponível em: <https://ucommerce.com.br/o-que-e-business-intelligence-e-para-que-serve-o-bi/>. Acesso em 20 de maio de 2021.
- [3] COSTA, Marcelo Augusto Souza da. **O processo de ETL na construção de conhecimento em uma aplicação de uma empresa seguradora.** Universidade Federal de Juiz de Fora. Juiz de Fora, 2009. Disponível em: <http://monografias.ice.ufjf.br/tcc-web/exibePdf?id=76>. Acesso em 21 de maio de 2021.
- [4] GODOI, Douglas. **Business Intelligence, conceito e definição.** Disponível em: <https://www.cetax.com.br/business-intelligence-tudo-o-que-voce-precisa-saber>. Acesso em 22 de maio de 2021.
- [5] KIMBALL, R.; CASERTA, J. **The Data Warehouse ETL Toolkit – Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data.** Indianápolis. Wiley Publishing, Inc., 2004. 490p.
- [6] TABLEAU. **O que é o Tableau?** 2021. Disponível em: <https://www.tableau.com/pt-br/why-tableau/what-is-tableau>. Acesso em 23 de maio de 2021.
- [7] LINK DASHBOARD DEPUTADOS. Disponível em: [https://public.tableau.com/profile/diogo.kabba\\_z#!/vizhome/ProjetoCECDA2019\\_2/StoryDeputados-GastosePerfil](https://public.tableau.com/profile/diogo.kabba_z#!/vizhome/ProjetoCECDA2019_2/StoryDeputados-GastosePerfil)