



**INSTITUTO
FEDERAL**
Brasília

Instituto Federal de Brasília

Campus Brasília

Curso Superior de Tecnologia em Sistemas para Internet

**TRATAMENTO E ANÁLISE DOS DADOS ABERTOS NA CÂMARA DOS
DEPUTADOS: UM ESTUDO DE CASO EM TECNOLOGIA, EMPRESAS E
PRODUTIVIDADE LEGISLATIVA**

Por

***DANIEL CRESPO DE SOUZA
LEANDRO SERPA RODRIGUES***

Tecnólogo

BRASÍLIA

2024

Daniel Crespo de Souza
Leandro Serpa Rodrigues

**TRATAMENTO E ANÁLISE DOS DADOS ABERTOS NA CÂMARA
DOS DEPUTADOS: UM ESTUDO DE CASO EM TECNOLOGIA,
EMPRESAS E PRODUTIVIDADE LEGISLATIVA**

*Trabalho apresentado ao Programa de Curso Superior de
Tecnologia em Sistemas para Internet da Instituto Federal
de Brasília como requisito parcial para obtenção do grau
de Tecnólogo em Sistemas de Internet .*

Orientadora: Ma. Roseane Salvio

Co-Orientador: Me. Fábio Ferraz Fernandez

BRASÍLIA
2024

S729 Souza, Daniel Crespo de.

Tratamento e análise dos dados abertos na Câmara dos Deputados: um estudo de caso em tecnologia, empresas e produtividade legislativa. / Daniel Crespo de Souza; Leandro Serpa Rodrigues. – Brasília, 2024.

58 f. : il. color.

Orientador: Roseane Salvio.

Coorientador: Fábio Ferraz Fernandez.

Trabalho de conclusão de curso (Graduação) – Instituto Federal de Educação, Ciência e Tecnologia de Brasília, Tecnologia em Sistemas para Internet, 2024.

1. Proposições. 2. Câmara dos Deputados. 3. Dashboard. I. Rodrigues, Leandro Serpa. II. Salvio, Roseane. (orient.). III. Fernandez, Fábio Ferraz. (coorient.). IV. Título.

CDU 004.65

Elaborado com os dados fornecidos pelo autor.

Daniel Crespo de Souza
Leandro Serpa Rodrigues

**Tratamento e Análise dos Dados Abertos na Câmara dos Deputados: Um Estudo de Caso
em Tecnologia, Empresas e Produtividade Legislativa**

Trabalho de conclusão de curso de graduação apresentado a Coordenação do Curso Superior de Tecnologia em Sistemas de Internet do Instituto Federal de Brasília – Campus Brasília, como requisito parcial para a obtenção do título de Tecnólogo em Sistemas de Internet.

Aprovado em: _____ de _____ de _____.

BANCA EXAMINADORA

Prof. Ma. Roseane Salvio
Computação/IFB

Prof. Me. Fábio Ferraz Fernandez
Computação/IFB

Prof. Esp. Hugo do Carmo Mendes Cesar
Computação/IFB

Prof.^a Ma. Kadidja Valéria Reginaldo de Oliveira
Computação/IFB

BRASÍLIA
2024

*Dedicamos este trabalho a nossa família, amigos, colegas,
professores e especialmente aos nossos orientadores, que
nos ajudaram nessa jornada até aqui. Que a ciência de
dados seja capaz de transmitir esse conhecimento para
todos.*

É o grau de comprometimento que determina o sucesso.

—J.K. ROWLING

Agradecimentos

Agradeço a minha orientadora Roseane Salvio, pela sabedoria com que me guiou nesta trajetória, também ao Co.Orientador Fábio Ferraz pela imensa oportunidade e disponibilidade de nos trilhar durante todo esse projeto.

Gostaria de deixar registrado também, o meu reconhecimento à minha família, amigos e aos meus colegas de trabalho, pois acredito que sem o apoio deles seria muito difícil vencer esse desafio. Enfim, a todos os que por algum motivo contribuíram para a realização deste trabalho.

Resumo

CRESPO, Daniel; SERPA, Leandro. TRATAMENTO E ANÁLISE DE DADOS ABERTOS NA CÂMARA DOS DEPUTADOS: UM ESTUDO DE CASO NA ÁREA DE TECNOLOGIA, EMPRESAS E PRODUTIVIDADE LEGISLATIVA. 2024. Trabalho de Conclusão de Curso (Graduação) – Tecnólogo em Sistemas para Internet. Instituto Federal de Brasília – Campus Brasília. Brasília/DF, 2023.

Devido ao crescente interesse de empresas de gestão de software nas atividades dos governantes, as instituições agora disponibilizam de maneira pública uma variedade de dados relacionados aos seus trabalhos. No contexto brasileiro, um desses conjuntos de informações refere-se às proposições da Câmara dos Deputados, abrangendo todas as matérias submetidas à avaliação do Senado, da Câmara ou do Congresso Nacional. Apesar da disponibilidade de diversos dados, as ferramentas de consulta existentes oferecem retornos apenas em formato textual. A utilização de gráficos interativos proporciona uma abordagem mais direta para a compreensão e análise desses dados disponíveis.

Este projeto específico concentrou-se em uma aplicação de visualização de dados, cujo objetivo é reunir informações abrangentes sobre proposições com ênfase em tecnologia, empresas e produtividade. A interface desenvolvida busca proporcionar uma experiência mais intuitiva e facilitada para o usuário por meio da construção de um dashboard. Diversas visualizações foram incorporadas à interface, abordando dados como o tipo de proposição, os partidos dos autores e os temas em questão. A aplicação oferece funcionalidades de filtragem, permitindo que o usuário refine as informações de acordo com a data de apresentação e os temas das proposições.

Palavras-chave: Proposições. Câmara dos Deputados. Dashboard.

Abstract

CRESPO, Daniel; SERPA, Leandro. TRATAMENTO E ANÁLISE DE DADOS ABERTOS NA CÂMARA DOS DEPUTADOS:UM ESTUDO DE CASO NA ÁREA DE TECNOLOGIA, EMPRESAS E PRODUTIVIDADE LEGISLATIVA. 2024. Trabalho de Conclusão de Curso (Graduação) – Tecnólogo em Sistemas para Internet. Instituto Federal de Brasília – Campus Brasília. Brasília/DF, 2023.

Because of the growing interest of a software management company in the activities of government officials, institutions now publicly provide a variety of data related to their work. In the Brazilian context, one of these sets of information pertains to propositions in the Chamber of Deputies, encompassing all matters submitted for evaluation by the Senate, the Chamber, or the National Congress. Despite the availability of diverse data, existing query tools only yield textual results, while the use of interactive infographics offers a more direct approach for understanding and analysis.

This specific project focused on a data visualization application with an emphasis on propositions related to technology, companies, and productivity. The developed interface aims to provide a more intuitive and streamlined experience for users through the construction of a dashboard. Various visualizations have been integrated into the interface, addressing data such as the type of proposition, the parties of the authors, and the relevant themes. The application offers filtering functionalities, allowing users to refine information based on the presentation date and proposition themes.

Keywords: Propositions. Chamber of Deputies. Dashboard.

Lista de Figuras

1.1	Imagem de dados brutos disponibilizados pela Câmara dos Deputados (CD) . . .	14
2.1	Palavras-chave	16
2.2	Imagem de números aleatórios	19
2.3	Imagem de números aleatórios com destaque no número 7	19
2.4	Imagem de um <i>Dashboard</i>	20
2.5	Imagem relacionada a visualização de dados	23
2.6	Comparativo entre os trabalhos	24
3.1	Ilustração do processo de dados	26
3.2	Sprints elaboradas	27
4.1	Ilustração de casos de uso	30
5.1	Imagem relacionada a proposições	33
5.2	Arquivo .csv em uma pasta no computador	33
5.3	Código de tratamento de dados parte 1	34
5.4	Código de tratamento de dados parte 2	34
5.5	Código de envio para o <i>Google Sheets</i>	35
5.6	Primeira versão do protótipo	36
5.7	Segunda versão do protótipo página 1	37
5.8	Segunda versão do protótipo página 2	37
6.1	Banco de Dados	40
6.2	Tela 1 do dashboard no Data Studio	41
6.3	Tela 2 do dashboard no Data Studio	41
6.4	Tela 3 do dashboard no Data Studio	42
6.5	Tela 4 do dashboard no Data Studio	42
6.6	Tela 5 do dashboard no Data Studio	43
6.7	Tela 6 do dashboard no Data Studio	43
6.8	Tela 7 do dashboard no Data Studio	44
6.9	Filtro do dashboard no Data Studio	44
6.10	Questionário do dashboard no Data Studio	45
6.11	Resultados pergunta 1 Google Forms	46
6.12	Resultados pergunta 2 Google Forms	46
6.13	Resultados pergunta 3 Google Forms	47
6.14	Resultados pergunta 4 Google Forms	47
6.15	Resultados pergunta 5 Google Forms	48

6.16 Resultados pergunta 6 Google Forms	48
6.17 Resultados pergunta 7 Google Forms	49
6.18 Resultados pergunta 8 Google Forms	49
6.19 Resultados pergunta 9 Google Forms	50
6.20 Resultados pergunta 10 Google Forms	50
6.21 Resultados pergunta 11 Google Forms	51
6.22 Resultados pergunta 12 Google Forms	51
6.23 Resultados pergunta 13 Google Forms	52

Lista de Acrônimos

API	Application Programming Interface	54
CD	Câmara dos Deputados	13
CSV	Conjunto de valores separados por vírgula	17
PEC	Propostas de Emenda à Constituição	17
PLP	Projetos de Lei Complementar	17
PL	Projetos de Lei Ordinária	17
PDC	Projetos de Decreto Legislativo	17
PRC	Projetos de Resolução	17
MPV	Medidas Provisórias	17
QUIS	Questionnaire for User Interaction Satisfaction	45
HCIL	Laboratório de Interação Humano-Computador	45

Sumário

1	Introdução	13
1.1	Problema	13
1.1.1	Objetivo geral	14
1.1.2	Objetivos específicos	15
1.1.3	Resultados Esperados	15
1.1.4	Estrutura do TCC	15
2	Conceitos gerais e revisão da literatura	16
2.1	Referencial teórico	16
2.1.1	Dados Abertos	16
2.1.2	Dados brutos e Tratamento de Dados	18
2.1.3	Análise e visualização de dados	18
2.1.4	Dashboard	20
2.1.5	Análise de Dados	21
2.2	Revisão de literatura	21
2.2.1	Análise de dados com R	21
2.2.2	Análise de dados abertos da Câmara	22
2.2.3	Coletando e analisando informações do Facebook com R	22
2.2.4	Visualização de dados de pessoas com doença de Parkinson	22
2.3	Comparativo entre os trabalhos	23
2.4	Tecnologias	24
2.4.1	Linguagem R	24
2.4.2	Google Sheets	25
2.4.3	Data Studio	25
3	Metodologia	26
3.1	Sprints SCRUM	27
4	Requisitos do sistema	29
4.1	Requisitos Funcionais	29
4.2	Requisitos não Funcionais	29
4.3	Casos de Uso	30
5	Desenvolvimento	32
5.1	Coleta de dados	32
5.2	Tratamento de dados	32
5.3	Banco de dados	35

5.4	Prototipação	35
5.4.1	Protótipo versão 1	36
5.4.2	Protótipo versão 2	36
5.5	Desenvolvimento no Data Studio	38
6	Apresentação e Análise dos Resultados	40
6.1	Banco de dados	40
6.2	Dashboard no Data Studio	40
6.3	Filtro de dados	44
6.4	Testes	44
6.4.1	Resultados dos testes de Usabilidade	46
6.4.2	Resultados após a compreensão dos dados da pesquisa	52
6.5	Feedbacks	52
7	Conclusões e Trabalhos Futuros	54
7.1	Trabalhos Futuros	54
7.1.1	API da Câmara dos Deputados	54
7.1.2	Implementações ao dashboard	55
	Referências	56

1

Introdução

O crescimento do acesso à informação é notável e garantir a qualidade e a quantidade desses dados é essencial (POSSAMAI; SOUZA, 2020). Isso se aplica não apenas à disponibilidade, mas também à liberdade de usar essas informações, expandindo a capacidade de entender legalmente como analisar, verificar, esclarecer e acompanhar dados governamentais de maneira contínua.

A crescente disponibilidade de conjuntos de dados informativos e ferramentas de software levou a uma maior dependência de visualizações de dados em muitas áreas (SILVA, 2021). Promovendo um conhecimento e explicando que a visualização de dados fornece uma maneira poderosa de comunicar descobertas baseadas em dados, motivar análises e detectar falhas (HARVARD, 2023).

A necessidade de entender os dados antes de transformá-los em visualização é destacada por (BENOIT, 2019). A ideia surgiu de traduzir as propostas legislativas textuais da Câmara dos Deputados (CD) em dados visuais, após receber esses dados em formato bruto e tratá-los para uso posterior em uma plataforma que permita a visualização de dados.

Esse tratamento, deve possuir um contexto, legitimidade, atender as expectativas do consumidor e obter um contexto onde os dados foram coletados (CRAVO; JOELSONS, 2020). Este trabalho consiste em fazer a utilização desses dados em uma plataforma *web*, que utilize os dados abertos da CD com ênfase em tecnologia, produtividade e empresas, especificamente, transformado os dados textuais em um formato visual, por meio de *dashboards* interativos, com foco em atender a uma empresa de gestão de software que pretende usar esses dados em seu benefício.

Essa abordagem visa não apenas atender às necessidades específicas dessa empresa, mas também contribuir para a compreensão mais ampla das dinâmicas legislativas relacionadas a tecnologia, produtividade e empresas.

1.1 Problema

A visualização de dados é um ramo do conhecimento com uma característica específica, que é não apresentar apenas uma mensagem predefinida de seus criadores, mas também ser

interpretada como uma ferramenta que possibilita que as pessoas possam tirar suas próprias conclusões sobre os dados (ROSA et al., 2021).

Dessa forma, é um caminho a ser considerado quando queremos apresentar conjunto de dados para determinado público-alvo. Existem diversas plataformas para conseguir fazer uma boa análise de dados, porém muitas delas contêm algumas limitações, seja no tipo de dados que essa plataforma consegue ler e também no estilo de apresentação após a produção de uma análise de dados. Um exemplo seria a plataforma Power BI, conforme explicado por (CESAR; COSTA, 2022), a limitação nas formas de visualizar a apresentação dos dados e a carência por tecnologias adequadas podem afetar o gerenciamento dos dados.

Considerando a dificuldade de compreensão das propostas legislativas presentes nos dados abertos da CD, buscamos maneiras de torná-las mais acessíveis. Uma empresa com foco em gestão de software tem o interesse em visualizar esses dados de uma forma dinâmica e mais compreensível, como resalta SHARDA et al. (2019), mas infelizmente os dados abertos da CD estão disponíveis apenas em forma textual como é demonstrado na figura 1.1.

Figura 1.1 Imagem de dados brutos disponibilizados pela CD

528	"PL 2956/2023";"Altera a Lei nº 12.830, de 20 de junho de 2013, para estabelecer a garantia do direito do investigado de ser ou
529	"PL 2960/2023";"Institui o Programa de Apoio a Famílias e Responsáveis por Pessoas Portadoras de Atrofia Muscular Espinhal (AME
530	"PL 2961/2023";"Altera a Lei nº 13.146, de 6 de julho de 2015, para estabelecer reserva de vagas para deficientes em concursos
531	"PL 2969/2023";"Altera o art. 25 da Lei nº 9.605, de 12 de fevereiro de 1998 que "dispõe sobre as sanções penais e administrati
532	"PL 3014/2023";"Altera o art. 8º da Lei nº 9.250, de 26 de dezembro de 1995, para incluir as despesas com compra, implante e ma
533	"PL 3007/2023";"Assegura ao estudante com síndrome de down o direito à matrícula em escola regular e, em escola da educação esp
534	"PL 4807/2023";"Dispõe sobre a publicidade e propaganda realizada por profissionais liberais.
535	";"";"Kim Kataguirí";"SP";"UNIÃO";"03/10/2023";"Aguardando Parecer";"http://www.camara.leg.br/proposicoesWeb/fichadetramitacao?
536	"PL 4818/2023";"Dispõe sobre o direito de matrícula e frequência de crianças diagnosticadas com diabetes nas escolas da Rede Pú
537	";"";"Flávia Moraes";"GO";"PDT";"03/10/2023";"Aguardando Designação de Relator";"http://www.camara.leg.br/proposicoesWeb/fichad
538	"PL 4822/2023";"Dispõe sobre a discricionariedade dos agentes públicos no uso de câmeras individuais para filmar e monitorar su
539	";"";"Daniela Reinehr";"SC";"PL";"03/10/2023";"Tramitando em Conjunto";"http://www.camara.leg.br/proposicoesWeb/fichadetramitac
540	"PL 4815/2023";"Reconhece o processo de produção de artesanato com a palha de Ouricuri do Pontal de Coruripe (AL) como manifest

Fonte: captura de tela elaborada pelo autor.

Estes dados após um tratamento, e compartilhados com uma visualização mais adequada, diferente da textual, ajudariam a empresa a entender as estratégias do governo e também como essa empresa de gestão de software poderia aproveitar essas políticas públicas em seu benefício para encontrar novos possíveis clientes.

Neste sentido, seria possível que uma plataforma web de análise de dados resolvesse esses problemas, sendo uma opção eficiente e de fácil compreensão e mais acessível que outras plataformas?

1.1.1 Objetivo geral

Implementar uma interface em uma plataforma web para análise de dados abertos na CD, com ênfase na pesquisa sobre tecnologia e produtividade de empresas, que contribua com uma visualização dos dados por meio de uma criação de *dashboards* interativos.

1.1.2 Objetivos específicos

- Coletar e catalogar os dados abertos relacionados às propostas legislativas com ênfase em tecnologia, empresas e produtividade na CD.
- Extrair informações relevantes sobre o conteúdo das proposições relacionadas à tecnologia, empresas e produtividade.
- Identificar os parlamentares e os partidos mais atuantes em proposições sobre tecnologia, empresas e produtividade.
- Classificar os tipos de proposições mais frequentes.
- Desenvolver uma interface de usuário intuitiva, para uma navegação pelos dados de forma eficiente.
- Filtrar informações referentes as propostas legislativas pela ementa.

1.1.3 Resultados Esperados

O projeto visa realizar um bom tratamento de dados, onde os dados oriundos da CD sejam transferidos de uma visualização textual para uma visual. Isso ocorrerá por meio de gráficos, com a ideia de conclusão de um dashboard interativo, com diversas funções de filtragem de dados.

1.1.4 Estrutura do TCC

O presente trabalho está organizado da seguinte forma: No Capítulo 1 é apresentada a introdução, incluindo: o tema, problema, objetivos, resultados esperados e a estrutura. No Capítulo 2 são descritos os principais trabalhos e literaturas relacionados ao tópico aqui abordado. No Capítulo 3 é apresentada a metodologia utilizada no projeto. No Capítulo 4 onde são apresentados os requisitos do sistema e os casos de uso, seguido do Capítulo 5 que está designado para o desenvolvimento do trabalho. No Capítulo 6 são apresentados os resultados obtidos durante o desenvolvimento do trabalho. No Capítulo 7 estão a conclusão e os trabalhos futuros.

O Capítulo 2, a seguir, tem o intuito de compreender como as pesquisas em livros, artigos, trabalhos correlatos e sites trouxeram contribuições imensas ao projeto e desenvolvimento deste trabalho.

2

Conceitos gerais e revisão da literatura

Nesse capítulo será abordado o referencial teórico que embasa esse trabalho, a revisão de literatura e a descrição das tecnologias utilizadas no seu desenvolvimento.

2.1 Referencial teórico

Nessa seção, serão descritos os livros, artigos e pesquisas que foram utilizados para a produção deste trabalho, destacando os conceitos de: dados abertos, visualização de dados, tratamento de dados e análise de dados.

Figura 2.1 Palavras-chave

Primeiro conjunto	Dados abertos visualização de dados
Segundo Conjunto	Tratamento de dados, linguagem R
Terceiro Conjunto	Análise de dados, Dashboard

Fonte: elaborado pelo autor.

2.1.1 Dados Abertos

"A política de Dados Abertos do Poder Executivo Federal foi instituída pelo Decreto nº 8.777, de 11 de maio de 2016."(Brasil. Governo Federal, 2023). A filosofia de Dados Abertos, conforme definida pela *Open Knowledge Foundation*, refere-se à prática de tornar dados

governamentais acessíveis ao público em formatos reutilizáveis. Essa abordagem visa aumentar a transparência, promover maior participação política dos cidadãos e estimular o desenvolvimento colaborativo de aplicações pela sociedade. Os dados abertos permitem que qualquer pessoa acesse, utilize, modifique e compartilhe livremente, sujeito a requisitos que preservem a origem e a natureza aberta dos dados Brasil. Governo Federal (2023).

Segundo SILVA (2023, p. 2).

O compartilhamento de dados no âmbito da APF (Administração pública federal) estimula que dados abertos e privados possam ser compartilhados aumentando a ligação entre os dados, evitando redundâncias e promovendo a interoperabilidade.

No contexto brasileiro, diversos órgãos da Administração Pública já adotam a prática de publicar seus dados na web, por meio de relatórios e balanços, proporcionando aos cidadãos a oportunidade de acompanhar os resultados das ações governamentais.

Os dados disponibilizados pela CD que utilizaremos são as proposições, e segundo o (Portal da Câmara dos Deputados, 2021), são matérias sujeitas à deliberação da Câmara, onde as proposições consideradas principais são:

- Propostas de Emenda à Constituição (PEC)
- Projetos de Lei Complementar (PLP)
- Projetos de Lei Ordinária (PL)
- Projetos de Decreto Legislativo (PDC)
- Projetos de Resolução (PRC)
- Medidas Provisórias (MPV)

Um fator importante a se ressaltar é que esses dados estão disponíveis em formato textual bruto chamado Conjunto de valores separados por vírgula (CSV), que se refere a uma lista de valores separados por vírgulas.

CAFé; SALES (2010) sugerem a organização de dados brutos de uma maneira a ser melhor compreendido, assim auxiliando na utilização desses dados que nesse trabalho se referem às atividades parlamentares na CD, precisamente as proposições legislativas disponíveis como dados abertos no site da CD.

Compreendendo essa informação, foi decidido utilizar as proposições mencionadas desenvolvendo um *dashboard* para transformar esse modelo textual em visual.

2.1.2 Dados brutos e Tratamento de Dados

O tratamento de dados vem ganhando muita importância no ambiente tecnológico. SILVA et al. (2020) observa que os dados abertos disponíveis do setor público, são obtidos em formato bruto, assim, tornando acessível e permitindo a sua reutilização com informações diferentes, gerando novas formas de entendimento desses dados.

Sobre o tratamento e uso dos dados brutos, aborda GALDINO et al. (2022) que um bom tratamento tornam os dados compreensíveis e a empresa com interesse na pesquisa passa a ter uma melhor visualização e entendimento após um tratamento eficaz de dados.

2.1.3 Análise e visualização de dados

A análise e visualização de dados vêm se tornando cada vez mais comum em meio ao mundo conectado, segundo NUSSBAUMER (2015). Com o grande aumento de dados disponíveis, surgiu uma necessidade de fazer com que esses dados tenham sentido.

É perceptível a importância fundamental da habilidade de transformar dados, especialmente aqueles provenientes de fontes textuais, como os registros da CD, em representações visuais compreensíveis.

A capacidade de visualizar e explicar esses dados de maneira gráfica não apenas facilita a compreensão, mas também torna as informações mais acessíveis e envolventes para o público em geral.

Como afirma (LANKOW et al., 2012) (tradução nossa).

Todas as visualizações contêm tais atributos, e usá-los adequadamente para transmitir informações é a chave para a comunicação visual. Nossos cérebros são capazes de reconhecer e processar muitos desses sinais visuais simultaneamente por meio de um processo chamado processamento pré-atentivo.¹

A visualização de dados se torna muito importante, e impacta drasticamente em quanto tempo é possível entender uma visualização.

Esses atributos mencionados por LANKOW et al. (2012) podem ser cores, formas, tamanhos, posições, entre outros. Quando esses atributos são usados corretamente, eles podem transmitir informações de maneira eficaz. Isso é muito importante na comunicação visual, pois permite que pessoas entendam complexidades e padrões nos dados de maneira intuitiva e rápida.

Segundo LANKOW et al. (2012) o termo processamento pré-atentivo é utilizado para descrever a capacidade do cérebro de reconhecer e processar informações visuais rapidamente e sem esforço consciente. Por exemplo, um gráfico de barras, não há necessidade de contar individualmente cada barra para saber qual é a maior. O cérebro reconhece isso instantaneamente através do processamento pré-atentivo.

¹All visualizations contain such attributes, and using them properly to convey information is the key to visual communication. Our brains are able to recognize and process many of these visual cues simultaneously through a course of action called preattentive processing.

Conforme demonstrado nas Figuras 2.2 e 2.3, é compreensível que a visualização de dados se torna interessante para o uso cotidiano, e que combinada com um tratamento de dados, se alinham para a construção impressionante de formas de visualização, como os gráficos.

Figura 2.2 Imagem de números aleatórios

2	1	4	3	9	5	6	7	8	2	3	6	5	9	4	0	1
6	7	9	3	4	9	0	5	6	2	5	8	4	0	5	2	6
9	8	2	6	3	5	9	3	2	9	3	7	2	6	3	4	8
8	1	6	2	3	8	7	9	5	0	2	3	9	2	8	4	3
0	9	1	8	5	4	2	9	4	7	4	6	8	4	0	2	9
3	9	2	7	3	6	6	5	2	9	4	0	4	9	4	8	6
5	2	4	3	6	4	8	1	0	3	9	4	8	4	7	3	2
8	6	2	3	0	8	7	3	6	2	5	4	4	8	3	5	0

Fonte: Figura retirada de LANKOW et al. 2012.

A Figura 2.2 traz números aleatórios sem destaques em sua visualização.

Figura 2.3 Imagem de números aleatórios com destaque no número 7

2	1	4	3	9	5	6	7	8	2	3	6	5	9	4	0	1
6	7	9	3	4	9	0	5	6	2	5	8	4	0	5	2	6
9	8	2	6	3	5	9	3	2	9	3	7	2	6	3	4	8
8	1	6	2	3	8	7	9	5	0	2	3	9	2	8	4	3
0	9	1	8	5	4	2	9	4	7	4	6	8	4	0	2	9
3	9	2	7	3	6	6	5	2	9	4	0	4	9	4	8	6
5	2	4	3	6	4	8	1	0	3	9	4	8	4	7	3	2
8	6	2	3	0	8	7	3	6	2	5	4	4	8	3	5	0

Fonte: Figura retirada de LANKOW et al. 2012.

A Figura 2.3 traz números aleatórios com destaques em sua visualização.

Como é destacado no artigo ALBUQUERQUE et al. (2024) federações da indústria do Brasil são responsáveis por fornecer informações importantes e efetivas para tomadas de decisões. Com o objetivo de analisar as estruturas de visualização de informações ALBUQUERQUE et al. (2024) utilizaram conceitos, técnicas e tecnologias em visualização de dados para criar seus painéis interativos e divulgar informações em *dashboards*.

Mantendo assim a ênfase no que busca esse projeto, utilizando a análise de dados para auxiliar a visualização do *dashboard*.

2.1.4 Dashboard

Adquirindo o conhecimento necessário em entender como acontece a visualização de Dados, nesse projeto, foi decidido utilizar um *dashboard* para transmitir essa visualização.

Mas o que seria um *dashboard*? Um *dashboard* é um painel de controle em uma única tela, onde você consegue observar e analisar todas as informações que são necessárias para atingir um ou mais objetivos. Nesse projeto, essa tela seria dividida com gráficos e algumas explicações sobre os dados que contêm nesse gráfico.

Um gráfico segundo BENOIT (2019) é uma forma visual de representar dados ou informações, utilizando símbolos, cores, formas, linhas, barras e outros elementos gráficos para transmitir informações de maneira clara e rápida. Eles servem para ilustrar e comparar dados, mostrar relações entre diferentes conjuntos de dados e resumir informações complexas de uma maneira visualmente simples de compreender. Existem diversos tipos de gráficos como podemos observar na Figura 2.4, incluindo: gráficos de barras, linhas, pizza entre outros.

Figura 2.4 Imagem de um *Dashboard*



Fonte: Figura retirada de LANKOW et al. (2012).

De acordo com FEW (2006), *dashboards* são uma forma de visualização de dados em

forma de gráficos e com pouco texto. Este formato visual permite interpretar e visualizar uma análise mais eficaz dos dados, como é apresentado na Figura 2.4.

Nesse projeto, foi idealizado o uso de informações sobre os dados da CD, e o que esses dados estão descrevendo em uma forma visual com gráficos e atribuído uma observação embaixo para maior compreensão dessa informação. A observação serve para fornecer informações adicionais ou esclarecimentos sobre os dados apresentados no gráfico e ajudar a entender melhor o contexto e a origem dos dados, tornando a visualização mais informativa e útil.

2.1.5 Análise de Dados

Como é destacado no artigo ALBUQUERQUE et al. (2024) federações da indústria do Brasil são responsáveis por fornecer informações importantes e efetivas para tomadas de decisões. Com o objetivo de analisar as estruturas de visualização de informações ALBUQUERQUE et al. (2024) utilizaram conceitos, técnicas e tecnologias em visualização de dados para criar seus painéis interativos e divulgar informações em *dashboards*.

Mantendo assim a ênfase no que busca esse projeto, utilizando a análise de dados para auxiliar a visualização do *dashboard*.

2.2 Revisão de literatura

Nesta seção, são explorados trabalhos relacionados que abordam o tratamento de dados utilizando a linguagem de programação R, e o desenvolvimento e construção de *dashboards*, detectando um padrão de interatividade melhor e com mais opções abertas para o desenvolvimento dos gráficos presentes no *dashboard*. São pesquisas referentes aos trabalhos associados à aplicação deste projeto. Ao filtrar a pesquisa com base na Figura 2.1, foram identificados diversos trabalhos relevantes e interessantes, os quais serão detalhados a seguir.

2.2.1 Análise de dados com R

NOGUEIRA (2021) desenvolveu um trabalho que chamou a atenção por alguns aspectos interessantes. Este trabalho de conclusão de curso visa ajudar o departamento de estatística da Universidade de Brasília a analisar os dados dos alunos e disciplinas de forma mais eficiente, rápida e personalizada. Usando R e Python com suas bibliotecas, o autor desenvolveu *dashboards* interativos e informativos que auxiliam o departamento de estatística a analisar dados de alunos e disciplinas de forma mais eficiente, permitindo visualizar dados de matrícula, horário, quantidade de alunos, professores, entre outros. Essa diferença entre as aplicações é interessante, pois pode auxiliar essa pesquisa em como desenvolver, analisar e apresentar os dados da Câmara.

Os projetos apresentam algumas diferenças. O que está sendo desenvolvido neste projeto é uma plataforma que lê dados abertos da CD e cria um *dashboard*, já o de (NOGUEIRA (2021)) é a criação de *dashboards* interativos e informativos para o Departamento de Estatística da Universidade de Brasília.

2.2.2 Análise de dados abertos da Câmara

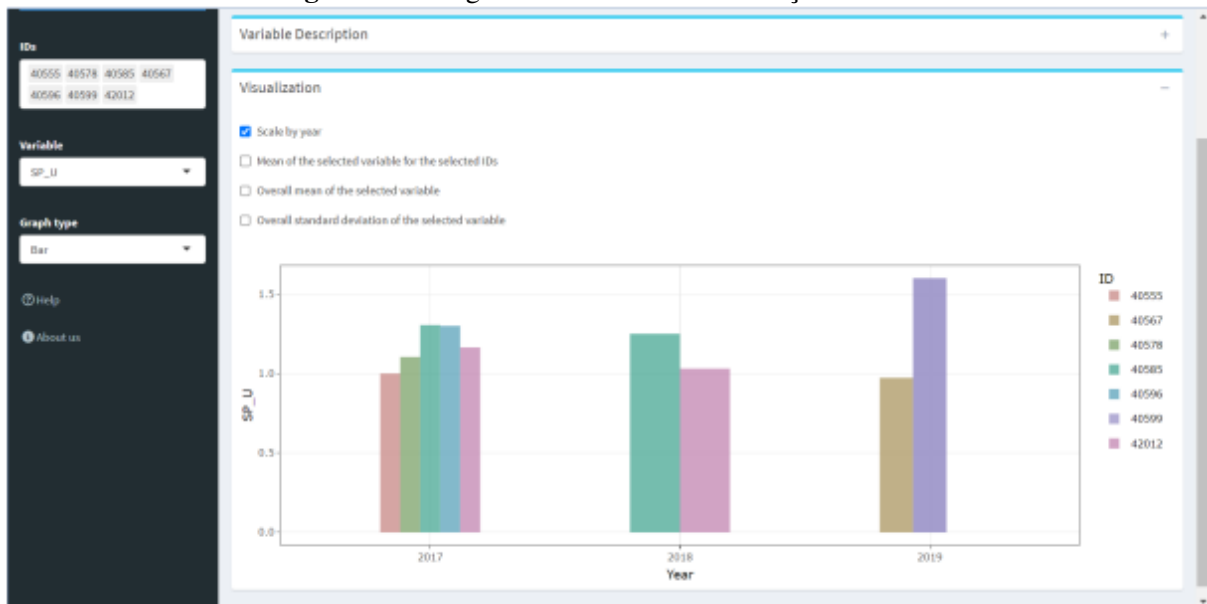
CAMPOS (2021) também trouxe uma visão diferente, que se refere aos dados abertos estarem disponíveis de forma textual, transformando assim uma visualização mais compreensível de um conjunto completo sobre o que os dados estão falando. Enquanto nesse projeto, com um acesso para diversas pessoas e produzindo uma plataforma compreensível para todos que acessarem serão transcritos esses dados textuais em um *dashboard* com visualização interativa. No projeto de CAMPOS (2021), também é citada a criação de uma plataforma, porém mais abrangente, onde é necessário fazer outra pesquisa dentro da própria plataforma para achar o que deseja. Diferentemente desta aplicação, onde o usuário escolherá a sua própria base de dados diretamente do site de dados abertos da Câmara e receberá a visualização referente à sua própria pesquisa.

2.2.3 Coletando e analisando informações do Facebook com R

Segundo FIGUEIREDO et al. (2018), o que ajudou na análise e tratamento de dados no R em seu trabalho, foi o desenvolvimento de uma ferramenta chamada DataMy que coleta, analisa e trata os dados de comentários de Facebook, assim separando os comentários negativos e positivos. Foi possível constatar que a técnica de mineração de texto em conjunto com a análise de sentimentos usada no tratamento das informações, permite a identificação de valores nas mensagens publicadas de forma ágil, interpretando-as e classificando-as por sua polaridade. Por meio disso, é possível ter uma ideia de como separar e interpretar os dados após o tratamento.

2.2.4 Visualização de dados de pessoas com doença de Parkinson

Em SANTOS (2023) foi desenvolvido um *dashboard* para visualizar os dados de pessoas com a doença de Parkinson. Foi implementado em linguagem R com apoio da biblioteca shiny para construção da aplicação web. Com esse trabalho surgiu uma ideia de como demonstrar os dados da CD, seguindo um padrão que melhor se encaixe nos requisitos de nosso cliente. Também uma outra forma de coleta de dados com R, utilizando para uma coleta mais ampla e um tratamento melhor das informações coletadas gerando um *dashboard*. A ideia de SANTOS (2023) trouxe um pensamento importante na forma de visualização de dados, e como demonstrar esses dados para o cliente final. Contribuindo também com a escolha de gráficos, posicionamento e legendas no formato em que são apresentadas, como na Figura 2.5.

Figura 2.5 Imagem relacionada a visualização de dados

Fonte: Figura retirada de SANTOS (2023).

2.3 Comparativo entre os trabalhos

A Figura 2.6 traz algumas diferenças entre os trabalhos correlatos apresentados na revisão de literatura. É possível compreender como foram utilizados os trabalhos correlatos e também as diferenças entre os trabalhos que foram apresentados.

NOGUEIRA (2021) utiliza um dashboard desenvolvido em Streamlit com gráficos diferentes, além de outra plataforma para apresentação dos dados. Já neste trabalho é desenvolvido em uma plataforma chamada Data Studio, para confecção da interface e de como os dados serão apresentados ao usuário.

Enquanto CAMPOS (2021) trouxe a implementação de um banco de dados chamado MongoDB em seu trabalho, este trabalho trouxe a plataforma do Google Sheets para armazenar dados.

FIGUEIREDO et al. (2018) apresenta uma visualização mais simples apenas com alguns cards, diferente deste, que trouxe diversos tipos de visualização com gráficos e tabelas.

E por fim, SANTOS (2023) trouxe pouco detalhamento e filtros em seu trabalho, e neste, é destacado a utilização de diversos filtros que podem ser utilizados de diversas formas.

Figura 2.6 Comparativo entre os trabalhos

Referências	Objetivo de Estudo	Funções existentes	Funções diferenciais deste trabalho
Nogueira (2021)	Análise de dados com R	Dashboard com o Streamlit.	Dashboard com o Data Studio.
Campos (2021)	Análise de dados abertos da Câmara	O banco de dados MongoDB.	A plataforma Google Sheets.
Figueiredo et al. (2018)	Coletando e analisando informações do Facebook com R	Visualizações com cards, sem outras formas de visualizações variadas.	Múltiplas formas de visualização de dados como gráficos e tabelas.
Santos (2023)	Visualização de dados de pessoas com doença de Parkinson	Falta de detalhamento de filtros.	Filtros mais detalhados e variados.

Fonte: elaborado pelo autor.

2.4 Tecnologias

Nesta seção, são descritas as tecnologias utilizadas para o desenvolvimento do trabalho. Como a linguagem R, o Google Sheets e o Data Studio.

2.4.1 Linguagem R

Como abrange BITTENCOURT et al. (2019) a linguagem R se torna um poderoso aliado para utilizar em análise e tratamento de dados onde os pacotes e bibliotecas em R disponíveis na internet, se mostraram bastante úteis para o uso de manipulação de dados. Tendo em vista que pacotes e bibliotecas são ferramentas que ajudam em tarefas específicas, no caso desse projeto, auxiliando em tratamento de dados.

A linguagem R utilizada juntamente com o *software* do RStudio facilitou e auxiliou na criação de códigos em R que tem total contribuição no tratamento dos dados brutos, oriundos da CD. Com essa linguagem de programação, utilizando o conjunto de bibliotecas do Tidyverse (2021), foi produzido um tratamento de dados específico para construção do *dashboard* final.

Com isso, foi produzido um *Script* de tratamento de dados e logo depois, por meio da linguagem R, enviados os dados tratados para o Google Sheets, onde os dados foram armazenados.

2.4.2 Google Sheets

O *Google Sheets* ou Planilhas do Google como é explicado por APRIANI; AAN; SAPUTRA (2022) é uma ferramenta web que proporciona aos usuários criar, editar, alterar dados e compartilhar planilhas

Em conjunto com essa pesquisa, foi mencionado por APRIANI; AAN; SAPUTRA (2022) que o *Google Sheets*, também pode ser usado como banco de dados, fornecendo assim, uma conexão melhor com o Data Studio, sendo uma forma mais simples de se usar um banco de dados e mais compreensível, funcionando apenas com uma conta do Google e acesso a internet.

2.4.3 Data Studio

O Data Studio foi utilizado para o desenvolvimento do *dashboard*, tendo em mente as funcionalidades em que ele agrega. Por ser uma plataforma do grupo Google, não é necessário hospedar o site, como é explicado por OLIVEIRA; LICHTNOW (2022).

JUNIOR (2021) traz essa perspectiva sobre o Data Studio:

O Data Studio é uma ferramenta online e gratuita que transforma bases de dados em relatórios e painéis informativos de fácil e compreensível leitura, com uma gama de personalizações e com possibilidade de compartilhamento.

É importante destacar, que, o desenvolvimento da interface deve ser produzido, e a conexão de dados também deve ser feita, para assim, gerar os gráficos que serão apresentados nesse *dashboard*.

O próximo capítulo apresenta a metodologia deste trabalho, e os passos para adotados alcançar seus objetivos.

3

Metodologia

A pesquisa está classificada como exploratória qualitativa, pois ela tem como objetivo explorar e compreender a natureza dos fenômenos envolvidos no problema proposto, utilizando métodos qualitativos na coleta de dados e análise.

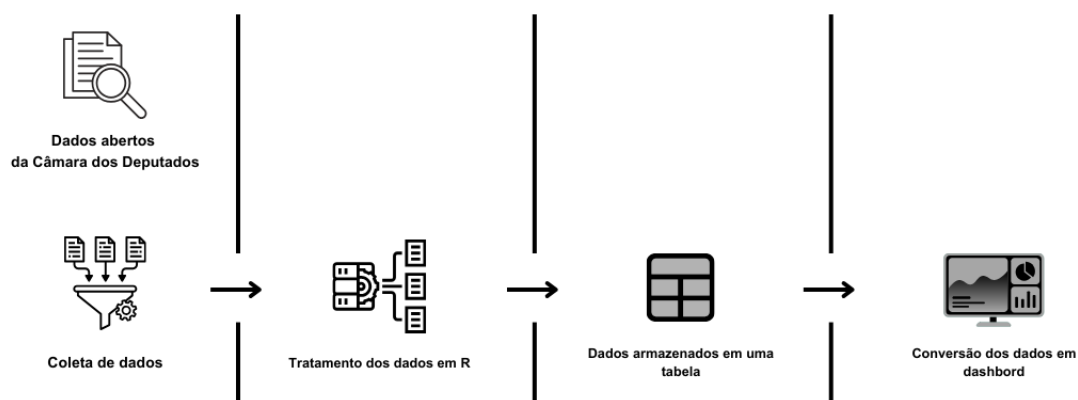
Pesquisas bibliográficas foram utilizadas para analisar e interpretar as informações existentes sobre o tema em estudo, possibilitando o desenvolvimento do projeto com base em conhecimentos científicos e teóricos consolidados.

Segundo SEMERANO (2024) o procedimento de coleta e análise tem como foco atender as metodologias como o SCRUM, que foi adotada para idealizar e acompanhar esse trabalho.

A metodologia SCRUM foi escolhida pela versatilidade de uso onde retroceder em partes da pesquisa pode ser fundamental para a sua conclusão. Diferente da metodologia do Kanban, que está mais focada em realizar processos do que realizar mais trabalho porém em um menor tempo como o SCRUM.

Será demonstrado aqui o processo em que esses dados foram utilizados até chegar na versão final da pesquisa, conforme a Figura 3.1.

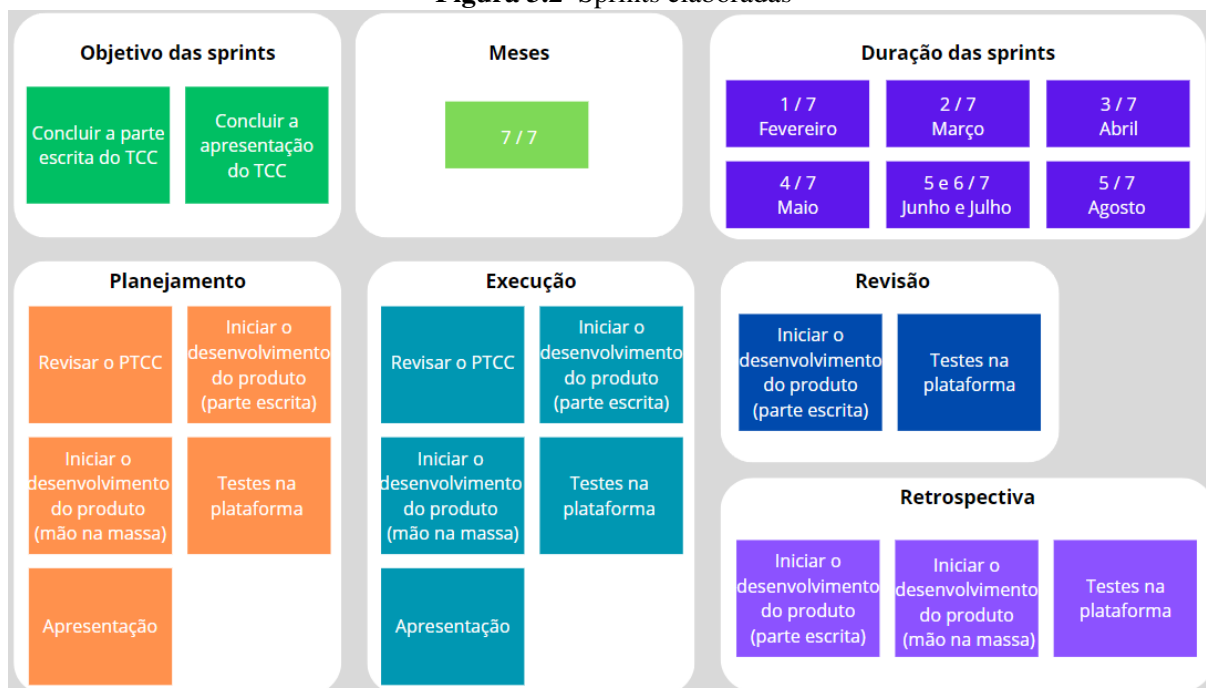
Figura 3.1 Ilustração do processo de dados



A Figura 3.1 traz os processos que foram utilizados durante o desenvolvimento dessa pesquisa demonstrando como foi pensado e estruturado o passo a passo até chegar ao resultado final.

3.1 Sprints SCRUM

Figura 3.2 Sprints elaboradas



Fonte: *print screen* realizada pelo autor.

A Figura 3.2 foi desenvolvida para nos auxiliar durante o processo de conclusão deste trabalho, onde foi separado em partes.

Primeiramente foram selecionados os objetivos das sprints, seguindo da duração em meses que foram necessárias para concluí-las e por fim o tempo de duração de cada uma.

No planejamento, foram definidas as etapas principais para cada sprint, estabelecendo os momentos de revisão, desenvolvimento da parte escrita e testes na plataforma. Esse planejamento também determinou as atividades que seriam divididas entre a execução prática do produto e o progresso da parte escrita, assegurando que cada tarefa fosse cumprida dentro do prazo.

Durante a execução, houve uma separação clara entre as atividades de desenvolvimento, revisões e testes. Duas sprints, em particular, exigiram atenção especial devido à sua complexidade e ao tempo necessário para implementar mudanças significativas. Esses ajustes foram essenciais para garantir a qualidade tanto da parte escrita quanto do produto final.

Um ponto que merece destaque é a duração das sprints, onde o volume de trabalho precisou ser estendido para dois meses na quarta sprint. Essa extensão foi necessária porque, nesse ponto, houve a necessidade de retroceder e corrigir partes do projeto, principalmente relacionadas à parte escrita, para alinhar o desenvolvimento com os padrões esperados.

Além disso, a revisão e os testes na plataforma foram processos contínuos, que acompanharam as fases de execução, permitindo ajustes em tempo real. Por fim, algumas sprints exigiram um retorno aos estágios anteriores durante o processo, garantindo que todas as correções fossem feitas de forma adequada antes da conclusão.

O presente trabalho consistiu em desenvolver um projeto de tratamento e análise de dados da CD com o auxílio de uma plataforma *web* para a criação de um *dashboard* personalizado para uma empresa de gestão de software. Nesse projeto foi utilizado a linguagem R, o conjunto de bibliotecas do *Tidyverse* e a plataforma Data Studio.

4

Requisitos do sistema

Neste capítulo serão apresentados os requisitos do sistema, que contempla os requisitos funcionais e não funcionais e também os casos de uso do sistema produzido aqui, a partir da plataforma do Data Studio.

4.1 Requisitos Funcionais

- RF1. Conexão com Fontes de Dados: Permite conectar-se a diversas fontes de dados, como Google Sheets, entre outras.
- RF2. Criação de Relatórios Personalizados: Possibilita a criação de relatórios e dashboards personalizados com gráficos, tabelas e outros elementos de visualização.
- RF3. Filtros e Segmentações Dinâmicas: Oferece a capacidade de aplicar filtros e segmentações dinâmicas para refinar os dados exibidos em relatórios.
- RF4. Compartilhamento e Colaboração: Facilita o compartilhamento de relatórios com outros usuários e permite a edição colaborativa em tempo real.
- RF5. Atualização Automática de Dados: Os relatórios podem ser configurados para atualização automática com base nas mudanças nas fontes de dados.
- RF6. Controle de Acesso: Permite configurar permissões de acesso para diferentes usuários, definindo quem pode visualizar ou editar os relatórios.
- RF7. Interatividade: Controle de data e gráficos clicáveis que permitem a exploração de dados detalhados.
- RF8. Modelagem de Dados: Suporte para cálculos personalizados, transformação e modelagem de dados diretamente dentro da plataforma.

4.2 Requisitos não Funcionais

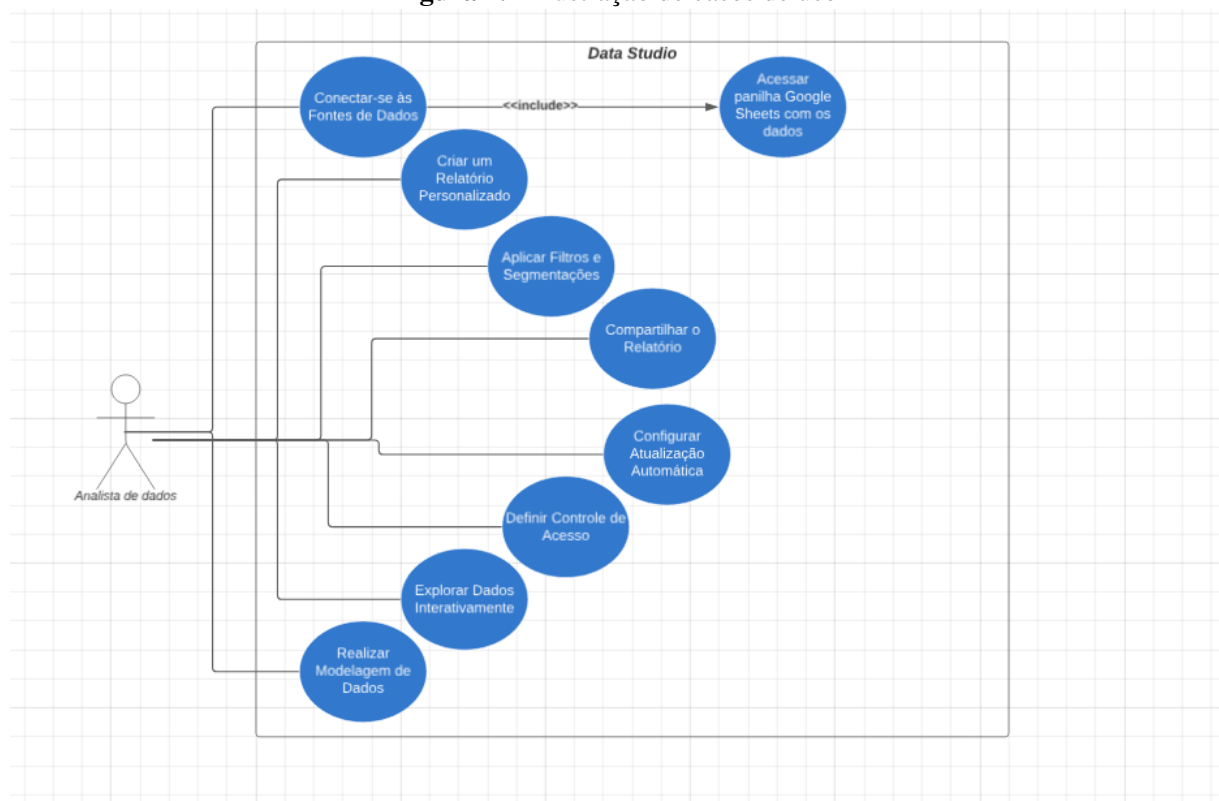
- RNF1. Usabilidade: Interface intuitiva e fácil de usar, com uma curva de aprendizado baixa para iniciantes.

- RNF2. Desempenho: Resposta rápida ao carregar e renderizar grandes volumes de dados, dependendo da performance das bases de dados conectadas.
- RNF3. Escalabilidade: Capaz de lidar com grandes conjuntos de dados e múltiplas fontes sem comprometer a performance.
- RNF4. Segurança: Implementa padrões de segurança robustos, incluindo autenticação Google, criptografia de dados e controle de acesso baseado em permissões.
- RNF5. Compatibilidade: Compatível com navegadores modernos e dispositivos móveis, permitindo acesso e edição de relatórios de qualquer lugar.
- RNF6. Confiabilidade: Alta disponibilidade com baixo risco de falhas ou tempo de inatividade, garantindo que os relatórios estejam acessíveis o tempo todo.
- RNF7. Manutenção: Suporte contínuo e atualizações automáticas pela equipe do Google, melhorando a funcionalidade e corrigindo bugs regularmente.

4.3 Casos de Uso

Aqui será demonstrado o caso de uso idealizado para o nosso sistema.

Figura 4.1 Ilustração de casos de uso



Fonte: Elaborada pelo autor.

Conforme a Figura 4.1 é apresentado o diagrama de caso de uso pela visão do usuário, que é a empresa de gestão de software, onde o mesmo irá acessar a plataforma através de um link e poderá: acessar o site da Câmara dos Deputados, selecionar um período, interagir com os gráficos, filtra os dados e redefinir as informações.

Este diagrama representa um caso de uso para o Data Studio, onde o ator principal é o Analista de Dados. O fluxo demonstra as ações que o analista pode realizar no Data Studio para criar relatórios personalizados e interativos. Abaixo está a descrição das principais ações identificadas no diagrama:

- Conectar-se às Fontes de Dados: O analista de dados se conecta a diversas fontes de dados, como planilhas, para obter as informações necessárias.
- Criar um Relatório Personalizado: Utilizando os dados conectados, o analista constrói relatórios personalizados que atendem às necessidades específicas do cliente.
- Aplicar Filtros e Segmentações: Para refinar os dados exibidos nos relatórios, o analista aplica filtros e segmentações, melhorando as informações apresentadas.
- Explorar Dados Interativamente: O analista explora os dados de forma interativa dentro do relatório, realizando análises mais profundas.
- Acessar Planilha Google Sheets com os Dados: Como uma ação incluída, o analista acessa dados diretamente de uma planilha do Google Sheets, que é uma das fontes de dados integráveis no Data Studio.
- Realizar Modelagem de Dados: O analista pode reestruturar ou modelar os dados dentro do Data Studio para melhorar a análise ou apresentação dos resultados.
- Definir Controle de Acesso: Para garantir a segurança e privacidade das informações, o analista define controles de acesso, determinando quem pode visualizar ou editar o relatório.
- Configurar Atualização Automática: O analista configura atualizações automáticas dos dados para garantir que as informações estejam sempre atualizadas no relatório.
- Compartilhar o Relatório: Após a criação e configuração, o relatório pode ser compartilhado com outros usuários, permitindo que tomadores de decisão acessem as informações.

5

Desenvolvimento

Este capítulo está destinado ao desenvolvimento do projeto, aqui é apresentado como foi desenvolvido o projeto, etapa por etapa.

5.1 Coleta de dados

Na parte inicial do projeto juntamente com a empresa de gestão de software, foram designados os dados necessários e informativos que irão compor nosso *dashboard*. Os dados escolhidos são as proposições da CD com ênfase nos conteúdos de:

1. Ciência, Tecnologia e Inovação.
2. Finanças Públicas e Orçamento.
3. Indústria, Comércio e Serviços.
4. Meio Ambiente e Desenvolvimento Sustentável.
5. Trabalho e Emprego.
6. Turismo.

Estes dados são baixados no site da CD na página de dados abertos. Os dados são baixados separadamente de acordo com proposições, classificação temática e autores das proposições como é demonstrado na Figura 5.1, ou seja, são baixadas três bases de dados.

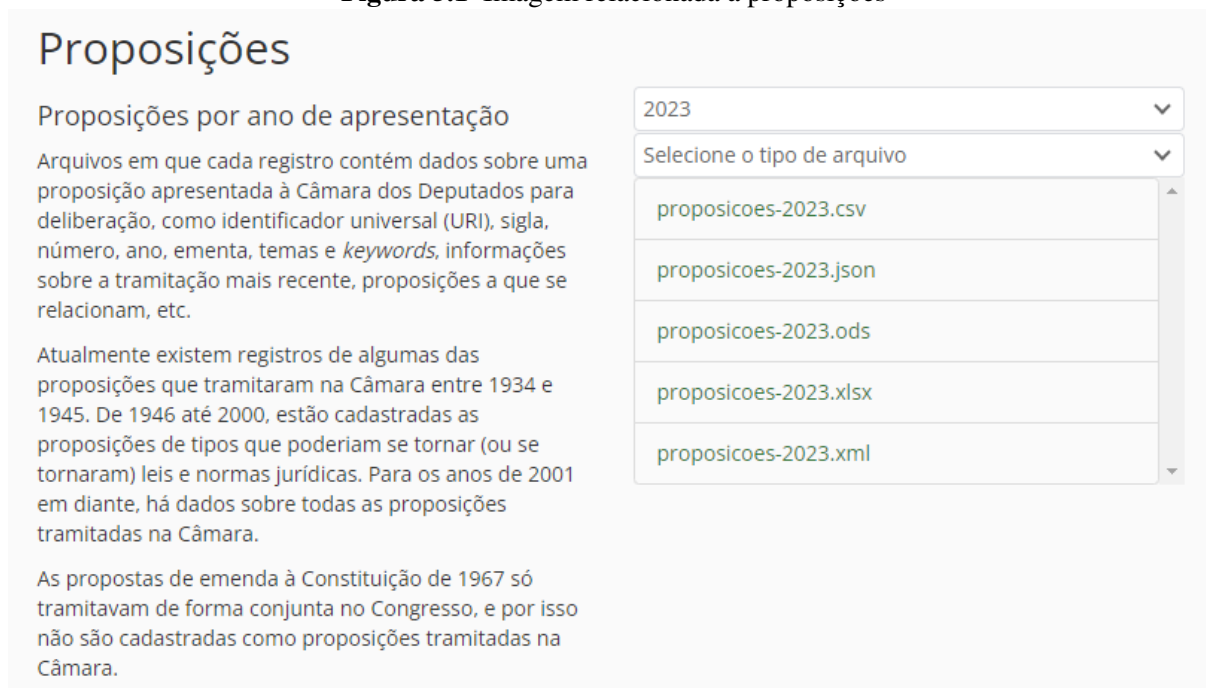
As bases de dados são baixadas em formato CSV, como é observado na Figura 5.1 isso garante com que a utilização desse formato auxilie para o tratamento desses dados.

Após o download dessas bases, são recebidos três arquivos em formato CSV, já no computador, eles são observados como é apresentado na Figura 5.2

5.2 Tratamento de dados

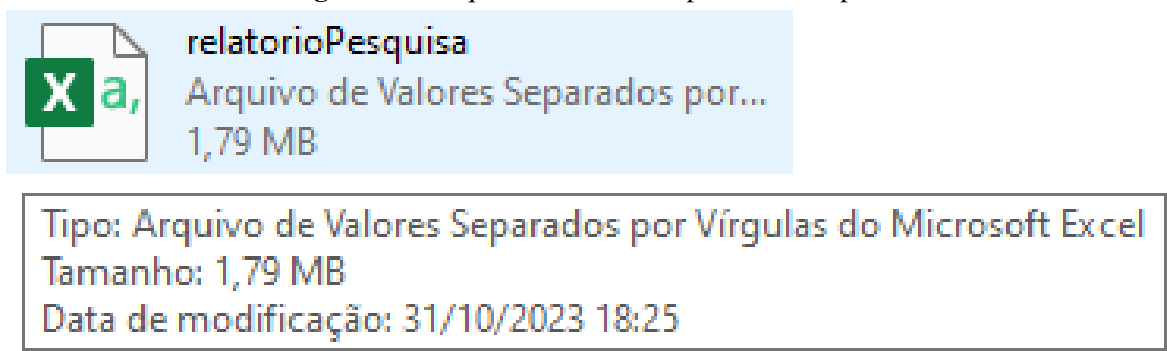
O tratamento de dados se inicia dentro do programa RStudio, onde são digitados os códigos na linguagem R que são usados para tratar os dados e permitir que sejam de compreensíveis para enviar para o banco de dados, que nessa pesquisa foi o *Google Sheets*.

Figura 5.1 Imagem relacionada a proposições



Fonte: Câmara dos Deputados (2021). *print screen* realizada pelo autor.

Figura 5.2 Arquivo .csv em uma pasta no computador



Fonte: *print screen* realizada pelo autor.

Inicialmente é necessário rodar o pacote de bibliotecas do Tidyverse (2021) que será de grande ajuda com suas funções já pré construídas. Utilizando as bases baixadas da CD precisamos conseguir ler, separar e organizar os dados em tabelas, isso individualmente para cada uma das três bases, após isso essas bases são unidas com o *left join* e são armazenadas em apenas um **data frame**, que é uma estrutura de dados muito parecida com uma tabela.

Podemos observar o código que foi utilizado para atingir o resultado na Figura 5.3.

Após essa primeira fase do tratamento de dados é necessário modificar a estrutura de alguns dados presentes na nossa base, isso serve para que a leitura individual dos dados possam ser compreendidas e analisadas.

A construção desse tratamento de dados foi de acordo com os temas designados pela empresa de gestão de software, resultando no código presente na Figura 5.4.

Figura 5.3 Código de tratamento de dados parte 1

```

purrr::map(c('rvest', 'stringr', 'dplyr', 'magrittr', 'lubridate', 'purrr', 'httr', 'tidyverse', 'readr'), library, character.only = TRUE)

temas2023 <- 'https://dadosabertos.camara.leg.br/arquivos/proposicoesTemas/csv/proposicoesTemas-2023.csv' %>%
  read_csv2(locale = locale(encoding = "UTF-8")) %>%
  mutate(id = str_remove(uriProposicao, 'https\\:\\\\{2}dadosabertos\\.camara\\.leg\\.br\\生api\\生v2\\生proposicoes\\生')) %>%
  relocate(id, .before = uriProposicao)

prop2023 <- file.choose() %>% read_csv2(locale = locale(encoding = "UTF-8"))

Autor2023 <- 'https://dadosabertos.camara.leg.br/arquivos/proposicoesAutores/csv/proposicoesAutores-2023.csv' %>%
  read_csv2(locale = locale(encoding = "UTF-8"))

prop2023 %>%
  mutate(id = as.character(id)) %>%
  left_join(temas2023, by = 'id') -> propTemas2023

Autor2023 %>%
  rename(id = idProposicao) %>%
  mutate(id = as.character(id)) %>%
  left_join(propTemas2023, by = 'id') -> propAutorTemas2023

```

Fonte: *print screen* realizada pelo autor.

Figura 5.4 Código de tratamento de dados parte 2

```

ArquivoFinal <- propAutorTemas2023 %>%
  select(id, dataApresentacao, descricaoTipo, ementa, tema, nomeAutor, siglaPartidoAutor, siglaUFAutor, siglaTipo.x, ultimoStatus_apreciacao)

ArquivoFinalV2 <- ArquivoFinal %>%
  filter(tema == c("Ciência, Tecnologia e Inovação", "Turismo", "Indústria, Comércio e Serviços", "Trabalho e Emprego", "Meio Ambiente e Desenvolvimento Sustentável", "Finanças Públicas e Orçamento" ))

ArquivoFinalV3 <- ArquivoFinalV2 %>%
  mutate(dataApresentacao = str_remove(dataApresentacao, '\\d{2}\\生.+'))

ArquivoFinalV4 <- ArquivoFinalV3 %>%
  mutate(dataApresentacao = ymd(dataApresentacao))

```

Fonte: *print screen* realizada pelo autor.

Finalizando a parte do tratamento de dados, agora podemos enviar a base de dados para

o nosso banco de dados que será o *Google Sheets*.

5.3 Banco de dados

Presente no banco de dados, inicialmente é recebido a base de dados tratada a partir do RStudio, que será enviada com base no código da Figura 5.5, que demonstra a *url* onde está a base de dados, um autenticador de email e por fim, envia a base de dados para aba "prop" no *Google Sheets* os dados que estão presentes na variável "prop".

Figura 5.5 Código de envio para o *Google Sheets*

```
ss <- 'https://docs.google.com/spreadsheets/d/1CTtaMIm1iLgxuddB97wBkemKhE_38kHmzFRoPfkf4_M/edit#gid=0'

googlesheets4::gs4_auth(email = 'email@gmail.com')
googlesheets4::sheet_write(ss = ss, sheet = 'prop', data = prop)
```

Fonte: Elaborada pelo autor.

Assim, obtemos o banco em perfeito estado, já separado com as seguintes tabelas:

1. Id
2. Data
3. Título
4. Ementa
5. Tema
6. Autor
7. Partido
8. UF
9. Tipo
10. Situação
11. Link

Dessa forma, chegamos na parte de utilizar esse banco de dados no Data Studio para a criação do nosso *dashboard*.

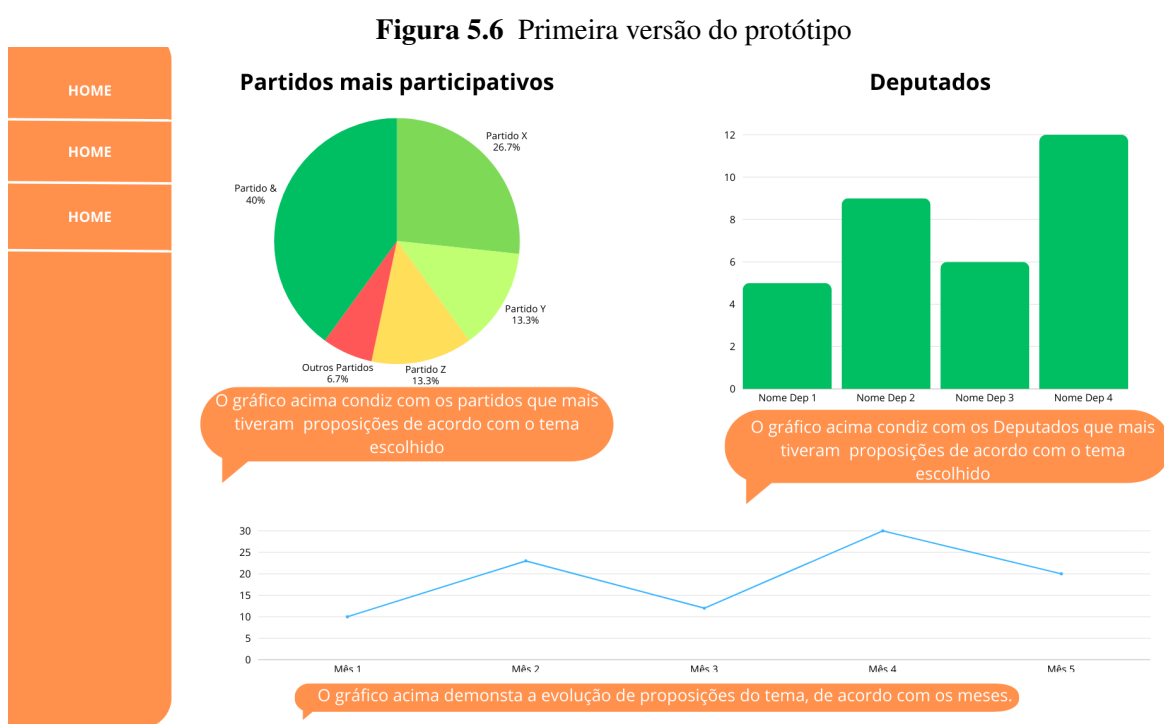
5.4 Prototipação

Antes de desenvolver o *dashboard* em si, primeiramente é necessário construir protótipos, e averiguar se estão conforme o desejo da empresa que vai utilizar. Para isso, com dados fictícios, foram construídas duas versões usando plataformas diferentes.

5.4.1 Protótipo versão 1

A primeira versão do nosso protótipo foi desenvolvida no Canva, um site que permite editar e criar imagens e apresentações. Partindo desse ponto, foi criado a primeira versão, com ideias simples, como personalizar o *dashboard* com cores ligadas a empresa e a CD.

Foi aperfeiçoada a ideia de que, os gráficos precisam fazer sentido, pois eles estão analisando e transcrevendo dados textuais para um formato visual. Baseado nessas ideias iniciais, foi produzido a primeira versão do *dashboard* que é essa presente na Figura 5.6



Fonte: Elaborada pelo autor.

A Figura 5.6 traz um *dashboard* simples com valores fictícios e três gráficos, sendo um de pizza, um de barras e um de linhas, que foram produzidos a partir do Canva.

5.4.2 Protótipo versão 2

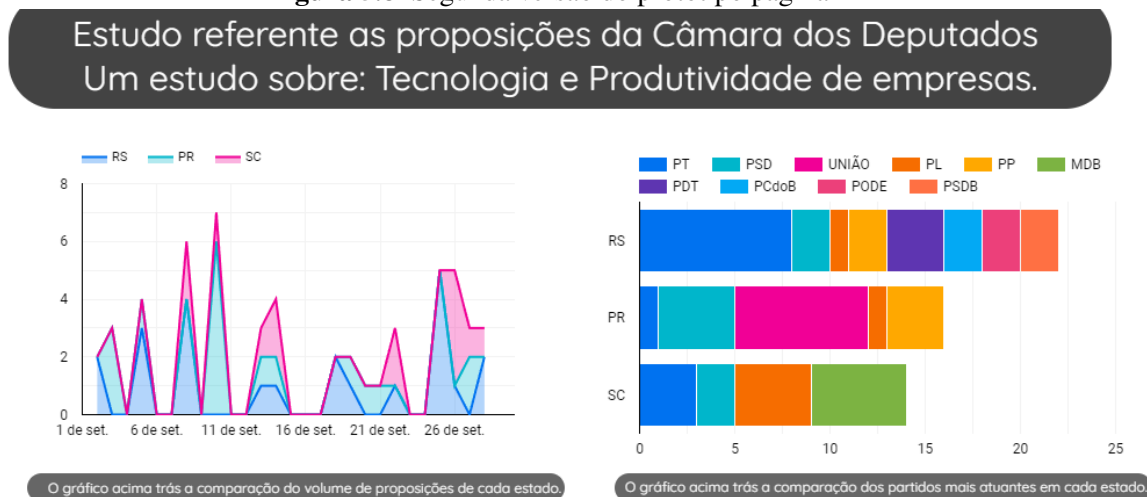
A segunda versão do protótipo foi desenvolvida a partir do Data Studio, como uma forma de compreender a plataforma que foi utilizada e as funções que ela oferece. Mantendo a ideia de usar dados fictícios e apresentar a plataforma para a empresa de gestão de software, foi possível entender quais seriam seus pedidos, modificações, dúvidas e curiosidades sobre o uso e a distribuição dos dados no *dashboard*.

Nessa segunda versão de prototipação, foi acolhida a ideia de páginas, demonstrando diferentes tipos de dados em cada uma delas, como é possível observar na Figura 5.7 e também na Figura 5.8, sendo respectivamente, a página 1 e 2 do segundo protótipo.

Figura 5.7 Segunda versão do protótipo página 1

Fonte: Elaborada pelo autor.

Na Figura 5.7 está a primeira página do segundo protótipo produzido deste trabalho, onde é demonstrado um gráfico de pizza, um gráfico de barras e um controle de período no *dashboard*.

Figura 5.8 Segunda versão do protótipo página 2

Fonte: Elaborada pelo autor.

Na Figura 5.8 está a segunda página do segundo protótipo produzido deste trabalho, onde é demonstrado um gráfico de evolução e um gráfico de barras no *dashboard*.

A prototipação foi de grande utilidade e informação para auxiliar a melhoria em alguns aspectos importantes.

Após essa segunda versão, foi produzido um estudo e reconhecimento dos desejos do cliente para implementação na versão final. Esse estudo ajudou muito, pois, a presença no evento "Campus Party Brasília", onde foram produzidas diversas palestras sobre dados, auxiliou em um conhecimento para visualizar melhor os dados por meio de gráficos e como melhorar esteticamente o *dashboard* desse projeto.

5.5 Desenvolvimento no Data Studio

O desenvolvimento começa a partir da conexão dos dados presentes no *Google Sheets* com a plataforma do Data Studio. A partir desse ponto foi iniciado a criação da interface com o estilo de uma página apenas na versão final, onde foi produzida integração dos gráficos e tabelas no *dashboard*.

Com o conhecimento adquirido e com a ajuda dos estudos apresentados no capítulo 2, foi compreendido que é necessário ressaltar a forma de apresentar estes dados de uma maneira visual. Foi produzida a ideia de contar uma história, uma narrativa, seguindo um padrão de compreensão para que seu conteúdo seja entendido da melhor forma possível.

Com a ajuda da empresa de gestão de software, foi pensado na ideia estrutural de começar em um interesse maior sobre o que as proposições estão oferecendo e logo depois, utilizar os dados desmembrados a partir disso, assim foi idealizado uma estrutura com os seguintes passos:

1. O volume das proposições.
2. Situações das proposições.
3. Os temas mais utilizados nas proposições.
4. As proposições acompanhadas de filtros.
5. Temas de interesse de cada partido.
6. Evolução mensal dos tipos de proposições.
7. Volume de proposições de acordo com seus desmembramentos.
8. Filtros acompanhados de volume de proposições.
9. Evolução mensal dos temas das proposições.

A montagem do *dashboard* permite escolher os gráficos que serão úteis em ajudar a contar uma história, explicando e compreendendo o que esses dados estão demonstrando.

A criação de gráficos é bem simples, já que existem modelos para utilizar, porém é importante ressaltar o uso correto dos dados para não transmitir uma informação incongruente com o estudo que está apresentando.

Para o início da conexão dos dados, como métrica base foram utilizados os identificadores (ID) do nosso banco de dados a partir de cada proposição, fazendo assim, uma contagem de todos os identificadores que são diferentes, obtendo dessa forma resultados que tenham sentido.

Designado a métrica utilizada, foram configurados os outros gráficos, ajudando a transmitir esse conhecimento um pouco de cada vez. Também foi utilizado a mesma ideia da métrica para construção de tabelas, onde, por meio de filtros, o cliente pode escolher a informação mais relevante que ele deseja acessar. Para esse acesso, serão disponibilizados links clicáveis que estão diretamente ligados ao site da Portal da Câmara dos Deputados (2021), demonstrando assim, uma melhor veracidade dos dados.

6

Apresentação e Análise dos Resultados

O trabalho se encerra com a finalização do *dashboard* e sua entrega, que se encontra online no Data Studio. Nesse sentido, o projeto cumpre todas as definições produzidas anteriormente e é capaz de tratar dados brutos, compreender esses dados, e por fim, transformá-los em gráficos ou tabelas compreensíveis permitindo a sua visualização.

6.1 Banco de dados

A Figura 6.1 representa o resultado obtido após concluir o tratamento de dados e o envio da base para o banco de dados no Google Sheets. A partir desse banco de dados foi feita a conexão dele com o Data Studio.

Figura 6.1 Banco de Dados

	A	B	C	D	E	F	G	H	I	J	K
1	id	data	título	ementa	tema	autor	partido	UF	tipo	situacao	LinkV2
2	2416327	2023-12-21	Projeto de Lei	Obriga as empre Indústria, Comércio e Serviços		Marx Beltrão	PP	AL	PL	Proposição Suje	https://www.cam
3	2416327	2023-12-21	Projeto de Lei	Obriga as empre Trabalho e Emprego		Marx Beltrão	PP	AL	PL	Proposição Suje	https://www.cam
4	2416309	2023-12-21	Projeto de Lei	Inclui o índice ID Finanças Públicas e Orçamento		Padovani	UNIÃO	PR	PL	Proposição Suje	https://www.cam
5	2416052	2023-12-20	Projeto de Lei	Altera os arts. 2º Finanças Públicas e Orçamento		Amom Mandel	CIDADANIA	AM	PL	Proposição Suje	https://www.cam
6	2415798	2023-12-20	Projeto de Lei	Altera a Lei nº 7, Trabalho e Emprego		Célio Studart	PSD	CE	PL	Proposição Suje	https://www.cam
7	2415644	2023-12-20	Requerimento di	Requer informaç Meio Ambiente e Desenvolvimento Sustentável		Marcos Pollon	PL	MS	RIC	Indefinida	https://www.cam
8	2415611	2023-12-20	Projeto de Lei	Altera a Lei nº 9.5 Finanças Públicas e Orçamento		Filipe Barros	PL	PR	PL	Proposição Suje	https://www.cam
9	2415465	2023-12-20	Projeto de Lei C	Prevê medidas c Finanças Públicas e Orçamento		Gabriel Nunes	PSD	BA	PLP	Indefinida	https://www.cam
10	2415324	2023-12-19	Projeto de Lei	Altera a Lei nº 11 Indústria, Comércio e Serviços		Duda Ramos	MDB	RR	PL	Proposição Suje	https://www.cam
11	2415177	2023-12-19	Requerimento di	Requer informaç Indústria, Comércio e Serviços		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
12	2415177	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Gilson Marques	NOVO	SC	RIC	Indefinida	https://www.cam
13	2415175	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Gilson Marques	NOVO	SC	RIC	Indefinida	https://www.cam
14	2415173	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Adriana Ventura	NOVO	SP	RIC	Indefinida	https://www.cam
15	2415172	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Adriana Ventura	NOVO	SP	RIC	Indefinida	https://www.cam
16	2415167	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
17	2415166	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
18	2415165	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
19	2415163	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
20	2415162	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Gilson Marques	NOVO	SC	RIC	Indefinida	https://www.cam
21	2415158	2023-12-19	Requerimento di	Requer informaç Ciência, Tecnologia e Inovação		Gilson Marques	NOVO	SC	RIC	Indefinida	https://www.cam
22	2415157	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Adriana Ventura	NOVO	SP	RIC	Indefinida	https://www.cam
23	2415144	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Marcel van Hatt	NOVO	RS	RIC	Indefinida	https://www.cam
24	2415136	2023-12-19	Requerimento di	Requer informaç Finanças Públicas e Orçamento		Gilson Marques	NOVO	SC	RIC	Indefinida	https://www.cam

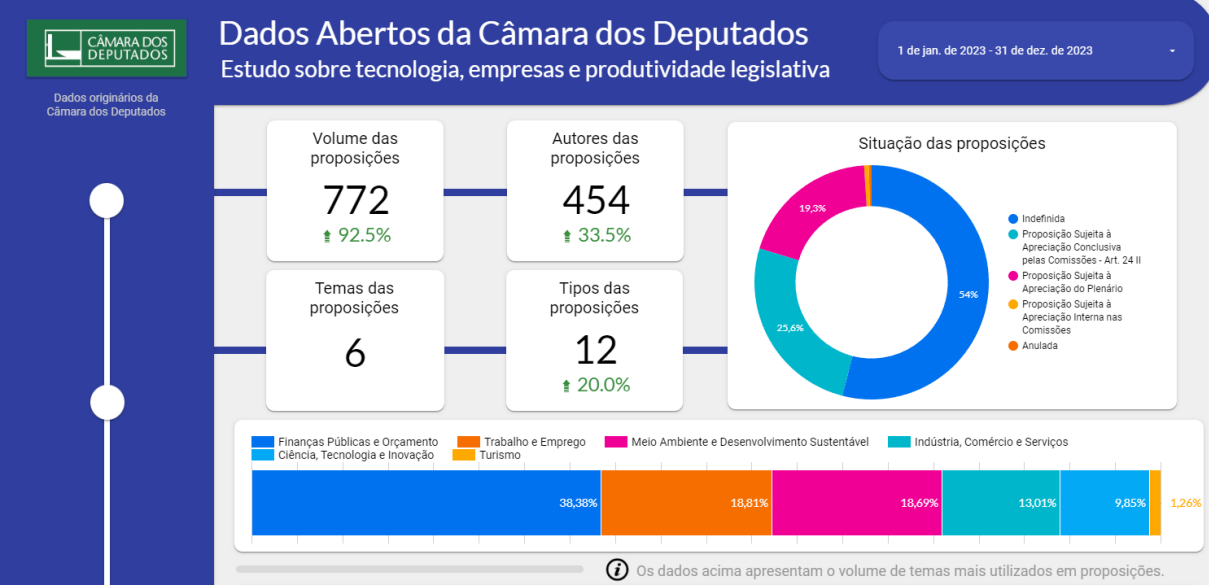
Fonte: Elaborada pelo autor.

A Figura 6.1 contém os dados da CD que estão tratados e armazenados no Google Sheets.

6.2 Dashboard no Data Studio

O desenvolvimento do *dashboard* no Data Studio, de acordo com a narrativa idealizada no capítulo 3 no item 5.5 resultou nas seguintes Telas: Figura 6.2, Figura 6.3, Figura 6.4, Figura 6.5, Figura 6.6, Figura 6.7 e Figura 6.8.

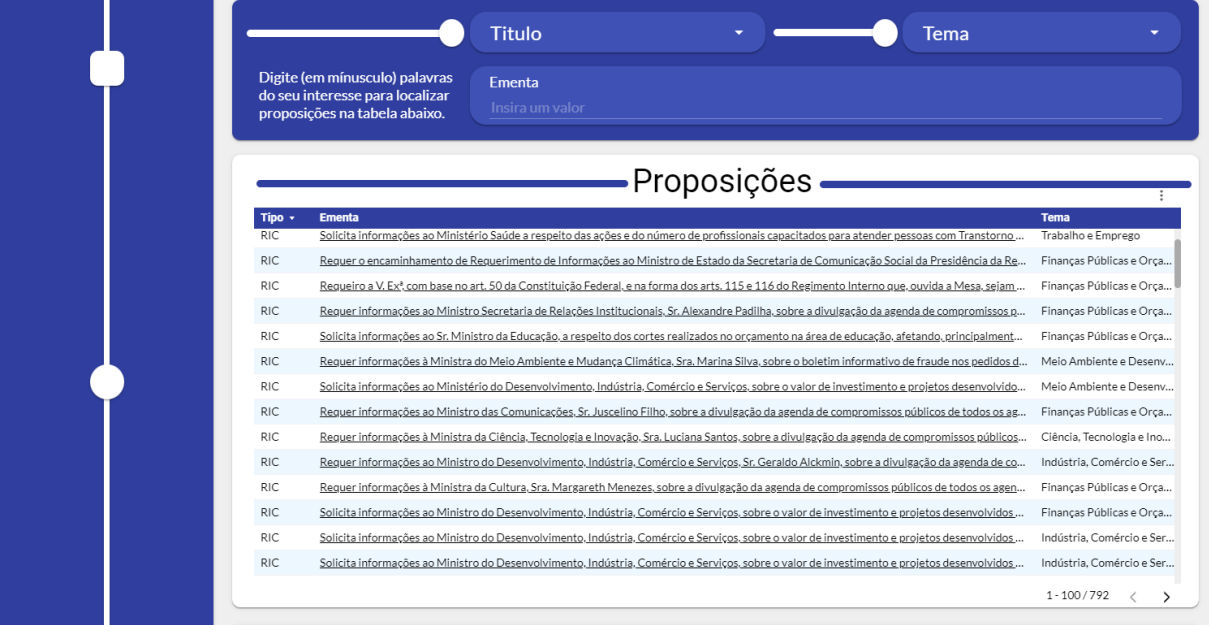
Figura 6.2 Tela 1 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

Trazendo a primeira visão do usuário ao *dashboard* a Figura 6.2 apresenta gráficos de barras e de rosquinha e quadros com informações sobre as proposições.

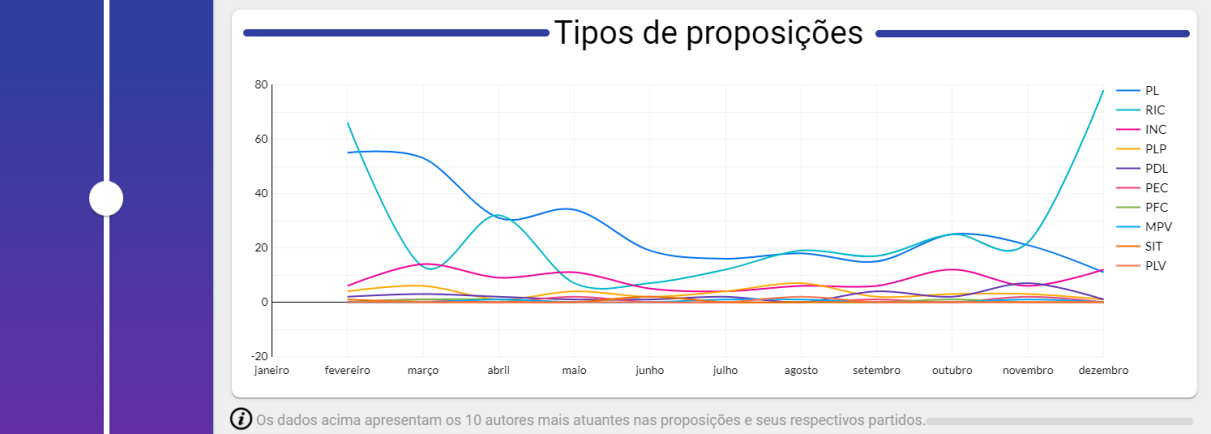
Figura 6.3 Tela 2 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

A Figura 6.3 apresenta alguns filtros que o usuário pode inserir e procurar por proposições específicas na tabela, a partir do título, tema ou ementa da proposição.

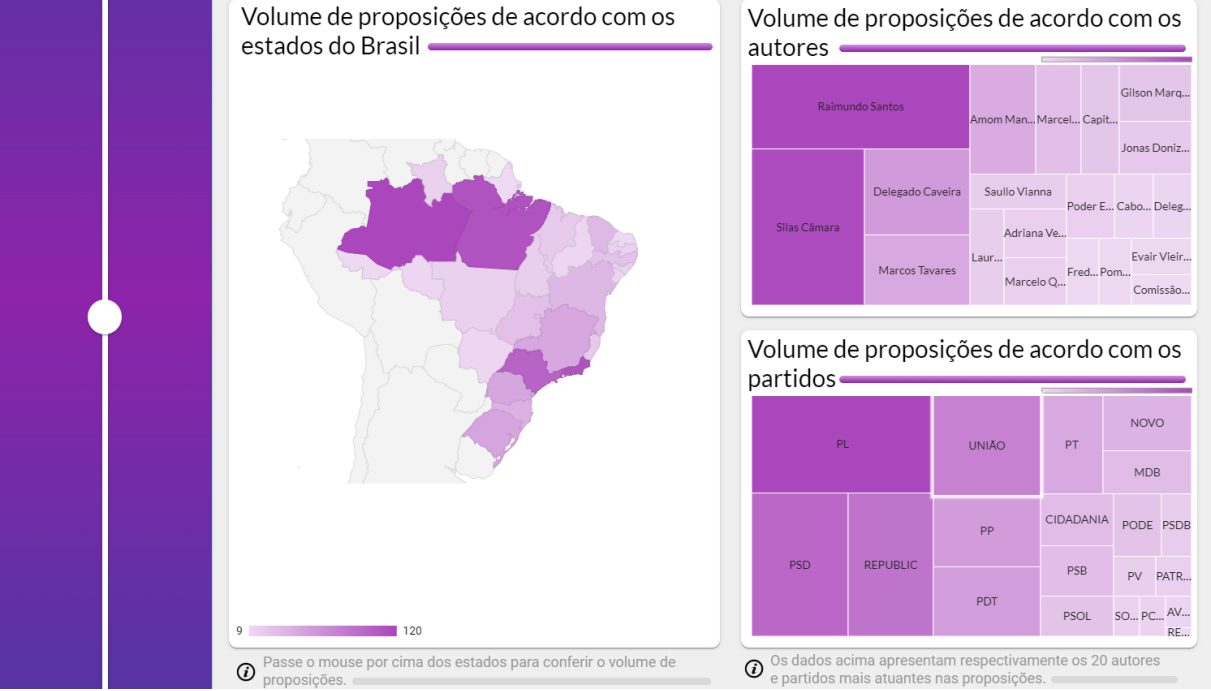
Figura 6.4 Tela 3 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

Apresentando um gráfico da linhas, a Figura 6.4 traz os tipos de proposições mais utilizados de acordo com os meses do ano de 2023.

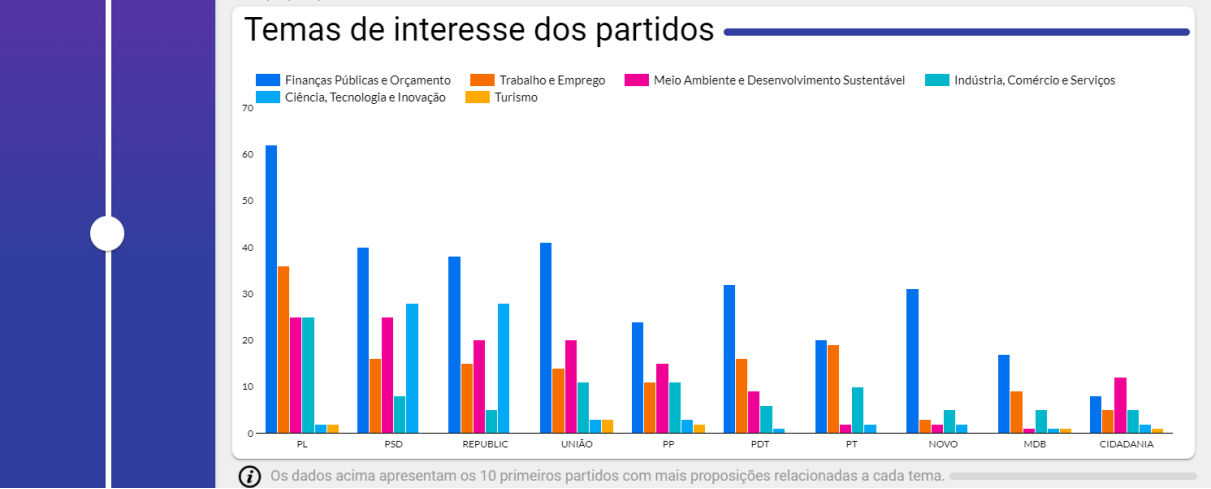
Figura 6.5 Tela 4 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

A Figura 6.5 envolve um mapa do Brasil, demonstrando de acordo com a intensidade da cor roxa, os estados que mais propuseram proposições. Também dois gráficos de *treemap* demonstrando os autores e partidos com mais proposições.

Figura 6.6 Tela 5 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

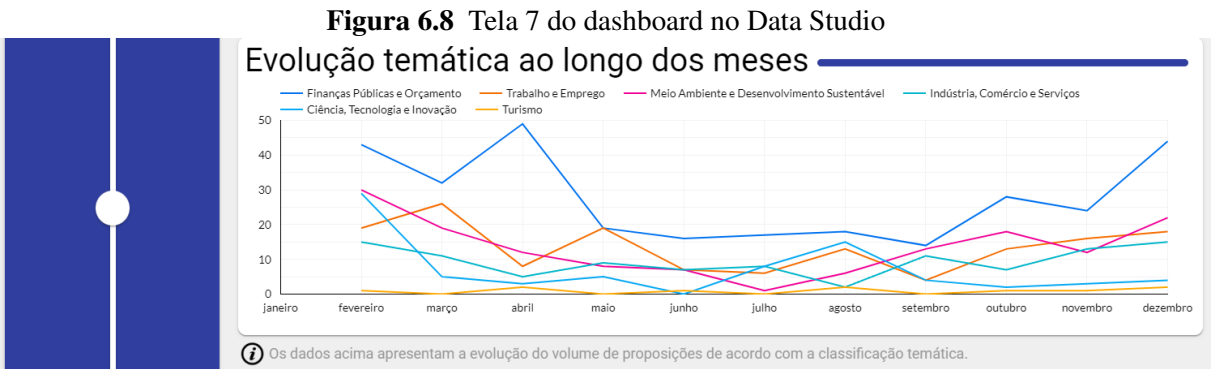
Já com um gráfico de barras a Figura 6.6 traz os temas de maior interesse dos dez primeiros partidos que mais apresentaram proposições a CD

Figura 6.7 Tela 6 do dashboard no Data Studio



Fonte: Elaborada pelo autor.

Também com mais possibilidades de filtros, a Figura 6.7 surpreende com uma tabela diferente que apresenta também, o volume de proposições de cada autor de acordo com o tema das proposições.

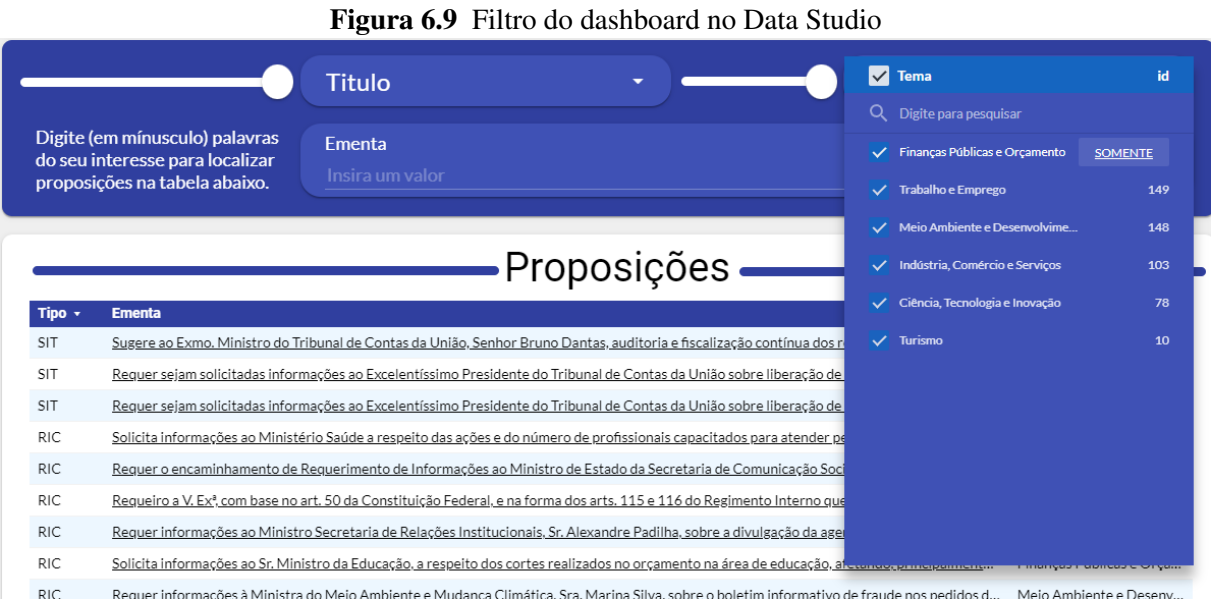


Fonte: Elaborada pelo autor.

A Figura 6.8 trouxe um gráfico de linhas de acordo com os temas mais discutidos durante os meses do ano de 2023.

6.3 Filtro de dados

No Data Studio, também está disponível os filtros de dados, que se encontram nas Figuras 6.3 e 6.7, possibilitando filtrar os dados, tanto pelo nome da coluna que estão disponíveis, ou apenas clicando nos gráficos, adequando assim, a interatividade, como é possível observar na Figura 6.9.



Fonte: Elaborada pelo autor.

6.4 Testes

Foram feitos testes de usabilidade no *dashboard* na plataforma do Data Studio, e estão sendo apresentados aqui os testes realizados.

Com base na ferramenta Questionnaire for User Interaction Satisfaction (QUIS), que foi criada por diversos pesquisadores do Laboratório de Interação Humano-Computador (HCIL) da Universidade de Maryland, foi utilizado a referência de FILARDI; TRAINA (2008) com métodos recomendados e aplicados pelos autores sobre a ferramenta, e foi produzido um questionário. Seu objetivo é medir a satisfação dos usuários, concentrando-se em aspectos específicos da interface entre o usuário e o computador.

Entendendo isso, foi decidido usar uma série de perguntas resumidas do QUIS, adaptada ao *dashboard* produzido no Data Studio. As perguntas elaboradas contêm uma série de avaliações que vão de 1 a 5, como 1, sendo a pior avaliação e 5 sendo a melhor avaliação. Foi utilizado a plataforma *Google Forms* para o auxílio com as perguntas e respostas do questionário presente na Figura 6.10.

Figura 6.10 Questionário do dashboard no Data Studio

Pergunta	Avaliação
Sobre a Tela	
A forma e tamanho das letras	1 (Difícil de ler) - 5 (Fácil de ler)
A organização das informações	1 (Confusa) - 5 (Clara)
Sobre os gráficos	
Quão agradável você achou a interface dos gráficos?	1 (Desagradável) - 5 (Agradável)
Quão simples foi entender os gráficos?	1 (Complexo) - 5 (Simples)
Quão simples foi utilizar os gráficos?	1 (Complexo) - 5 (Simples)
Quão completa foi a informação apresentada nos gráficos?	1 (Completa) - 5 (Incompleta)
Sobre os filtros	
Quão simples foi entender os filtros?	1 (Complexo) - 5 (Simples)
Quão simples foi utilizar os filtros?	1 (Complexo) - 5 (Simples)
Sobre a aprendizagem	
Quão fácil foi aprender a usar os gráficos?	1 (Difícil) - 5 (Fácil)
Quão fácil é lembrar como usar os gráficos?	1 (Difícil) - 5 (Fácil)
Sobre o sistema	
Quão rápida foi a resposta dos gráficos?	1 (Lenta) - 5 (Rápida)
Quão rápida foi a resposta dos filtros?	1 (Lenta) - 5 (Rápida)
Quão confiáveis foram os gráficos?	1 (Não confiável) - 5 (Confiável)

Fonte: Elaborada pelo autor.

6.4.1 Resultados dos testes de Usabilidade

A partir da pesquisa feita com 11 participantes, contando com estudantes de graduação e a empresa de gestão de software, foram coletadas respostas de forma anônima e recebidas as suas avaliações pelo *Google Forms* sendo possível entender a usabilidade do *dashboard* produzido.

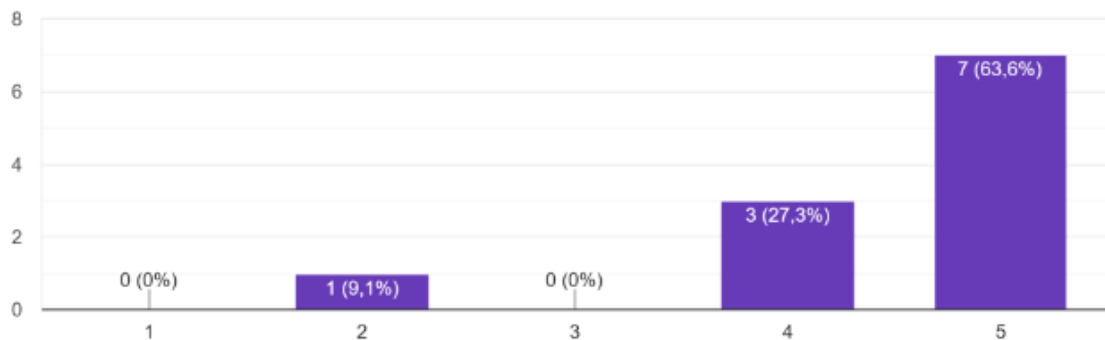
Sobre a tela do *dashboard* foram designadas duas questões, uma sobre o formato e tamanho das letras e outra sobre a organização das informações.

Percebeu-se na Figura 6.11 que dos 11 participantes da pesquisa 7 avaliaram bem o formato e tamanho das letras. Em uma escala de votações de 1 a 5 significa um ótimo resultado, mas entende-se que aumentar o tamanho das letras pode melhorar essa avaliação.

Figura 6.11 Resultados pergunta 1 Google Forms

A forma e tamanho das letras

11 respostas



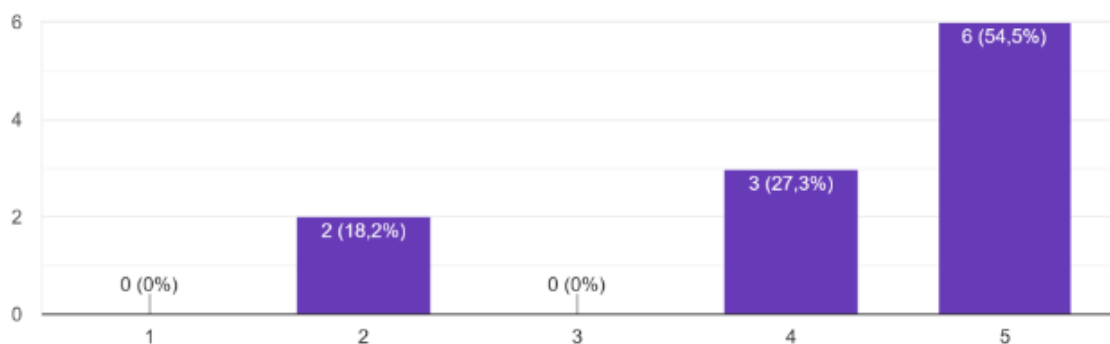
Fonte: Dados da pesquisa.

Na Figura 6.12 apenas 54% dos participantes avaliaram a organização das informações como muito boa. Então é necessário melhorar as informações e como elas são apresentadas.

Figura 6.12 Resultados pergunta 2 Google Forms

A organização das informações

11 respostas



Fonte: Dados da pesquisa.

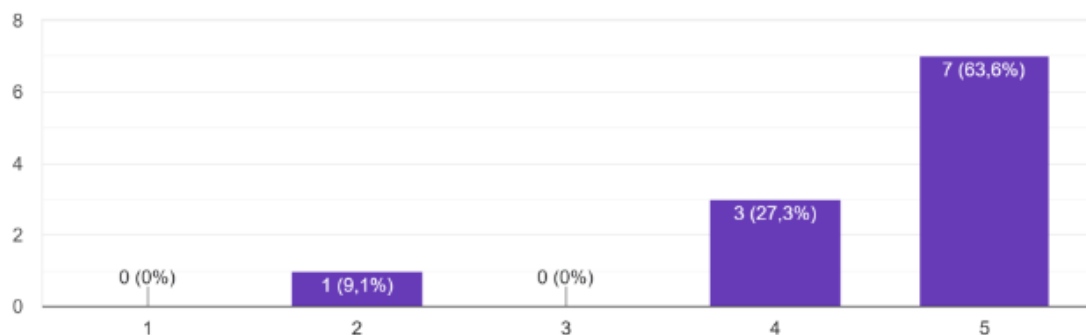
Em relação aos gráficos apresentados no *dashboard* foram elaboradas quatro perguntas para saber a opinião dos participantes da pesquisa, sobre a interface, o entendimento, e a utilização do *dashboard*.

Apenas 7 dos 11 participantes avaliaram a interface como agradável na Figura 6.13. Melhorar a interatividade poderia tornar a ainda mais agradável.

Figura 6.13 Resultados pergunta 3 Google Forms

Quão agradável você achou a interface dos gráficos?

11 respostas



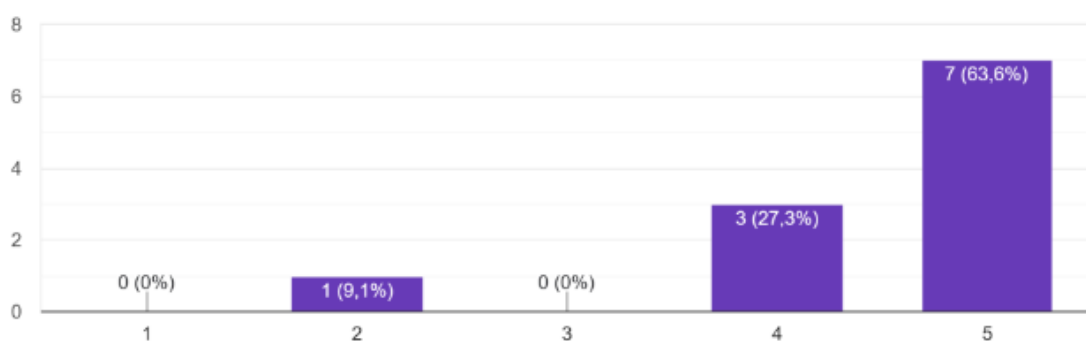
Fonte: Dados da pesquisa.

Sobre o fácil entendimento na Figura 6.14, 63,6% dos participantes avaliaram os gráficos como simples de entender. Escolhendo gráficos que tenham uma visualização mais simples, ou também mudando a coloração poderia tornar a compreensão mais fácil para os usuários.

Figura 6.14 Resultados pergunta 4 Google Forms

Quão simples foi entender os gráficos?

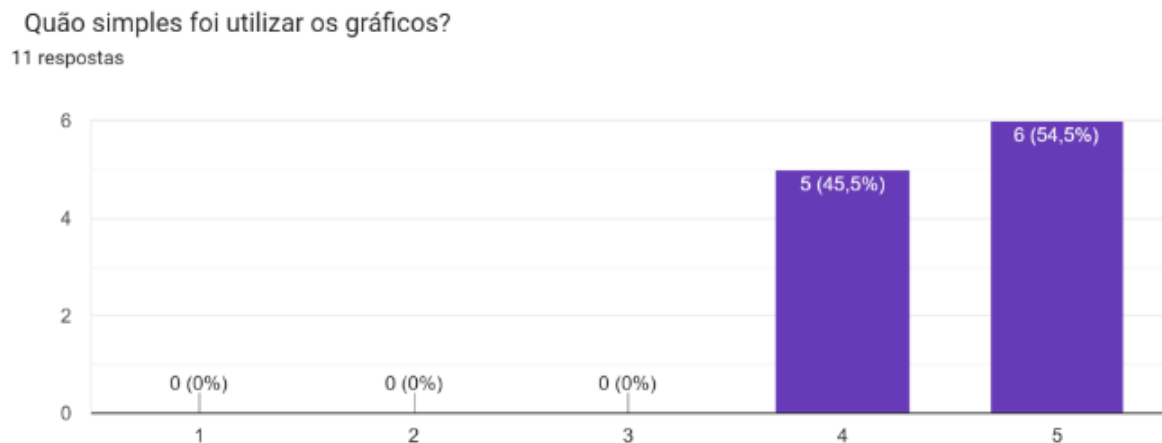
11 respostas



Fonte: Dados da pesquisa.

Sobre o uso dos gráficos na Figura 6.15, 6 dos participantes avaliaram como simples de utilizar e os outros 5 avaliaram com nota 4, compreendendo que a usabilidade mu dos gráficos é muito positiva.

Figura 6.15 Resultados pergunta 5 Google Forms



Fonte: Dados da pesquisa.

A informação dos gráficos como completa foi abordada por 54,5% dos participantes que avaliaram na Figura 6.16. É possível melhorar a forma como as informações são apresentadas, usando outros gráficos ou tabelas para comparação entre os dados.

Figura 6.16 Resultados pergunta 6 Google Forms

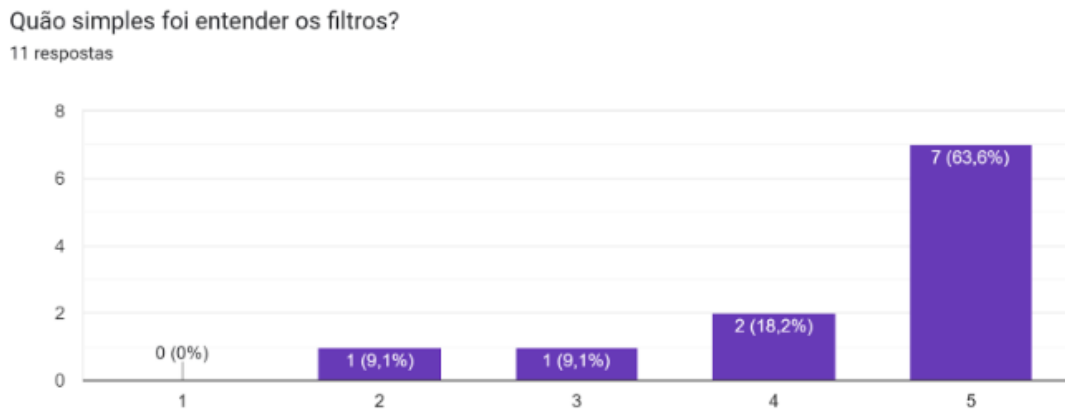


Fonte: Dados da pesquisa.

Sobre os filtros apresentados no *dashboard* foram elaboradas duas perguntas afim de compreender o uso dos filtros pelos participantes.

Em uma escala de 1 a 5 na Figura 6.17, apenas 7 dos participantes avaliaram os filtros como simples de entender. Melhorar a compreensão dos filtros é uma alternativa, assim deixando mais claro em como fazer a utilização deles.

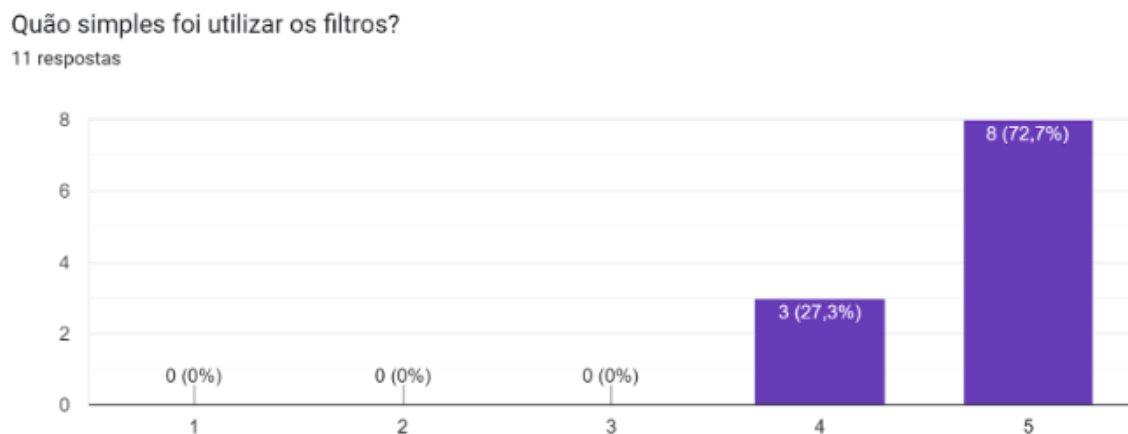
Figura 6.17 Resultados pergunta 7 Google Forms



Fonte: Dados da pesquisa.

Sobre o uso dos filtros na Figura 6.18, 72,7% dos participantes avaliaram os filtros como simples de utilizar enquanto os outros 27,3%, avaliaram com nota 4, trazendo resultados positivos quanto a utilização dos filtros.

Figura 6.18 Resultados pergunta 8 Google Forms



Fonte: Dados da pesquisa.

A outra seção de perguntas procurou avaliar quanto a aprendizagem dos participantes da pesquisa após o uso e entendimento do *dashboard* produzido.

Em uma escala de 1 a 5 na Figura 6.19, apenas 54,5% dos participantes avaliaram os filtros como fácil de aprender. Uma apresentação de como os filtros são utilizados poderia garantir maiores notas no aprendizado dos filtros.

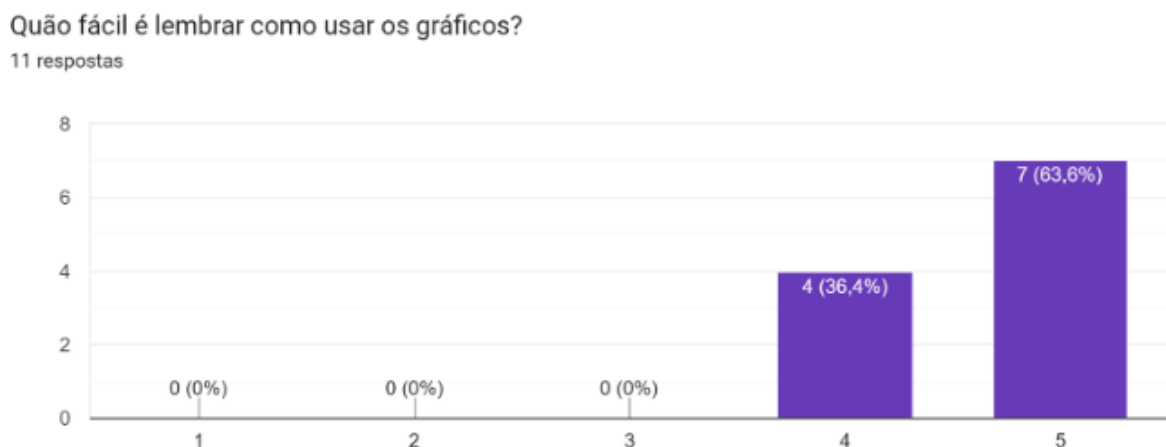
Figura 6.19 Resultados pergunta 9 Google Forms



Fonte: Dados da pesquisa.

Na Figura 6.20 7 participantes avaliaram os gráficos como fáceis de lembrar. Uma boa opção seria transformar os filtros e deixá-los mais aparentes no dashboard, mudando sua cor ou aparência.

Figura 6.20 Resultados pergunta 10 Google Forms



Fonte: Dados da pesquisa.

E por fim, gostaríamos de entender com a ajuda dos participantes as responsividade do sistema do Data Studio, de acordo com a interface produzida e como eles avaliam o uso da plataforma.

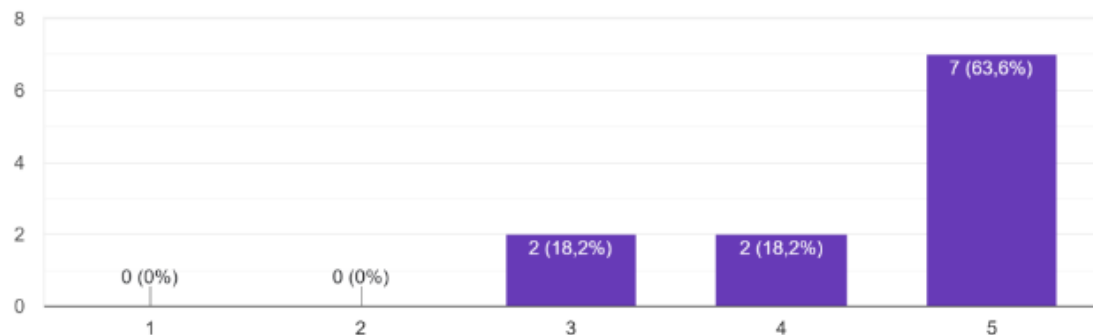
Com a avaliação em 63% de notas 5 na Figura 6.21, foi notado que a velocidade de respostas dos gráficos não foi gratificante para todos os participantes da pesquisa.

Já que a plataforma do Data Studio é disponibilizada online, é possível entender que problemas de internet podem ter afetado essas avaliações.

Figura 6.21 Resultados pergunta 11 Google Forms

Quão rápida foi a resposta dos gráficos?

11 respostas



Fonte: Dados da pesquisa.

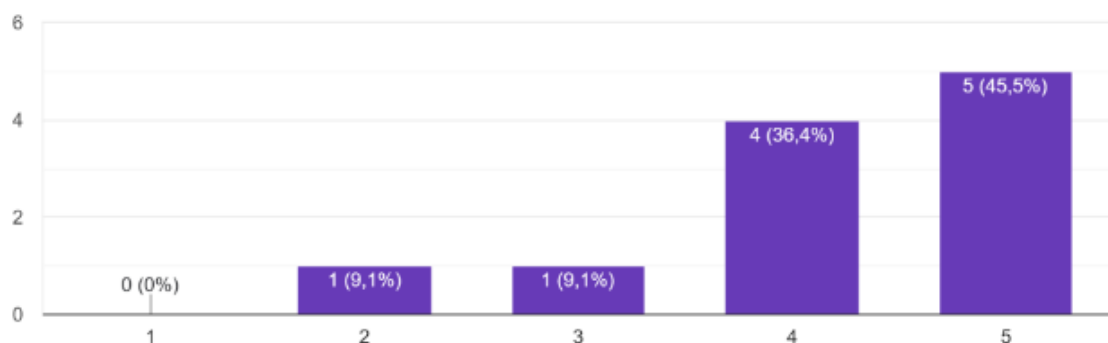
Sobre a velocidade de respostas dos filtros na Figura 6.22, 5 participantes apenas avaliaram em 5. É compreensível que os filtros exigem muito da plataforma.

Separar os filtros em páginas diferentes, poderia fazer o sistema rodar mais facilmente em relação aos filtros atribuídos ao *dashboard*.

Figura 6.22 Resultados pergunta 12 Google Forms

Quão rápida foi a resposta dos filtros?

11 respostas



Fonte: Dados da pesquisa.

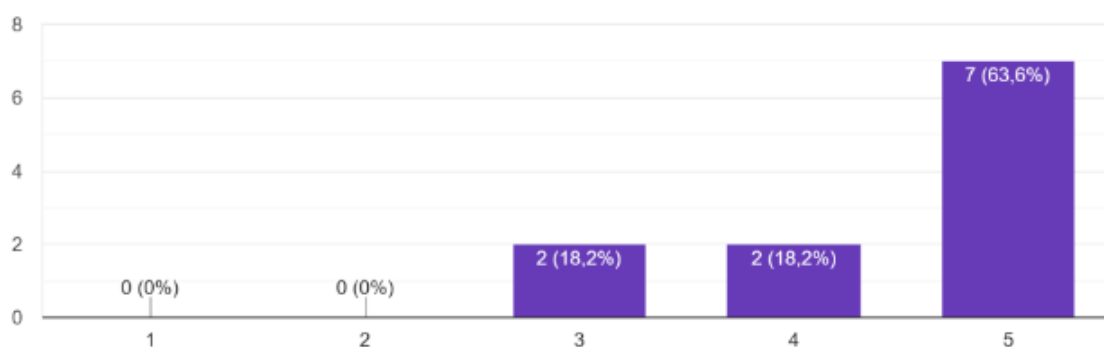
A confiabilidade dos gráficos foi avaliada em 7 votos como confiáveis na Figura 6.23, entendendo que os participantes acreditam em como os dados são informados.

Porém uma breve apresentação dos dados em formato bruto e demonstrando de onde eles estão sendo retirados, seria útil para obter uma avaliação melhor sobre a confiabilidade.

Figura 6.23 Resultados pergunta 13 Google Forms

Quão confiáveis foram os gráficos?

11 respostas



Fonte: Dados da pesquisa.

6.4.2 Resultados após a compreensão dos dados da pesquisa

Encerrando as perguntas pelo Google Forms, foi identificado que a pesquisa recebeu a maior parte de notas positivas, entendendo que o *dashboard* é interessante e intuitivo, que os gráficos são bem analisados e os filtros são compreensíveis, a aprendizagem vai sendo escalonada e intuitiva diante dos participantes.

É compreensível que melhorias na distribuição e visualização dos gráficos ajudariam em como o sistema responde a interface produzida nesse *dashboard*. Em um contra ponto, foi possível observar que se houvesse um treinamento com os participantes que responderam o questionário, seria mais fácil entender como o sistema funciona e como utilizar o *dashboard* de maneira simples, assim seriam obtidos um maior valor de notas positivas.

6.5 Feedbacks

A partir das necessidades de visualizar os dados da CD a empresa de gestão de software retornou sobre como utilizou e compreendeu o sistema do Data Studio, gerando *feedbacks* sobre o uso e a compreensão do *dashboard* que contém os dados provenientes da Câmara dos Deputados.

De acordo com a empresa, a evolução dos *dashboards* apresentados foi surpreendente e positiva, atendendo às expectativas estabelecidas. A transformação dos dados textuais em visuais foi particularmente elogiada, pois facilitou a compreensão das informações.

No entanto, surgiu um questionamento sobre o recorte temporal dos dados onde a empresa gostaria de ter uma coleta de dados mais abrangente após a conclusão da pesquisa.

Além disso, foi ressaltado que a decisão sobre como os dados são apresentados deve seguir um ciclo de uso dos dados, enfatizando a importância de compreender os projetos e seus tipos, conforme as temáticas escolhidas pela empresa, presentes na CD. A partir desse ponto, a utilização dos filtros foi destacada como uma ferramenta muito útil, permitindo uma análise mais aprofundada dos projetos em relação a temas específicos.

A interatividade com os gráficos também foi discutida. Alguns membros da equipe enfrentaram dificuldades com a interatividade dos gráficos, pois não estavam familiarizados com o *dashboard*. A empresa sugeriu a realização de uma apresentação do *dashboard* para nivelar o conhecimento e garantir que toda a equipe possa utilizar o sistema com um conhecimento uniforme e eficaz.

7

Conclusões e Trabalhos Futuros

Com a conclusão do banco de dados, o desenvolvimento da interface e a conexão dos dados no Data Studio, a empresa de gestão de software adquire a capacidade de visualizar os dados de forma dinâmica. O *dashboard* gerado a partir desse projeto oferece diferentes tipos de gráficos e filtros personalizáveis para atender às necessidades específicas de pesquisa da empresa. Além disso, o sistema está disponível na *web*, permitindo acesso remoto por meio de um *link* fornecido para empresa.

Nesse sentido, agora a empresa de gestão de software consegue compreender melhor os dados da CD que foram tratados e transformados de um formato textual, para um formato visual, por meio de um *dashboard* dinâmico e compreensível que atendeu as necessidades pela empresa designadas. Essa abordagem eficiente simplifica a interpretação dos dados e auxilia na tomada de decisões.

7.1 Trabalhos Futuros

Atualmente o sistema tem um limite de dados que são informados para a empresa, contendo apenas a base de dados do ano de 2023 que foi baixada pelo site da CD e o *dashboard* na versão mais simples do Data Studio. Uma sugestão seria implementar algumas ideias para um futuro desenvolvimento dessa pesquisa, como a automação e a manipulação de novos formatos de visualização dos dados.

7.1.1 API da Câmara dos Deputados

Utilizar a Application Programming Interface (API) da Câmara dos Deputados deixando as atualizações do *dashboard* acontecendo diariamente faria uma automação desse projeto, onde seria necessário apenas a manutenção do código caso houver alguma mudança, com isso, implementando o uso da API ao banco de dados. Com um código configurado para a API, o sistema não necessitaria mais de downloads das bases de dados no site da câmara, pois faria essa atualização de forma dinâmica e automática.

7.1.2 Implementações ao dashboard

Apesar do *dashboard* ter uma grande quantidade de gráficos e filtros personalizáveis, seria interessante colocar novas formas de visualizar os dados, utilizando versões pagas do Data Studio. Uma visualização interessante seria a nuvem de palavras de acordo com as ementas, seria necessário uma alteração no código para fazer essa atribuição, porém deixaria o *dashboard* com um aspecto mais futurista e inovador. Seria possível criar essa nova versão do código utilizando um filtro para remover palavras simples como conectivos.

- ALBUQUERQUE, J. A. E. de et al. A visualização da informação nos observatórios de informação das Federações da Indústria do Brasil. **Revista Informação na Sociedade Contemporânea**, [S.l.], v.8, p.e34745–e34745, 2024.
- APRIANI, D.; AAN, M.; SAPUTRA, W. E. Data visualization using google data studio. **International Journal of Cyber and IT Service Management**, [S.l.], v.2, n.1, p.11–19, 2022.
- BENOIT, G. **Introduction to Information Visualization**: transforming data into meaningful information. London: Rowman & Littlefield, 2019.
- BITTENCOURT, P. A. S. et al. A linguagem R e sua importância na recolha e análise dos dados. In: ATAS DO CONGRESSO INTERNACIONAL SOBRE AVALIAÇÃO NO ENSINO SUPERIOR. **Anais...** [S.l.: s.n.], 2019. p.179.
- Brasil. Governo Federal. **Sobre Dados Abertos**. 2023.
- CAFÉ, L. M. A.; SALES, R. d. **Organização da informação**: conceitos básicos e breve fundamentação teórica. Brasília DF: IBICT, 2010. 335p. In: Passeios no bosque da informação: estudos sobre representação e organização da informação e do conhecimento.
- CAMPOS, L. d. A. Interface Visual Interativa para Dados Abertos sobre Proposições na Câmara de Deputados. , [S.l.], 2021.
- CESAR, L.; COSTA, M. **Power BI vs Google Data Studio: qual é a melhor ferramenta de análise de dados?** Disponível em: <https://blog.nectarcrm.com.br/power-bi-ou-data-studio/> Acesso em: 16 nov. 2023.
- CRAVO, D. C.; JOELSONS, M. A importância do CDC no tratamento de dados pessoais de consumidores no contexto de pandemia e de vacatio legis da LGPD. **Revista de Direito do Consumidor**, [S.l.], 2020.
- Câmara dos Deputados. **Propostas legislativas**. Disponível em: <https://www.camara.leg.br/buscaProposicoesWeb/pesquisaAvancada> Acesso em: 09 abr. 2024.
- FEW, S. **Information Dashboard Design: the efective visual communication of data**. 1.ed. [S.l.]: O'Reilly, 2006.
- FIGUEIREDO, T. N. et al. Coletando e analisando informações do Facebook com a linguagem R. , [S.l.], 2018.
- FILARDI, A. L.; TRAINA, A. J. M. Montando questionários para medir a satisfação do usuário: avaliação de interface de um sistema que utiliza técnica de recuperação de imagens por conteúdo. , [S.l.], 2008.
- GALDINO, R. et al. Contribuições da bibliometria ao tratamento de dados institucionais não sistematizados de produção científica: o caso do instituto federal de educação, ciência e tecnologia de são paulo (ifsp). **Brazilian Journal of Information Science**, [S.l.], n.16, p.7, 2022.

- HARVARD, U. **Data Science: visualization**. Available at <https://pll.harvard.edu/course/data-science-visualization/2023-10> acessado 16 de novembro de 2023.
- JUNIOR, P. C. T. Business Intelligence para leigos: análise do desenvolvimento de uma ferramenta de business intelligence na escola de aprendizes-marinheiros no espírito santo. **Seminário de Gestão Organizacional Contemporânea-SEGOC**, [S.l.], v.1, 2021.
- LANKOW, J. et al. **Infographics: the power of visual storytelling**. [S.l.]: Wiley, 2012. (Business, marketing).
- NOGUEIRA. Dashboard da Graduação do Departamento de Estatística Shiny e Streamlit. , [S.l.], 2021.
- NUSSBAUMER, C. **Storytelling With Data: a data visualization guide for business professionals**. 1.ed. [S.l.]: John Wiley Sons, Inc, 2015.
- OLIVEIRA, R. B. P. de; LICHTNOW, D. Uma Análise Comparativa das Ferramentas Microsoft Power BI e Google Data Studio. In: XVII ESCOLA REGIONAL DE BANCO DE DADOS. **Anais...** [S.l.: s.n.], 2022. p.127–130.
- Portal da Câmara dos Deputados. **Acesso à Informação**. Disponível em: https://www2.camara.leg.br/transparencia/aceso-a-informacao/copy_of_perguntas-frequentes/processo-legislativo Acesso em: 25 nov. 2023.
- POSSAMAI, A. J.; SOUZA, V. G. de. Transparência e dados abertos governamentais: possibilidades e desafios a partir da lei de acesso à informação. **Administração Pública e Gestão Social**, [S.l.], 2020.
- ROSA, W. P. et al. Representação visual semântica das pesquisas desenvolvidas nos cursos de graduação e pós-graduação em Biblioteconomia e Ciência da Informação da UFRGS. **Revista Brasileira de Biblioteconomia e Documentação**. São Paulo. Vol. 17 (2021), p. 1-19, [S.l.], 2021.
- SANTOS. Desenvolvimento de Dashboard para Visualização de Dados de Pessoas com Doença de Parkinson. , [S.l.], 2023.
- SEMERANO, V. Z. O Papel Estratégico do Analista de Qualidade (QA) em Equipes SCRUM, Kanban e Scrumban no Desenvolvimento de Software Ágil. **Advances in Global Innovation & Technology**, [S.l.], v.2, n.2, p.32–45, 2024.
- SHARDA, R. et al. **Business Intelligence e Análise de Dados para gestão do negócio**. 4.ed. Porto Alegre: Bookman, 2019.
- SILVA, A. d. A. P. et al. Qualidade da Informação dos dados governamentais abertos: análise do portal de dados abertos brasileiro. **Revista Gestão em Análise**, [S.l.], v.9, n.1, p.31–47, 2020.
- SILVA, F. C. C. da. **Gestão de dados científicos**. [S.l.]: Interciencia, 2021.
- SILVA, P. N. Dados governamentais abertos sobre a COVID-19 no Brasil: avaliação do potencial de reúso. **Em Questão**, [S.l.], v.29, p.e-124820, 2023.
- Tidyverse. **R packages for data science**. Disponível em: <https://www.tidyverse.org/> Acesso em: 16 nov. 2023.