

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

NÚCLEO DE EDUCAÇÃO A DISTÂNCIA Pós-graduação *Lato Sensu* em Analytics e Business Intelligence

RELATÓRIO TÉCNCO

ACOMPANHAMENTO DOS GASTOS DOS PARLAMENTARES

Willyan Silveira Neves

Belo Horizonte 2021

SUMÁRIO

1. Introdução	3
1.1. Contexto	3
1.2. Objetivos	3
1.3. Público alvo	3
2. Modelo de Dados	4
2.1. Modelo Dimensional	4
2.2. Fatos e Dimensões	5
3. Integração, Tratamento e Carga de Dados	7
4. Camada de Apresentação	14
4.1 Dashboard	14
4.2 Análises avançadas	14
5. Registros de Homologação	14
5. Conclusões	14
6 Links	15

1. Introdução

1.1. Contexto

Acompanhar os gastos públicos e suas propostas de leis dos integrantes da câmara dos deputados, deveria ser uma obrigação de todas as pessoas e existem vários fatores para que isso não aconteça. Um deles, é a forma como essa informação está disposta para consulta.

Para tanto, o site da câmara (https://dadosabertos.camara.leg.br/), disponibiliza uma secção onde é possível baixar um dataset que pode ser trabalhado por profissionais da área de BI afim de gerar informações que posteriormente podem ser fornecidas para o público em geral ou servir para consulta de alguma empresa.

São fornecidas as informações dos gastos, serviços onde esse dinheiro foi gasto, fornecedor, as proposições (propostas de leis) dos deputados, características dessas proposições , bem como outras informações do deputado como estado e partido etc.

1.2. Objetivos

Dado ao atual cenário político, onde cada vez mais as pessoas se interessam por esse assunto e para o cumprimento dos requisitos necessários para a conclusão do presente curso, este trabalho foi desenvolvido.

1.3. Público alvo

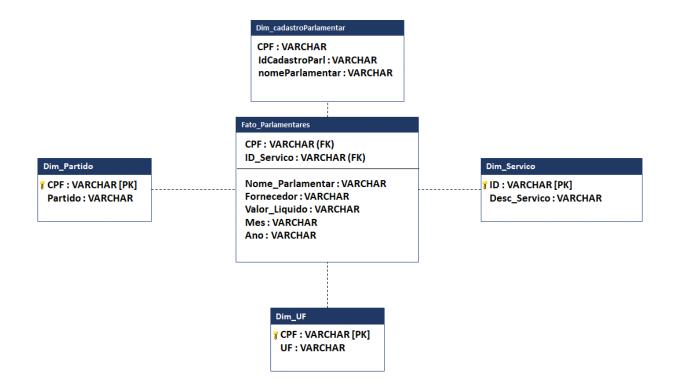
Os destinatários a solução criada serão qualquer pessoa ou instituição com acesso à internet e um navegador, seja ele por computador ou celular e que tenham o interesse no conhecimento dos gastos e proposições dos parlamentares.

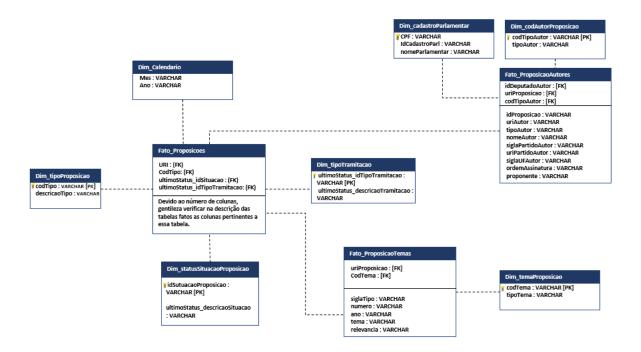
A aplicação também pode ser utilizada pelo App do power BI, que é o ambiente onde será desenvolvido as visões necessárias.

2. Modelo de Dados

2.1. Modelo Dimensional

O diagrama do modelo dimensional oferece uma visão sobre como os dados serão organizados para atender às necessidades identificadas por meio dos objetivos do projeto.





2.2. Fatos e Dimensões

Tabela fato "Fato_Parlamentares": Contém os registros dos gastos acompanhados do Nome do parlamentar, cpf, serviço utilizado, fornecedor, valor líquido, mês e ano.

Tabela fato "Fato_ProposicaoAutores": Contém os registros da autoria das propostas feitas pelos deputados, colunas idProposicao, uriProposicao, idDeputadoAutor, uriAutor, codTipoAutor, tipoAutor, nomeAutor, siglaPartidoAutor, uriPartidoAutor, siglaUFAutor, ordemAssinatura, proponente.

Tabela fato "Fato_ProposicaoTemas": Tabela que contém os temas das propostas, colunas uriProposicao, siglaTipo, numero, ano, codTema, tema, relevância.

Tabela fato "Fato_Proposicoes": Tabela que contém as propostas dos deputados bem como suas características, colunas id, uri, siglaTipo, numero, ano, codTipo, descricaoTipo, ementa, ementaDetalhada, Keywords, dataApresentacao, uriOrgaoNumerador, uriPropAnterior, uriPropPrincipal, uriPropPosterior urlInteiroTeor, urnFinal, ultimoStatus_dataHora, ultimoStatus_uriRelator, ultimoStatus_sequencia ultimoStatus_idOrgao, ultimoStatus_siglaOrgao ultimoStatus_uriOrgao, ultimoStatus_regime,

ultimoStatus_descricaoTramitacao, ultimoStatus_idTipoTramitacao, ultimoStatus_descricaoSituacao, ultimoStatus_idSituacao, ultimoStatus_despacho, ultimoStatus_url

Tabela "**Dim_Partido**": Tabela dimensão para referência dos partidos. Contém a chave primária CPF e o partido do parlamentar.

Tabela "**Dim_UF**": Tabela dimensão para referência dos estados pertencentes aos deputados. Contém a chave primária CPF e o UF (estado) do parlamentar.

Tabela "**Dim_Servico**": Tabela dimensão para referência dos tipos de serviços que os deputados podem utilizar. Contém a chave primária ID e a descrição do serviço utilizado pelo parlamentar.

Tabela "**Dim_cadastroParlamentar**": Tabela dimensão para referência dos deputados. Contém a chave primária IdCadastroParl, CPF e nomeParlamentar.

Tabela "**Dim_Calendario**": Tabela dimensão para referência das datas. Contém a tabela com as colunas de mês e ano.

Tabela "**Dim_codAutorProposicao**": Tabela dimensão para referência dos do autor da proposição. Contém a chave primária codTipoAutor e tipoAutor.

Tabela "**Dim_statusSituacaoProposicao**": Tabela dimensão para referência da situação da proposição. Contém a chave primária idSituacaoProposicao e ultimoStatus descricaoSituacao.

Tabela "**Dim_tipoProposicao**": Tabela dimensão para referência dos do tipo de proposição. Contém a chave primária codTipo e descricaoTipo.

Tabela "**Dim_temaProposicao**": Tabela dimensão para referência dos do tema em que a proposição se enquadra. Contém a chave primária codTema e tipoTema.

Tabela "**Dim_tipoTramitacao**": Tabela dimensão para referência do status da tramitação da proposição. Contém a chave primária ultimoStatus_idTipoTramitacao e ultimoStatus descricaoTramitacao.

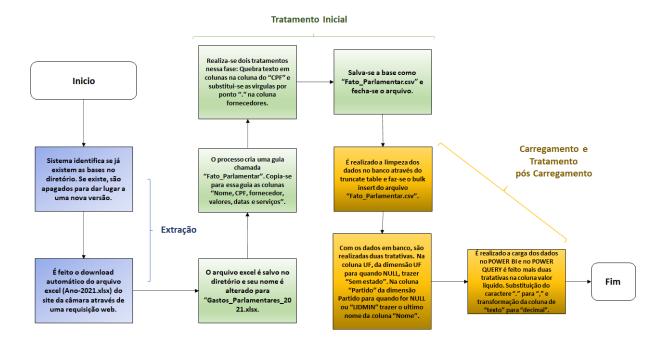
3. Integração, Tratamento e Carga de Dados

Descrição dos processos de transformação e carga de dados com imagens dos fluxos criados.

Todos os processos citados, ocorrem duas vezes, uma para o ano de 2020 e outra para o ano de 2021. Os processos de extração e carga se deram com código próprio criado na linguagem python. Já os processos de tratativa, se deram pela linguagem python, também com código próprio, no sql management studio e no power query no power bi.

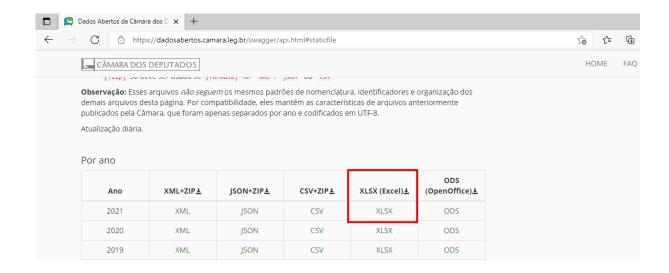
Todos os códigos estão disponíveis no link fornecido.

Ingestão de Dados e Processos ETL – Tabela Fato_Parlamentares

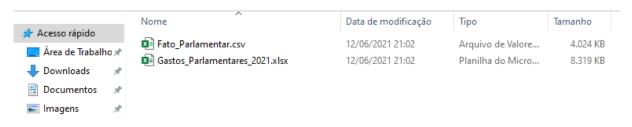


Página de download, site da câmara.

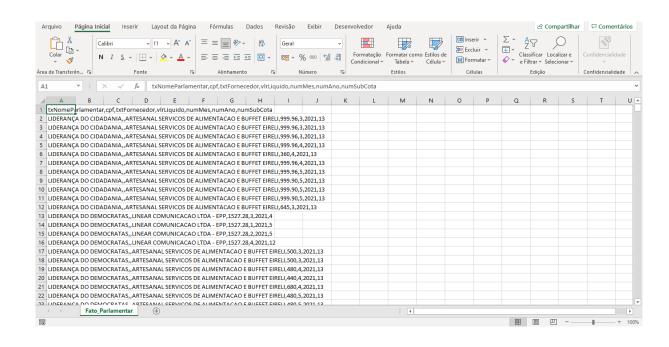
https://dadosabertos.camara.leg.br/swagger/api.html#staticfile. O arquivo baixado está marcado.



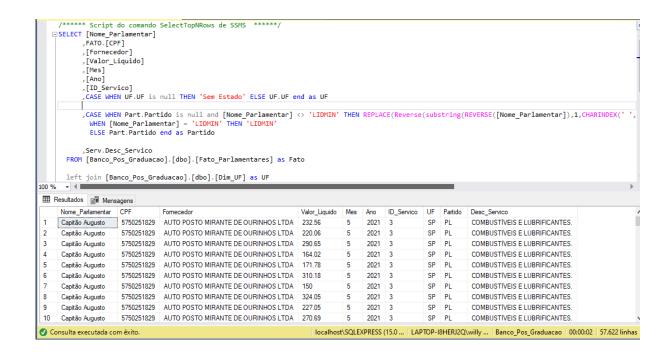
Na imagem abaixo, vemos o arquivo "Gastos_Parlamentares_2021.xlsx" que anteriormente se chamava "Anos-2021.xlsx". Vemos também o arquivo "Fato_Parlamentar.csv" tratado conforme o fluxo já informado.



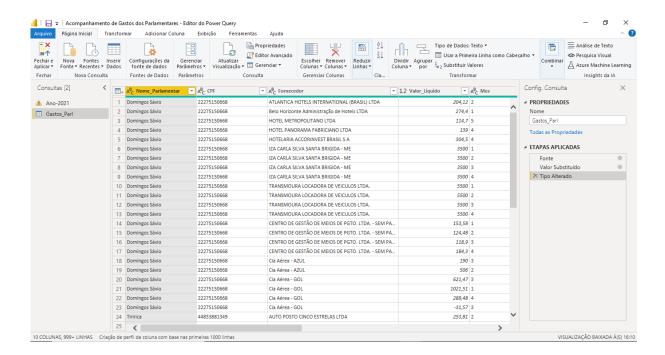
Base "Fato_Parlamentar.csv".



Dados em banco após o bulk insert.

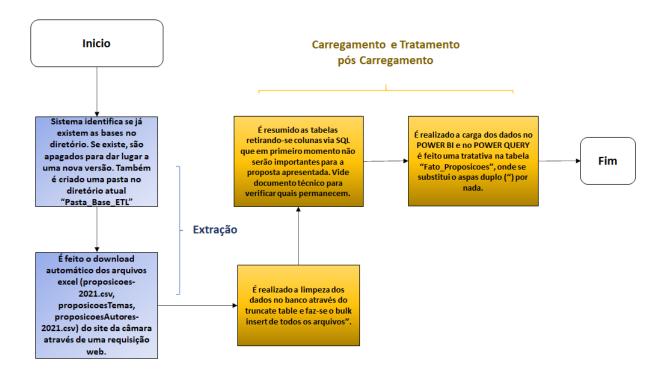


Dados tratados na tabela "Fato Parlamentares".



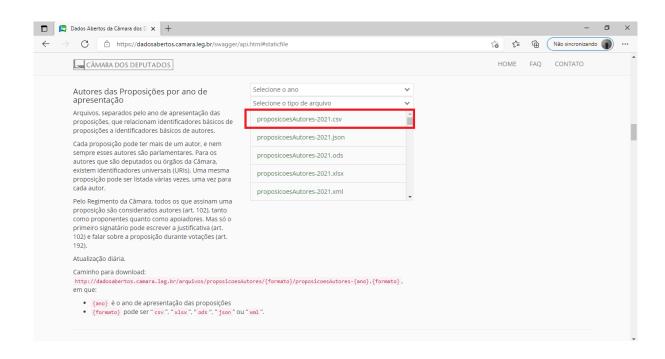
Para as tabelas a seguir, é utilizado o mesmo mecanismo de extração e carga contidos nos códigos python enviados, diferenciando se na tratativa dos dados que para estas tabelas, ocorrem no sql management studio e power query do power bi.

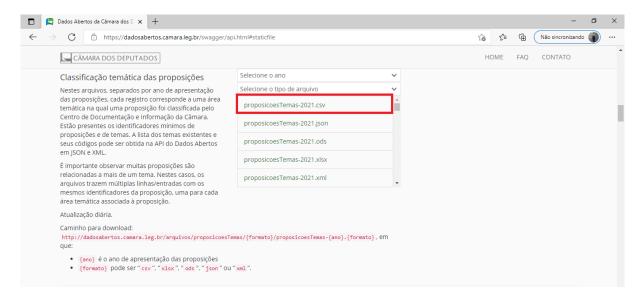
Ingestão de Dados e Processos ETL – Tabelas Fato_ProposicaoAutores, Fato_ProposicaoTemas e Fato_Proposicoes.

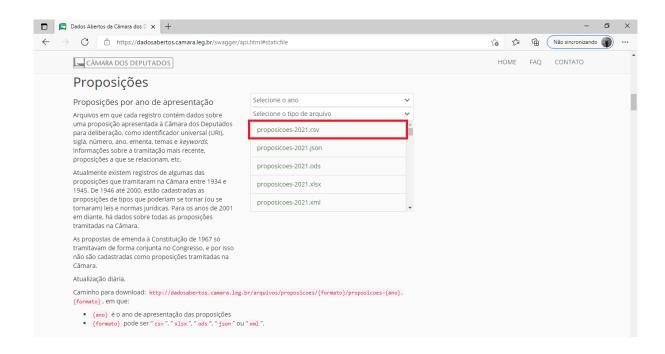


Página de download, site da câmara.

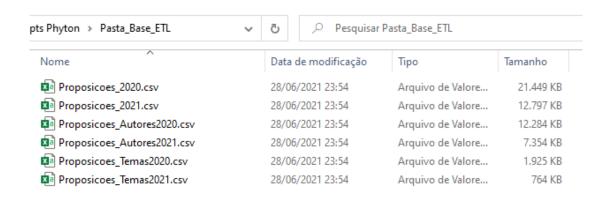
https://dadosabertos.camara.leg.br/swagger/api.html#staticfile. O arquivo baixado está marcado.



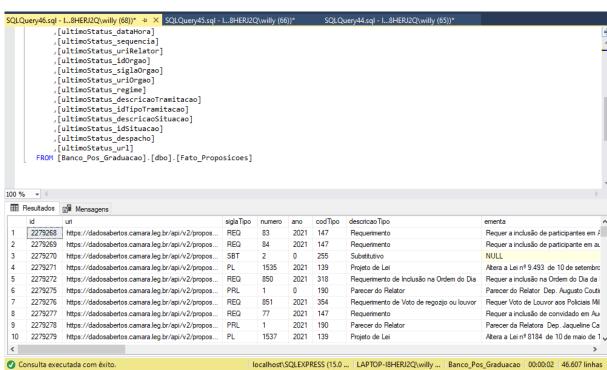


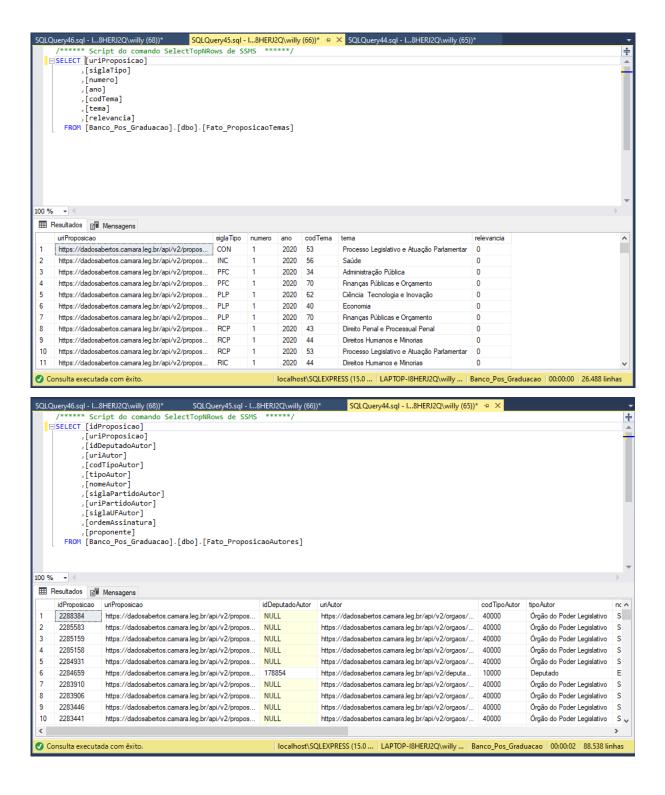


Arquivos extraídos e renomeados conforme abaixo.



Dados em banco.





Tratativas no SQL:

Tratativas na tabela Fato_ProposicaoAutores: Colunas permanecentes uriProposicao, idDeputadoAutor, codTipoAutor, tipoAutor, nomeAutor, siglaPartidoAutor.

Tratativas na tabela Fato_ProposicaoTemas: Colunas permanecentes uriProposicao, ano, codTema, tema

Tratativas na tabela Fato_Proposicoes: Colunas permanecentes uri, ano, codTi-po,descricaoTipo, ementa, dataApresentacao,urlInteiroTeor, ultimoStatus_dataHora, ultimoStatus_descricaoTramitacao, ultimoStatus_idTipoTramitacao, ultimoStatus_descricaoSituacao, ultimoStatus_idSituacao, ultimoStatus_despacho

Tratativas no power query:

Na Fato_proposicoes, é realizado a substituição do caractere aspas duplo (") por nada.

4. Camada de Apresentação

4.1 Dashboard

Descreva os painéis de controle (dashboard) criados detalhando as visões estratégica, tática e operacional e como são interligadas. Descreva os indicadores e suas fórmulas de cálculo, as métricas, as dimensões de análise, as visualizações empregadas (mapas, gráficos, gauges, entre outros), bem como os filtros aplicáveis à solução, com base nas dimensões utilizadas.

4.2 Análises avançadas

Padrões e/ou Tendências: análises de dados criadas a partir da aplicação de técnicas de aprendizado de máquina. As análises podem ser independentes ou incorporadas no dashboard criado anteriormente.

5. Registros de Homologação

Testes da solução desenvolvida mostrando que o dado apresentado no dashboard é o mesmo dos sistemas fonte. Estes testes são apresentados, normalmente, por meio de consultas SQL feitas contra as fontes de dados e pelo confrontamento dos resultados com o que é exibido nas visualizações de dados apresentadas nos dashboards

5. Conclusões

Este item deve apresentar os seguintes itens: (1) análise crítica apresentando os achados mais relevantes nos dados, feitos a partir do uso do dashboard e das experiências adquiridas no processo de desenvolvimento; (2) proposta de intervenção, possíveis ações a serem tomadas por gestores do contexto analisado no intuito de melhorar o desempenho da organização; e, por fim, (3) as lições aprendidas no processo que possam ajudar nos projetos futuros.

Aponte, ainda, as limitações do trabalho e possíveis pontos de extensão para que outros possam utilizar como ponto de partida em novos projetos.

6. Links

https://drive.google.com/drive/folders/1O5WbvMiLeD3gzzZUYZVi_5rb8Hsf-2-f?usp=sharing

REFERÊNCIAS

Como um projeto aplicado, as referências, levantadas no processo de revisão bibliográfica, são opcionais e estimulamos que o responsável pelo projeto apresente aquilo que for relevante para complementar o entendimento do trabalho. Dessa forma, relacione-as de acordo com o modelo a seguir.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.