

MVP de Engenharia de Dados

Profs. Victor Almeida e Silvio Alonso

Aluno: José Guilherme Campos Machado

Descrição

Esse trabalho é referente ao MVP da 3ª sprint (Engenharia de Dados) do curso de Pós-Graduação em Ciência de Dados e Analytics da PUC-RJ

Nele será desenvolvida uma pipeline de dados utilizando tecnologias de nuvem na plataforma da AWS. A pipeline irá conter a busca, coleta, modelagem, carga e análise dos dados.

Objetivo escolhido

Encontrar no Portal da Transparência do governo brasileiro informações sobre as verbas destinadas pelos políticos a suas bases eleitorais.



Podemos elencar como possíveis perguntas a serem respondidas:

- 1) Quais políticos conseguiram aprovar mais emendas orçamentárias?
- 2) Quais políticos aprovaram maiores valores?
- 3) Para quais setores foram destinadas mais verbas?
- 4) Quais estados receberam mais verbas?

Se for possível tentarei mapear se políticos de situação conseguem aprovar mais emendas orçamentárias que políticos de oposição.

Detalhamento

1. Busca pelos dados

Para responder as perguntas citadas acima busquei no Portal da Transparência do governo brasileiro informações detalhadas sobre emendas parlamentares referentes a orçamento.

No site é possível existe a possibilidade de filtrar as informações mas também existe um arquivo único com as informações disponíveis a partir de 2014.

<https://portaldatransparencia.gov.br/download-de-dados/emendas-parlamentares>

Optei por baixar o arquivo único

Dados Abertos (Trechos retirados do site)

“O Portal Mapa Brasil Transparente disponibiliza as informações das consultas em formato apropriado para download e processamento. O usuário poderá baixar as informações referentes às avaliações independentes e fazer todos os cruzamentos e análises desejados.

O novo Portal da Transparência reforça, assim, com novos recursos e mais informações, sua razão de ser uma ferramenta que permita ao cidadão, de forma cada vez mais eficiente, fiscalizar e assegurar a boa e correta aplicação dos recursos públicos federais.

O acesso ao Portal não requer usuário nem senhas, sendo permitido a qualquer cidadão navegar pelas páginas de forma livre, bem como visualizar e utilizar os dados disponíveis da forma que melhor lhe convier.”

2. Coleta

https://github.com/JGCMachado/Notebooks/blob/main/MVP3_Emendas_Parlamentares_Work.ipynb