UNIVERSIDADE DE BRASÍLIA

Faculdade do Gama

Sistemas de Banco de Dados 2

**Trabalho Final (TF)**

**Tema B - Mineração de Dados (*Data Mining - DM*)**

**João Lucas Pinto Vasconcelos - 190089601**

**Mateus Orlando Medeiros Ribeiro - 211062259**

**Zenilda Pedrosa Vieira - 212002907**

Brasília, DF

2024

**Base de Dados**

A base de dados escolhida para esse projeto do trabalho final, sugerida por João Vasconcelos, foi a obtida dos dados abertos da Câmara dos Deputados, no local virtual <https://dadosabertos.camara.leg.br/swagger/api.html>. Nesse local responsável em manter a transparência dos gastos dos deputados e partidos é possível verificar os gastos da câmara dos deputados na atual legislatura, que está vigente desde o dia 01/02/2023 e terminará em 28/02/2027. Para a base de dados do projeto foram coletados apenas os gastos do ano de 2023 e separados os gastos estranhos que totalizaram 6% da base de dados.

Como justificativa dessa escolha, ponderou-se a relevância do tema na realidade atual e sua presença no dia a dia de todos os cidadãos. É possível realizar uma mineração dos dados disponíveis e retirar conhecimentos úteis para uma análise detalhada dos gastos dos deputados, dos partidos e até por tipo de gastos, utilizando as técnicas de agrupamento, classificação e regressão de *data mining*.

**Diagrama Entidade-Relacionamento (DE-R)**

Diagrama

Descrição gerada automaticamente

**Diagrama Lógico de Dados (DLD)**

Diagrama

Descrição gerada automaticamente

**Dicionário de Dados**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: PARTIDOS** | | | | |
| **Descrição:** Define todos os partidos presentes na câmara dos deputados na legislatura atual. | | | | |
| **Atributo** | | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| sigla | | chave primária  obrigatório | varchar | 50 | Sigla do partido |
| nome | |  | varchar | 255 | Nome do partido |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: DEPUTADOS** | | | | |
| **Descrição:** Define todos os deputados de todos os partidos da legislatura atual. | | | | |
| **Atributo** | | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| id | | chave primária  obrigatório | int |  | Identificador do deputado |
| nome | |  | varchar | 255 | Nome do deputado |
| partido | | Chave estrangeira  obrigatório | varchar | 50 | Sigla do partido ao qual o deputado pertence |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: GASTOS** | | | | |
| **Descrição:** Define todos os gastos realizados por deputados no ano de 2023. | | | | |
| **Atributo** | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| id | chave primária  obrigatório | int |  | Identificador do gasto |
| id\_deputado | chave estrangeira  obrigatório | int |  | Identificador do deputado |
| data |  | date | 50 | Data da despesa |
| valor |  | float |  | Valor da despesa |
| tipoDespesa |  | varchar | 255 | Tipo da despesa |
| codDocumento |  | int |  | Código do documento gerado pelo gasto |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: GASTOS\_ESTRANHOS** | | | | |
| **Descrição:** Define todos os gastos realizados por deputados no ano de 2023 com código do documento nulo ou valor do gasto negativo. | | | | |
| **Atributo** | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| id | chave primária  obrigatório | int |  | Identificador do gasto |
| id\_deputado | chave estrangeira  obrigatório | int |  | Identificador do deputado |
| data |  | date |  | Data da despesa |
| valor |  | float |  | Valor da despesa |
| tipoDespesa |  | varchar | 255 | Tipo da despesa |
| codDocumento |  | int |  | Código do documento gerado pelo gasto |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: VW\_GASTOS\_MENSAL\_DEPUTADOS** | | | | |
| **Descrição:** Agrega o gasto mensal de cada deputado, apresentando o total de gastos em cada mês, a média dos gastos mensais, juntamente com o partido ao qual ele pertence, destacando os deputados que mais gastaram por mês (ordem decrescente de gastos). | | | | |
| **Atributo** | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| id\_deputado | chave primária  obrigatório | int |  | Identificador do deputado |
| nome\_deputado |  | varchar | 255 | Nome do deputado |
| sigla\_partido | chave primária  obrigatório | varchar | 50 | Sigla do partido referente ao deputado |
| nome\_partido |  | varchar | 255 | Nome do partido referente ao deputado |
| ano |  | int | 4 | Ano da despesa |
| mes |  | int | 2 | Mês da despesa |
| total\_gastos\_mensal |  | float |  | Total dos gastos naquele mês |
| media\_gastos\_mensais |  | float |  | Média dos gastos daquele mês por quantidade total de gastos |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entidade: VW\_TOTAL\_PARTIDO\_TIPO\_GASTOS** | | | | |
| **Descrição:** Agrega os gastos totais por partido e por tipo de despesa de toda a câmara dos deputados. Fornece uma análise detalhada dos gastos totais de cada partido, categorizados por tipo de despesa. | | | | |
| **Atributo** | **Propriedades do atributo** | **Tipo de dado** | **Tamanho** | **Descrição** |
| sigla\_partido | chave primária  obrigatório | varchar | 50 | Sigla do partido |
| nome\_partido |  | varchar | 255 | Nome do partido |
| tipoDespesa |  | varchar | 255 | Tipo da despesa |
| total\_gastos |  | float |  | Total dos gastos |