Plano de Análise do Projeto

1. Objetivo do Documento

Este documento tem como objetivo detalhar o plano de análise para o projeto de Engenharia e Análise de Dados com T-SQL. Ele traduz as perguntas de negócio em requisitos de dados, processos de análise e define os resultados esperados para responder a cada questão.

2. Perguntas de Negócio (Escopo da Análise)

As seguintes perguntas guiaram a modelagem e a análise dos dados:

- Quem são os maiores fornecedores dos deputados?
- Quais setores econômicos mais lucram com a cota parlamentar (via CNAE)?
- Há concentração de mercado (poucas empresas dominam)?
- Quais partidos/estados concentram mais gastos em determinados setores?
- Existe sazonalidade nos gastos (ex.: anos eleitorais, períodos específicos)?
- Deputados mais gastadores em publicidade têm comportamento diferente em votações/presenças?
- Empresas fornecedoras atuam em vários estados ou concentram contratos em regiões específicas?

3. Requisitos de Dados

O projeto utiliza um conjunto de dados robusto, arquitetado em um esquema em estrela, com base nas seguintes fontes:

- Cota Parlamentar Brasil.io: Fonte principal de dados sobre despesas parlamentares, contendo informações detalhadas sobre cada gasto.
- Câmara dos Deputados Dados Abertos: Utilizada para obter informações sobre deputados, partidos e legislaturas.
- Receita Federal CNPJs: Utilizada para classificar os fornecedores por setor econômico (CNAE).

4. Plano de Análise por Pergunta

1. Quem são os maiores fornecedores dos deputados?

- Tabelas: FatoGastoParlamentar e Fornecedor.
- Análise: Realizar um JOIN entre as tabelas. Agrupar os dados por ID_Fornecedor e Nome, somar o campo Valor na tabela de fato e ordenar de forma decrescente para ranquear os maiores fornecedores.

2. Quais setores econômicos mais lucram com a cota parlamentar (via CNAE)?

- o **Tabelas:** FatoGastoParlamentar, Fornecedor e Setor.
- Análise: Fazer JOINs para conectar as três tabelas. Agrupar por ID_Setor e
 DescricaoCNAE, somar o campo Valor e ordenar de forma decrescente.
- 3. Existe concentração de mercado (poucas empresas dominam)?
 - o **Tabelas:** FatoGastoParlamentar, Fornecedor e Setor.
 - Análise: Complementar a análise anterior. Para cada setor, calcular a porcentagem do gasto total que cada fornecedor representa. Isso identificará se poucos fornecedores concentram a maioria dos gastos em um setor específico.
- 4. Quais partidos/estados concentram mais gastos em determinados setores?
 - o **Tabelas:** FatoGastoParlamentar, Deputado, Partido e Setor.
 - Análise: Fazer os JOINs para conectar as tabelas. Agrupar os dados por Sigla do partido, UF do deputado e Nome do setor. Somar o Valor para encontrar as maiores concentrações de gastos.
- 5. Existe sazonalidade nos gastos (ex.: anos eleitorais)?
 - o **Tabelas:** FatoGastoParlamentar e Tempo.
 - Análise: Agrupar os dados por Dia, Mês e Ano da dimensão de tempo. Somar o campo Valor para analisar a série histórica. Também é possível analisar o campo AnoEleitoral da dimensão de tempo para comparar os gastos nesse período com os demais anos.
- 6. Deputados mais gastadores em publicidade têm comportamento diferente em votações/presenças?
 - Tabelas: FatoGastoParlamentar, SessaoVotacao, CategoriaDespesa e Deputado.
 - Análise: Primeiro, filtrar os gastos na tabela FatoGastoParlamentar usando o ID_Categoria para a categoria de "publicidade". Conectar o resultado com a tabela SessaoVotacao via ID_Deputado. Analisar a frequência de presença e o padrão de votos (Sim/Não) dos deputados que se destacam por gastos elevados em publicidade.
- 7. Empresas fornecedoras atuam em vários estados ou se concentram em regiões específicas?
 - o **Tabelas:** FatoGastoParlamentar, Fornecedor **e** Deputado.
 - Análise: Fazer os JOINs necessários. Agrupar por ID_Fornecedor e, para cada um, contar o número de UFs (estados) únicos de origem dos deputados que contrataram esse fornecedor.