

# Atividade Final da Disciplina de Modelagem de Dados

## Objetivo:

Construir um modelo de dados relacional e multidimensional completo e robusto para os dados de candidatos e votação do primeiro turno das eleições municipais de 2024, utilizando os dados disponíveis no site do TSE, com foco na análise, qualidade dos dados e performance.

## Etapas:

### 1. Download e Organização dos Dados:

#### 1.1. Acesso e Download:

Após o acesso ao site do TSE (Tribunal Superior Eleitoral) na área de dados abertos e realizei o download dos seguintes arquivos:

- Consulta\_Candidatos\_2024.zip
- Bens\_Candidato\_2024.zip
- Redes\_Sociais\_Candidato\_2024.zip
- Motivo\_Cassacao\_Candidato\_2024.zip
- Votacao\_Nominal\_MunZona\_2024.zip

#### 1.2. Organização:

- Uma pasta no meu computador chamada "Atividade\_Modelagem\_Dados" foi criada e dentro dela subpastas foram criadas para cada arquivo baixado, por exemplo, "Candidatos", "Bens", "Redes\_Sociais", "Cassacao" e "Votacao".
- Cada arquivo zip foi descompactado na sua respectiva subpasta.

#### 1.3. Leitura e Compreensão:

- Uma leitura dos arquivos "Leiame.pdf" presentes em cada subpasta foi realizada para entender a estrutura dos dados, os tipos de campos, os códigos utilizados e as relações entre os arquivos.
- Uma análise do dicionário de dados foi realizada para cada arquivo.

#### 1.4. Filtragem por UF:

Foram selecionados os arquivos que possuem a sigla do estado de Minas Gerais, quais sejam Consulta\_Candidatos\_2024\_MG.csv, Bens\_Candidato\_2024\_MG.csv, Redes\_Sociais\_Candidato\_2024\_MG.csv, Motivo\_Cassacao\_Candidato\_2024\_MG.csv, Votacao\_Nominal\_MunZona\_2024\_MG.csv. Os arquivos selecionados e movidos para a pasta "Atividade\_Modelagem\_Dados"

## 2. Modelagem Conceitual (MER):

### 2.1. Identificação de Entidades:

Uma análise dos arquivos de dados filtrados foi realizada e identifiquei as entidades principais, quais sejam:

1. Candidato
2. Rede Social
3. Bem do Candidato
4. Eleição
5. Cargo
6. Partido
7. Coligação/Federação
8. Votação
9. Município
10. Motivo Cassação.

## **2.2. Definição de Atributos:**

Para cada entidade, liste os atributos relevantes, classificando-os em:

### **Candidato**

- **Atributos:**

- SQ\_CANDIDATO (chave primária)
- NR\_CANDIDATO
- NM\_CANDIDATO
- NM\_URNA\_CANDIDATO
- NM\_SOCIAL\_CANDIDATO
- NR\_CPF\_CANDIDATO
- DS\_EMAIL
- DT\_NASCIMENTO
- NR\_TITULO\_ELEITORAL\_CANDIDATO
- CD\_GENERO (2: Masculino, 4: Feminino)
- CD\_GRAU\_INSTRUCAO (1 a 8 conforme descrição)
- CD\_ESTADO\_CIVIL (1, 3, 5, 7, 9)
- CD\_COR\_RACA (01 a 06 conforme descrição)
- CD\_OCUPACAO
- DS\_OCUPACAO
- SG\_PARTIDO
- SQ\_COLIGACAO
- CD\_ELEICAO

### **Rede Social**

- **Atributos:**

- SQ\_CANDIDATO (chave estrangeira)
- NR\_ORDEM\_REDE\_SOCIAL (chave primária composta com SQ\_CANDIDATO)
- DS\_URL

### **Bem do Candidato**

- **Atributos:**

- SQ\_CANDIDATO (chave estrangeira)
- NR\_ORDEM\_BEM\_CANDIDATO (chave primária composta com SQ\_CANDIDATO)
- CD\_TIPO\_BEM\_CANDIDATO
- DS\_TIPO\_BEM\_CANDIDATO
- DS\_BEM\_CANDIDATO
- VR\_BEM\_CANDIDATO
- DT\_ULT\_ATUAL\_BEM\_CANDIDATO
- HH\_ULT\_ATUAL\_BEM\_CANDIDATO

### **Eleição**

- **Atributos:**

- CD\_ELEICAO (chave primária)
- AA\_ELEICAO
- DT\_ELEICAO
- CD\_TIPO\_ELEICAO (1, 2, 3)
- NM\_TIPO\_ELEICAO
- DS\_ELEICAO
- TP\_ABRANGENCIA (Municipal, Estadual, Federal)
- CD\_CARGO

## Cargo

- **Atributos:**

- CD\_CARGO (chave primária)
- DS\_CARGO

## Partido

- **Atributos:**

- SG\_PARTIDO (chave primária)
- NR\_PARTIDO
- NM\_PARTIDO

## Coligação/Federação

- **Atributos:**

- SQ\_COLIGACAO (chave primária)
- NM\_COLIGACAO
- DS\_COMPOSICAO\_COLIGACAO

## Votação

- **Atributos:**

- SQ\_CANDIDATO (chave estrangeira)
- CD\_ELEICAO (chave estrangeira)
- CD\_MUNICIPIO (chave estrangeira)
- NR\_TURNO
- NR\_ZONA
- QT\_VOTOS\_NOMINAIS
- QT\_VOTOS\_NOMINAIS\_VALIDOS
- ST\_VOTO\_EM\_TRANSITO
- CD\_SIT\_TOT\_TURNO
- DS\_SIT\_TOT\_TURNO

## Município

- **Atributos:**

- CD\_MUNICIPIO (chave primária)
- NM\_MUNICIPIO
- SG\_UF

## Motivo Cassação

- **Atributos:**

- SQ\_CANDIDATO (chave estrangeira)
- CD\_MOTIVO
- DS\_MOTIVO
- TP\_MOTIVO
- DS\_TP\_MOTIVO

### 2.3. Relacionamentos:

Os relacionamentos identificados entre as entidades foram:

- Candidato tem uma ou mais Redes Sociais (1:N).
- Candidato possui uma ou mais declarações de Bens (1:N).
- Candidato participa de uma Eleição (N:1).
- Candidato é associado a um Partido ou a uma Coligação/Federação (N:1 para Partido e N:1 para Coligação/Federação).

- Candidato tem registros de Votação em diferentes zonas e municípios (1).
- Candidato pode ter um ou mais Motivos de Cassação (1).
- Redes sociais possui um candidato (N:1)
- Os Bens possuem um candidato (N:1)
- A eleição possui vários cargos (1:N)
- A eleição possui votos (1:N)
- A eleição possui candidatos (1:N)
- Vários cargos podem estar em uma eleição (N:1).
- O partido possui vários candidatos (1:N)
- A Coligação/Federação podem ter vários candidatos (1:N)
- Votação ocorre em um Município específico e tem uma relação com Candidato (N:1 para ambos).
- Município pode ter várias votações
- O motivo da cassação poder ter vários candidatos (1:N)

#### 2.4. Representação Gráfica:

O diagrama MER foi gerado de acordo com as entidades e relacionamento identificados acima:

Entidades	Atributos	Relacionamentos
1. Candidato	SQ_CANDIDATO	<ul style="list-style-type: none"> <li>- Candidato tem uma ou mais Redes Sociais (1:N).</li> <li>- Candidato possui uma ou mais declarações de Bens (1:N).</li> <li>- Candidato participa de uma Eleição (N:1).</li> <li>- Candidato é associado a um Partido ou a uma Coligação/Federação (N:1 para Partido e N:1 para Coligação/Federação).</li> <li>- Candidato tem registros de Votação em diferentes zonas e municípios (1).</li> <li>- Candidato pode ter um ou mais Motivos de Cassação (1).</li> </ul>
	NR_CANDIDATO	
	NM_CANDIDATO	
	NM_URNA_CANDIDATO	
	NM_SOCIAL_CANDIDATO	
	NR_CPF_CANDIDATO	
	DS_EMAIL	
	DT_NASCIMENTO	
	NR_TITULO_ELEITORAL_CANDIDATO	
	CD_GENERO	
	CD_GRAU_INSTRUCAO	
	CD_ESTADO_CIVIL	
	CD_COR_RACA	
	CD_OCUPACAO	
	DS_OCUPACAO	
2. Rede Social	SG_PARTIDO	<ul style="list-style-type: none"> <li>- Redes sociais possui um candidato (N:1)</li> </ul>
	SQ_COLIGACAO	
	CD_ELEICAO	
3. Bem do Candidato	SQ_CANDIDATO	<ul style="list-style-type: none"> <li>- Os Bens possuem um candidato (N:1)</li> </ul>
	NR_ORDEM_REDE_SOCIAL	
	DS_URL	
	CD_TIPO_BEM_CANDIDATO	
	DS_TIPO_BEM_CANDIDATO	
	DS_BEM_CANDIDATO	

Entidades	Atributos	Relacionamentos
	VR_BEM_CANDIDATO DT_ULT_ATUAL_BEM_CANDIDATO HH_ULT_ATUAL_BEM_CANDIDATO	
4. Eleição	CD_ELEICAO AA_ELEICAO DT_ELEICAO CD_TIPO_ELEICAO NM_TIPO_ELEICAO DS_ELEICAO TP_ABRANGENCIA	- A eleição possui vários cargos (1:N) - A eleição possui votos (1:N) - A eleição possui candidatos (1:N)
5. Cargo	CD_CARGO DS_CARGO	- Vários cargos podem estar em uma eleição (N:1).
6. Partido	SG_PARTIDO NR_PARTIDO NM_PARTIDO	- O partido possui vários candidatos (1:N)
7. Coligação/ Federação	SQ_COLIGACAO NM_COLIGACAO DS_COMPOSICAO_COLIGACAO	- A Coligação/Federação podem ter vários candidatos (1:N)
8. Votação	SQ_CANDIDATO CD_ELEICAO CD_MUNICIPIO NR_TURNO NR_ZONA QT_VOTOS_NOMINAIS QT_VOTOS_NOMINAIS_VALIDOS ST_VOTO_EM_TRANSITO CD_SIT_TOT_TURNO DS_SIT_TOT_TURNO	- Votação ocorre em um Município específico e tem uma relação com Candidato (N:1 para ambos).
9. Município	CD_MUNICIPIO NM_MUNICIPIO SG_UF	- Município pode ter varias votações
10. Motivo Cassação	SQ_CANDIDATO CD_MOTIVO DS_MOTIVO TP_MOTIVO DS_TP_MOTIVO	- O motivo da cassação poder ter vários candidatos (1:N)

### **3. Modelagem Lógica (DER):**

#### **3.1. Conversão do MER:**

O MER foi transformado em um Diagrama Entidade-Relacionamento (DER), aplicando as regras de mapeamento do modelo conceitual para o modelo lógico relacional.

#### **3.2. Normalização:**

As formas normais (1FN, 2FN, 3FN) foram aplicadas para eliminar redundâncias e garantir a integridade dos dados.

#### **3.3. Definição de Tipos de Dados:**

Os tipos de dados e tamanhos foram definidos para cada atributo, considerando os domínios dos dados e as restrições de integridade.

##### **1. Tabela Candidato**

Atributo	Tipo de Dados	Tamanho	Descrição
SQ_CANDIDATO	INT	-	Identificador único do candidato (PK)
NR_CANDIDATO	INT	-	Número do candidato
NM_CANDIDATO	VARCHAR	100	Nome completo do candidato
NM_URNA_CANDIDATO	VARCHAR	50	Nome do candidato na urna
NM_SOCIAL_CANDIDATO	VARCHAR	50	Nome social do candidato
NR_CPF_CANDIDATO	CHAR	11	CPF do candidato
DS_EMAIL	VARCHAR	100	Email do candidato
DT_NASCIMENTO	DATE	-	Data de nascimento do candidato
NR_TITULO_ELEITORAL_CANDIDATO	CHAR	12	Número do título eleitoral
CD_GENERO	TINYINT	-	Código do gênero (2: Masculino, 4: Feminino)
CD_GRAU_INSTRUCAO	TINYINT	-	Código do grau de instrução
CD_ESTADO_CIVIL	TINYINT	-	Código do estado civil
CD_COR_RACA	TINYINT	-	Código da cor/raça
CD_OCUPACAO	INT	-	Código da ocupação
DS_OCUPACAO	VARCHAR	50	Descrição da ocupação
SG_PARTIDO	CHAR	5	Sigla do partido
SQ_COLIGACAO	INT	-	Sequencial da coligação
CD_ELEICAO	INT	-	Código da eleição

##### **2. Rede Social**

Armazena os perfis de redes sociais dos candidatos.

Atributo	Tipo de Dados	Tamanho	Descrição
SQ_CANDIDATO	INT	-	Identificador do candidato (PK/FK)
NR_ORDEM_REDE_SOCIAL	SMALLINT	-	Número da ordem da rede social (PK)
DS_URL	VARCHAR	200	URL do perfil da rede social

### 3. Bem do Candidato

Armazena os bens dos candidatos.

Atributo	Tipo de Dados	Tamanho	Descrição
SG_CANDIDATO	INT	-	Identificador do candidato (PK/FK)
NR_ORDEM_BEM_CANDIDATO	SMALLINT	-	Número da ordem do bem (PK)
CD_TIPO_BEM_CANDIDATO	SMALLINT	-	Código do tipo de bem
DS_TIPO_BEM_CANDIDATO	VARCHAR	50	Descrição do tipo de bem
DS_BEM_CANDIDATO	VARCHAR	200	Descrição do bem
VR_BEM_CANDIDATO	DECIMAL	(15, 2)	Valor do bem
DT_ULT_ATUAL_BEM_CANDIDATO	DATE	-	Data da última atualização do bem
HH_ULT_ATUAL_BEM_CANDIDATO	TIME	-	Hora da última atualização do bem

### 4. Eleição

Armazena informações sobre as eleições.

Atributo	Tipo de Dados	Tamanho	Descrição
CD_ELEICAO	INT	-	Código da eleição (PK)
AA_ELEICAO	YEAR	-	Ano da eleição
DT_ELEICAO	DATE	-	Data da eleição
CD_TIPO_ELEICAO	TINYINT	-	Código do tipo de eleição
NM_TIPO_ELEICAO	VARCHAR	50	Nome do tipo de eleição
DS_ELEICAO	VARCHAR	100	Descrição da eleição
TP_ABRANGENCIA	VARCHAR	20	Abrangência da eleição (Municipal, Estadual, Federal)

### 5. Cargo

Armazena os cargos aos quais os candidatos concorrem.

Atributo	Tipo de Dados	Tamanho	Descrição
CD_CARGO	SMALLINT	-	Código do cargo (PK)
DS_CARGO	VARCHAR	50	Descrição do cargo

### 6. Partido

Armazena as informações dos partidos.

Atributo	Tipo de Dados	Tamanho	Descrição
SG_PARTIDO	CHAR	5	Sigla do partido (PK)
NR_PARTIDO	SMALLINT	-	Número do partido
NM_PARTIDO	VARCHAR	50	Nome do partido

## 7. Coligação/Federação

Armazena as coligações ou federações de partidos.

Atributo	Tipo de Dados	Tamanho	Descrição
SQ_COLIGACAO	INT	-	Sequencial da coligação/federação (PK)
NM_COLIGACAO	VARCHAR	100	Nome da coligação/federação
DS_COMPOSICAO_COLIGACAO	VARCHAR	200	Descrição da composição da coligação

## 8. Votação

Armazena os detalhes das votações dos candidatos.

Atributo	Tipo de Dados	Tamanho	Descrição
SQ_CANDIDATO	INT	-	Identificador do candidato (PK/FK)
CD_ELEICAO	INT	-	Código da eleição (PK/FK)
CD_MUNICIPIO	INT	-	Código do município (PK/FK)
NR_TURNO	TINYINT	-	Número do turno
NR_ZONA	SMALLINT	-	Número da zona eleitoral
QT_VOTOS_NOMINAIS	INT	-	Quantidade de votos nominais
QT_VOTOS_NOMINAIS_VALIDOS	INT	-	Quantidade de votos válidos nominais
ST_VOTO_EM_TRANSITO	BOOLEAN	-	Indicador de voto em trânsito
CD_SIT_TOT_TURNO	SMALLINT	-	Código da situação total do turno
DS_SIT_TOT_TURNO	VARCHAR	50	Descrição da situação total do turno

## 9. Município

Armazena os municípios onde os candidatos concorrem.

Atributo	Tipo de Dados	Tamanho	Descrição
CD_MUNICIPIO	INT	-	Código do município (PK)
NM_MUNICIPIO	VARCHAR	100	Nome do município
SG_UF	CHAR	2	Sigla da unidade federativa

## 10. Motivo Cassação

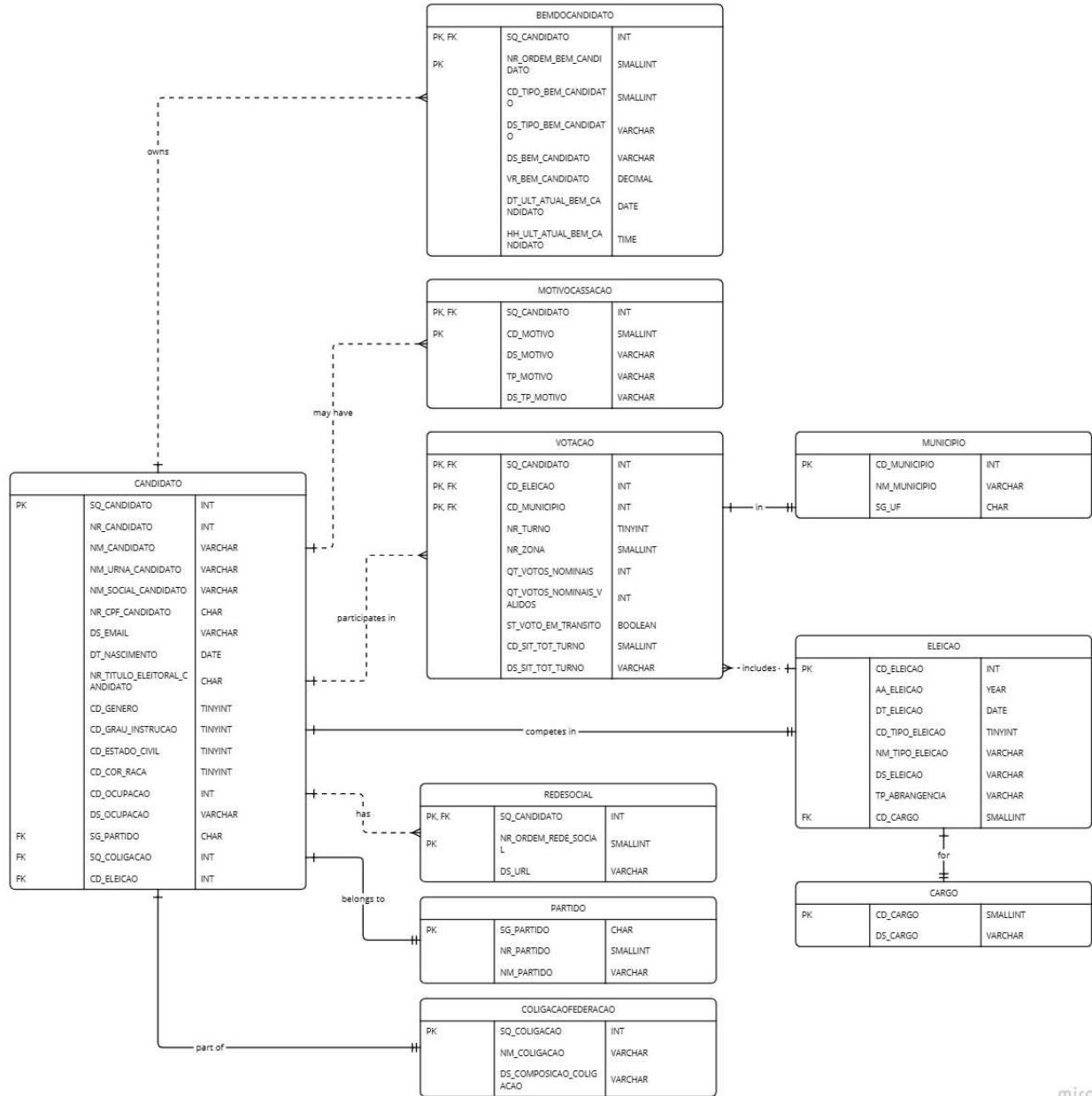
Armazena os motivos de cassação dos candidatos, se houver.

Atributo	Tipo de Dados	Tamanho	Descrição
SQ_CANDIDATO	INT	-	Identificador do candidato (PK/FK)
CD_MOTIVO	SMALLINT	-	Código do motivo de cassação (PK)
DS_MOTIVO	VARCHAR	200	Descrição do motivo
TP_MOTIVO	VARCHAR	50	Tipo do motivo
DS_TP_MOTIVO	VARCHAR	100	Descrição do tipo do motivo

### 3.4. Ferramenta de Modelagem:

A ferramenta MIRO foi utilizada para construir o DER.

## Diagrama Entidade-Relacionamento (DER)



miro

#### **4. Estudo de Volumetria :**

Para realizar um estudo de volumetria para um banco de dados relacional focando apenas na votação dos municípios do estado de Minas Gerais, devo avaliar o volume de dados que será armazenado em cada tabela principal, considerando o número de registros esperados e o tamanho de cada atributo nas tabelas principais relacionadas à votação.

Abaixo, segue o processo detalhado de como calculei a volumetria.

##### **4. 1. Identificar as Principais Tabelas**

Para um banco de dados relacional focado na votação dos municípios de Minas Gerais, considero as tabelas:

- **Candidato** (informações dos candidatos)
- **Eleição** (dados da eleição)
- **Cargo** (informações sobre cargos concorridos)
- **Partido** (dados dos partidos políticos)
- **Coligação** (informações de coligações)
- **Município** (dados dos municípios de Minas Gerais)
- **Votação** (dados da votação, que será a tabela mais volumosa)

##### **4.2. Definir Estimativas de Registro e Tamanho Médio**

Para calcular a volumetria de cada tabela, utilizo a fórmula:

$$\text{Volume Estimado} = \text{Nº de Registros} \times \text{Tamanho Médio por Registro}$$

##### **4.3. Volume Estimado por Tabela**

###### **Tabela Candidato**

- **Nº de Registros:** Suponha que Minas Gerais tenha aproximadamente 10.000 candidatos únicos ao longo dos anos.
- **Tamanho Médio por Registro:** Aproximadamente 250 bytes (considerando todos os atributos como nome, número, CPF, gênero, ocupação, etc.).
- **Volume Estimado:**  $10.000 \times 250 = 2.500.000$  bytes ou  $\sim 2.5$  MB.

###### **Tabela Eleição**

- **Nº de Registros:** Suponha que existam 5 eleições (considerando eleições municipais e estaduais ao longo de alguns anos).
- **Tamanho Médio por Registro:** Aproximadamente 100 bytes.
- **Volume Estimado:**  $5 \times 100 = 500$  bytes ou  $\sim 0.5$  KB.

###### **Tabela Cargo**

- **Nº de Registros:** Aproximadamente 20 tipos de cargos.
- **Tamanho Médio por Registro:** 50 bytes.
- **Volume Estimado:**  $20 \times 50 = 1.000$  bytes ou  $\sim 1$  KB.

###### **Tabela Partido**

- **Nº de Registros:** Suponha 35 partidos políticos.
- **Tamanho Médio por Registro:** 60 bytes.
- **Volume Estimado:**  $35 \times 60 = 2.100$  bytes ou  $\sim 2$  KB.

### Tabela Coligação

- **Nº de Registros:** Suponha 50 coligações registradas ao longo dos anos.
- **Tamanho Médio por Registro:** Aproximadamente 150 bytes.
- **Volume Estimado:**  $50 \times 150 = 7.500$  bytes ou ~7.5 KB.

### Tabela Município

- **Nº de Registros:** Minas Gerais possui 853 municípios.
- **Tamanho Médio por Registro:** 100 bytes (código do município, nome, e UF).
- **Volume Estimado:**  $853 \times 100 = 85.300$  bytes ou ~85 KB.

### Tabela Município

A tabela de votação representa o maior volume de dados, pois armazena a combinação de dados de candidatos, eleições, turnos, e municípios.

- **Nº de Registros:**
  - Nº de Candidatos: 10.000 (ao longo dos anos).
  - Nº de Eleições: 5.
  - Nº de Turnos: 2 (1º e 2º turnos).
  - Nº de Municípios: 853 (total de municípios em Minas Gerais).

$$\text{Nº de Registros} = 10.000 \times 5 \times 2 \times 853 = 85.300.000$$

- **Tamanho Médio por Registro:**

Aproximadamente 50 bytes (contendo identificador de candidato, eleição, município, turno, e número de votos).

- **Volume Estimado:**

$$85.300.000 \times 50 = 4.265.000.000 \text{ bytes} \approx 4.3 \text{ GB}$$

## Resumo da Volumetria

Tabela	Nº de Registros	Tamanho Médio por Registro	Volume Estimado
Candidato	10.000	250 bytes	2.5 MB
Eleição	5	100 bytes	0.5 KB
Cargo	20	50 bytes	1 KB
Partido	35	60 bytes	2 KB
Coligação	50	150 bytes	7.5 KB
Município	853	100 bytes	85 KB
Votação	85.300.000	50 bytes	4.3 GB

## Total Estimado de Volume de Dados

**Total Estimado:** Aproximadamente **4.3 GB**.

## 5. Modelagem Física:

### 5.1. Geração do Script SQL (DDL):

O script inclui comandos para:

- Criar as tabelas com seus respectivos atributos e tipos de dados.
- Definir as chaves primárias e estrangeiras.
- Criar índices para otimizar as consultas.
- Estabelecer restrições de integridade (NOT NULL, UNIQUE, CHECK).

```
-- Criação do banco de dados
CREATE DATABASE EleicaoDB;
GO

-- Configuração de uso do banco de dados
USE EleicaoDB;
GO

-- Configuração de charset e collation padrão
ALTER DATABASE EleicaoDB COLLATE Latin1_General_CI_AS;
GO

-- Criação das tabelas com comentários
-- Tabela Candidato
CREATE TABLE Candidato (
    SQ_CANDIDATO INT PRIMARY KEY, -- Identificador único do candidato
    NR_CANDIDATO INT NOT NULL, -- Número do candidato
    NM_CANDIDATO VARCHAR(100) NOT NULL, -- Nome completo do candidato
    NM_URNA_CANDIDATO VARCHAR(50) NOT NULL, -- Nome do candidato na urna
    NM_SOCIAL_CANDIDATO VARCHAR(50), -- Nome social do candidato
    NR CPF_CANDIDATO CHAR(11) UNIQUE NOT NULL, -- CPF do candidato (único)
    DS_EMAIL VARCHAR(100), -- Email do candidato
    DT_NASCIMENTO DATE, -- Data de nascimento do candidato
    NR_TITULO_ELEITORAL_CANDIDATO CHAR(12) UNIQUE, -- Número do título eleitoral (único)
    CD_GENERO TINYINT CHECK (CD_GENERO IN (2, 4)), -- Código do gênero (2: Masculino, 4: Feminino)
    CD_GRAU_INSTRUCAO TINYINT, -- Código do grau de instrução
    CD_ESTADO_CIVIL TINYINT, -- Código do estado civil
    CD_COR_RACA TINYINT, -- Código da cor/raça
    CD_OCUPACAO INT, -- Código da ocupação
    DS_OCUPACAO VARCHAR(50), -- Descrição da ocupação
    SG_PARTIDO CHAR(5), -- Sigla do partido
    SQ_COLIGACAO INT, -- Sequencial da coligação
    CD_ELEICAO INT, -- Código da eleição
        CONSTRAINT FK_Candidato_Partido FOREIGN KEY (SG_PARTIDO) REFERENCES Partido(SG_PARTIDO),
        CONSTRAINT FK_Candidato_Coligacao FOREIGN KEY (SQ_COLIGACAO) REFERENCES Coligacao_Federacao(SQ_COLIGACAO),
        CONSTRAINT FK_Candidato_Eleicao FOREIGN KEY (CD_ELEICAO) REFERENCES Eleicao(CD_ELEICAO)
);
GO

-- Tabela Rede Social
CREATE TABLE Rede_Social (
    SQ_CANDIDATO INT NOT NULL, -- Identificador do candidato
    NR_ORDEM_REDE_SOCIAL SMALLINT NOT NULL, -- Número da ordem da rede social
    DS_URL VARCHAR(200) NOT NULL, -- URL do perfil da rede social
    PRIMARY KEY (SQ_CANDIDATO, NR_ORDEM_REDE_SOCIAL),
    CONSTRAINT FK_RedeSocial_Candidato FOREIGN KEY (SQ_CANDIDATO) REFERENCES Candidato(SQ_CANDIDATO)
);
GO

-- Tabela Bem do Candidato
CREATE TABLE Bem_Candidato (
    SQ_CANDIDATO INT NOT NULL, -- Identificador do candidato
    NR_ORDEM_BEM_CANDIDATO SMALLINT NOT NULL, -- Número da ordem do bem
    CD_TIPO_BEM_CANDIDATO SMALLINT, -- Código do tipo de bem
    DS_TIPO_BEM_CANDIDATO VARCHAR(50), -- Descrição do tipo de bem
```

```

DS_BEM_CANDIDATO VARCHAR(200), -- Descrição do bem
VR_BEM_CANDIDATO DECIMAL(15, 2), -- Valor do bem
DT_ULT_ATUAL_BEM_CANDIDATO DATE, -- Data da última atualização do bem
HH_ULT_ATUAL_BEM_CANDIDATO TIME, -- Hora da última atualização do bem
PRIMARY KEY (SQ_CANDIDATO, NR_ORDEM_BEM_CANDIDATO),
CONSTRAINT FK_BemCandidato_Candidato FOREIGN KEY (SQ_CANDIDATO) REFERENCES
Candidato(SQ_CANDIDATO)
);
GO

-- Tabela Eleição
CREATE TABLE Eleicao (
    CD_ELEICAO INT PRIMARY KEY, -- Código da eleição
    AA_ELEICAO YEAR NOT NULL, -- Ano da eleição
    DT_ELEICAO DATE NOT NULL, -- Data da eleição
    CD_TIPO_ELEICAO TINYINT, -- Código do tipo de eleição
    NM_TIPO_ELEICAO VARCHAR(50), -- Nome do tipo de eleição
    DS_ELEICAO VARCHAR(100), -- Descrição da eleição
    TP_ABRANGENCIA VARCHAR(20) -- Abrangência da eleição (Municipal, Estadual, Federal)
);
GO

-- Tabela Cargo
CREATE TABLE Cargo (
    CD_CARGO SMALLINT PRIMARY KEY, -- Código do cargo
    DS_CARGO VARCHAR(50) NOT NULL -- Descrição do cargo
);
GO

-- Tabela Partido
CREATE TABLE Partido (
    SG_PARTIDO CHAR(5) PRIMARY KEY, -- Sigla do partido
    NR_PARTIDO SMALLINT UNIQUE NOT NULL, -- Número do partido (único)
    NM_PARTIDO VARCHAR(50) NOT NULL -- Nome do partido
);
GO

-- Tabela Coligação/Federação
CREATE TABLE Coligacao_Federacao (
    SQ_COLIGACAO INT PRIMARY KEY, -- Sequencial da coligação/federação
    NM_COLIGACAO VARCHAR(100), -- Nome da coligação/federação
    DS_COMPOSICAO_COLIGACAO VARCHAR(200) -- Descrição da composição da coligação
);
GO

-- Tabela Votação
CREATE TABLE Votacao (
    SQ_CANDIDATO INT NOT NULL, -- Identificador do candidato
    CD_ELEICAO INT NOT NULL, -- Código da eleição
    CD_MUNICIPIO INT NOT NULL, -- Código do município
    NR_TURNO TINYINT, -- Número do turno
    NR_ZONA SMALLINT, -- Número da zona eleitoral
    QT_VOTOS_NOMINAIS INT, -- Quantidade de votos nominais
    QT_VOTOS_NOMINAIS_VALIDOS INT, -- Quantidade de votos válidos
    ST_VOTO_EM_TRANSITO BIT, -- Indicador de voto em trânsito (1 = Sim, 0 = Não)
    CD_SIT_TOT_TURNO SMALLINT, -- Código da situação total do turno
    DS_SIT_TOT_TURNO VARCHAR(50), -- Descrição da situação total do turno
    PRIMARY KEY (SQ_CANDIDATO, CD_ELEICAO, CD_MUNICIPIO),
    CONSTRAINT FK_Votacao_Candidato FOREIGN KEY (SQ_CANDIDATO) REFERENCES
    Candidato(SQ_CANDIDATO),
    CONSTRAINT FK_Votacao_Eleicao FOREIGN KEY (CD_ELEICAO) REFERENCES
    Eleicao(CD_ELEICAO),
    CONSTRAINT FK_Votacao_Municipio FOREIGN KEY (CD_MUNICIPIO) REFERENCES
    Municipio(CD_MUNICIPIO)
);
GO

-- Tabela Município
CREATE TABLE Municipio (
    CD_MUNICIPIO INT PRIMARY KEY, -- Código do município
    NM_MUNICIPIO VARCHAR(100) NOT NULL, -- Nome do município

```

```

    SG_UF CHAR(2) NOT NULL -- Sigla da unidade federativa
);

GO

-- Tabela Motivo Cassação
CREATE TABLE Motivo_Cassacao (
    SQ_CANDIDATO INT NOT NULL, -- Identificador do candidato
    CD_MOTIVO SMALLINT NOT NULL, -- Código do motivo de cassação
    DS_MOTIVO VARCHAR(200), -- Descrição do motivo
    TP_MOTIVO VARCHAR(50), -- Tipo do motivo
    DS_TP_MOTIVO VARCHAR(100), -- Descrição do tipo do motivo
    PRIMARY KEY (SQ_CANDIDATO, CD_MOTIVO),
    CONSTRAINT FK_MotivoCassacao_Candidato FOREIGN KEY (SQ_CANDIDATO) REFERENCES
Candidato(SQ_CANDIDATO)
);
GO

-- Índices para otimização das consultas
CREATE INDEX IDX_Candidato_NR_CANDIDATO ON Candidato(NR_CANDIDATO);
CREATE INDEX IDX_Candidato_NR_CPF_CANDIDATO ON Candidato(NR_CPF_CANDIDATO);
CREATE INDEX IDX_Eleicao_AA_ELEICAO ON Eleicao(AA_ELEICAO);
CREATE INDEX IDX_Votacao_CD_MUNICIPIO ON Votacao(CD_MUNICIPIO);

```

## 6. Modelagem Multidimensional :

### 6.1. Definição do Esquema:

#### 6.2. Fato:

**Tabela Fato:** Votação

A tabela de fatos é a principal e contém métricas numéricas que podem ser analisadas em diferentes dimensões.

- **Granularidade:**  
O nível mais granular é por candidato, eleição, município e turno.
- **Métricas:**  
QT\_VOTOS\_NOMINAIS: Quantidade de votos nominais  
QT\_VOTOS\_NOMINAIS\_VALIDOS: Quantidade de votos válidos  
ST\_VOTO\_EM\_TRANSITO: Indicador de voto em trânsito, que pode ser agregado ou filtrado em análises

### 6.3. Dimensões:

As dimensões permitem analisar os dados eleitorais sob diferentes perspectivas.

- **Dimensão Candidato**

Para que eu possa analisar os dados de votação por atributos específicos dos candidatos.

**Atributos:**

- SQ\_CANDIDATO,
- NM\_CANDIDATO,
- NM\_URNA\_CANDIDATO,
- NM\_SOCIAL\_CANDIDATO,
- NR\_CPF\_CANDIDATO,
- CD\_GENERO,
- CD\_GRAU\_INSTRUCAO,
- CD\_ESTADO\_CIVIL,
- CD\_COR\_RACA,
- CD\_OCUPACAO,
- SG\_PARTIDO.

- **Dimensão Eleição**

Permite analisar os votos por características específicas de cada eleição.

**Atributos:**

- CD\_ELEICAO,
- AA\_ELEICAO,
- CD\_TIPO\_ELEICAO,
- NM\_TIPO\_ELEICAO,
- DS\_ELEICAO,
- TP\_ABRANGENCIA.

- **Dimensão Cargo**

Representa os cargos aos quais os candidatos estão concorrendo.

**Atributos:**

- CD\_CARGO,
- DS\_CARGO.

- **Dimensão Partido**

Informa a composição partidária dos candidatos.

**Atributos:**

- SG\_PARTIDO,
- NR\_PARTIDO,
- NM\_PARTIDO.

- **Dimensão Coligação/Federação**

Informa as coligações ou federações às quais os candidatos estão afiliados.

**Atributos:**

- SQ\_COLIGACAO,
- NM\_COLIGACAO,
- DS\_COMPOSICAO\_COLIGACAO.

- **Dimensão Município**

Permite análise por localização geográfica.

**Atributos:**

- CD\_MUNICIPIO,
- NM\_MUNICIPIO,
- SG\_UF.

- **Dimensão Tempo**

Ajuda a analisar a variação de votos ao longo do tempo.

**Atributos:**

- Data Completa,
- Ano,
- Mês,
- Dia da Semana

#### **6.4. Escolha do Formato:**

Para criar um modelo de dados para um data warehouse com os dados fornecidos, adoto a estrutura em estrela (**star schema**), que é amplamente usada para facilitar a consulta em análises OLAP (Processamento Analítico Online). O modelo contém uma tabela central de **fatos** que armazena as métricas de interesse, conectada a várias tabelas **dimensão**, que descrevem os atributos contextuais.

#### **6.5. Hierarquias:**

As hierarquias permitem a criação de estruturas para navegação e agregação de dados, facilitando a análise em diferentes níveis de granularidade. As hierarquias definidas para cada dimensão são:

- **Dimensão Candidato:** Gênero > Raça > Grau de Instrução
- **Dimensão Eleição:** Ano da Eleição > Tipo de Eleição
- **Dimensão Cargo:** Tipo de Cargo > Cargo
- **Dimensão Município:** Estado > Município
- **Dimensão Tempo:** Ano > Mês > Dia

## 6.6. Volumetria:

O estudo de volumetria envolve cálculos e estimativas para determinar o espaço que as tabelas ocuparão no disco, bem como as taxas de crescimento ao longo do tempo. Abaixo estão as etapas e considerações para realizar o estudo de volumetria no contexto de um data warehouse para dados eleitorais.

### 6.6.1. Definir os Valores para Cada Dimensão e Tabela de Fatos

Realizo uma estimativa inicial para cada uma das dimensões e para a tabela de fatos com base nos dados de votações de Minas Gerais.

### 6.6.2. Estimar o Volume de Dados

Para estimar a volumetria, uso a seguinte fórmula:

$$\text{Volume} = \text{Nº de Registros} \times \text{Tamanho Médio por Registro}$$

## Estimativas Iniciais

### Tabelas de Dimensão

1. **Dim\_Candidato**
  - **Nº de Registros:** Estimo que Minas Gerais tenha aproximadamente 10.000 candidatos únicos para diferentes eleições.
  - **Tamanho Médio por Registro:** Aproximadamente 250 bytes.
  - **Volume Estimado:**  $10.000 \times 250 = 2.500.000$  bytes ou ~2.5 MB.
2. **Dim\_Eleicao**
  - **Nº de Registros:** Estimo que existam 5 eleições (considerando eleições municipais e estaduais, com diferentes tipos de eleição ao longo dos anos).
  - **Tamanho Médio por Registro:** 100 bytes.
  - **Volume Estimado:**  $5 \times 100 = 500$  bytes ou ~0.5 KB.
3. **Dim\_Cargo**
  - **Nº de Registros:** Estimo aproximadamente 20 tipos de cargos.
  - **Tamanho Médio por Registro:** 50 bytes.
  - **Volume Estimado:**  $20 \times 50 = 1.000$  bytes ou ~1 KB.
4. **Dim\_Partido**
  - **Nº de Registros:** Estimo 35 partidos.
  - **Tamanho Médio por Registro:** 60 bytes.
  - **Volume Estimado:**  $35 \times 60 = 2.100$  bytes ou ~2 KB.
5. **Dim\_Coligacao**
  - **Nº de Registros:** Estimo 50 coligações ao longo dos anos.
  - **Tamanho Médio por Registro:** 150 bytes.
  - **Volume Estimado:**  $50 \times 150 = 7.500$  bytes ou ~7.5 KB.
6. **Dim\_Municipio**
  - **Nº de Registros:** Minas Gerais possui 853 municípios.
  - **Tamanho Médio por Registro:** 100 bytes.
  - **Volume Estimado:**  $853 \times 100 = 85.300$  bytes ou ~85 KB.

## 7. Dim\_Tempo

- **Nº de Registros:** Considerando 10 anos e 365 dias por ano, tenho aproximadamente 3.650 dias.
- **Tamanho Médio por Registro:** 20 bytes.
- **Volume Estimado:**  $3.650 \times 20 = 73.000$  bytes ou ~73 KB.

## Tabela de Fatos (Fato\_Votacao)

- **Nº de Registros:** A tabela de fatos geralmente é a maior, pois registra a combinação de todas as dimensões. Para calcular o número de registros, multiplico as possibilidades de combinação dos dados:

$\text{Nº de Registros} = \text{Nº de Candidatos} \times \text{Nº de Eleições} \times \text{Nº de Turnos} \times \text{Nº de Municípios}$

- Nº de Candidatos: 10.000 (considero todos os candidatos ao longo dos anos).
- Nº de Eleições: 5.
- Nº de Turnos: 2 (considerando 1º e 2º turnos).
- Nº de Municípios: 853 (total de municípios em Minas Gerais).

$\text{Nº de Registros} = 10.000 \times 5 \times 2 \times 853 = 85.300.000$  registros

- **Tamanho Médio por Registro:** Aproximadamente 40 bytes.
- **Volume Estimado:**

$85.300.000 \times 40 = 3.412.000.000$  bytes ≈ 3.4 GB

## Resumo da Volumetria

Tabela	Nº de Registros	Tamanho Médio por Registro	Volume Estimado
Dim_Candidato	10.000	250 bytes	2.5 MB
Dim_Eleicao	5	100 bytes	0.5 KB
Dim_Cargo	20	50 bytes	1 KB
Dim_Partido	35	60 bytes	2 KB
Dim_Coligacao	50	150 bytes	7.5 KB
Dim_Municipio	853	100 bytes	85 KB
Dim_Tempo	3.650	20 bytes	73 KB
Fato_Votacao	85.300.000	40 bytes	3.4 GB

## Total Estimado de Volume de Dados

**Total Estimado de Volume de Dados:** Aproximadamente 3.5 GB

## 6.7. Documentação:

Projeto um modelo de dados multidimensional para analisar dados eleitorais, focando nas interações de candidatos, eleições, partidos e localidades. A tabela fato centraliza as métricas de votação, enquanto as dimensões permitem que essas métricas sejam analisadas sob diversas perspectivas, como gênero, partido, cargo e localidade.

## Tabelas e Atributos Documentados

Tabela	Descrição	Principais Atributos
Fato_Votacao	Tabela fato principal que armazena os dados de votação por candidato, eleição e localidade.	QT_VOTOS_NOMINAIS , QT_VOTOS_VALIDOS
Dim_Candidato	Dimensão para análise por informações do candidato.	SQ_CANDIDATO , NM_CANDIDATO , CD_OUPACAO
Dim_Eleicao	Dimensão para análise de informações sobre as eleições.	CD_ELEICAO , AA_ELEICAO , CD_TIPO_ELEICAO
Dim_Cargo	Dimensão para análise dos cargos em disputa.	CD_CARGO , DS_CARGO
Dim_Partido	Dimensão que armazena os partidos dos candidatos.	SG_PARTIDO , NM_PARTIDO
Dim_Coligacao	Dimensão que armazena informações sobre as coligações dos candidatos.	SQ_COLIGACAO , NM_COLIGACAO
Dim_Municipio	Dimensão para análise geográfica por município e estado.	CD_MUNICIPIO , SG_UF
Dim_Tempo	Dimensão temporal para análise ao longo do tempo.	Ano , Mês , Dia

### 6.8 Geração do Script SQL (DDL):

O script deve incluir comandos para:

- Criar as tabelas com seus respectivos atributos e tipos de dados no modelo star schema.
- Definir as chaves primárias e estrangeiras.
- Criar índices para otimizar as consultas.
- Estabelecer restrições de integridade (NOT NULL, UNIQUE, CHECK).

```
-- Criação do banco de dados
CREATE DATABASE EleicoesDW
GO

-- Seleção do banco de dados
USE EleicoesDW
GO

-- Configuração de charset e collation
ALTER DATABASE EleicoesDW COLLATE Latin1_General_CI_AS
GO

-- Criação da Tabela de Fatos (Fato_Votacao)
CREATE TABLE Fato_Votacao (
    SQ_CANDIDATO INT NOT NULL,
    CD_ELEICAO INT NOT NULL,
    CD_CARGO SMALLINT NOT NULL,
    CD_MUNICIPIO INT NOT NULL,
    NR_TURNO TINYINT NOT NULL,
    QT_VOTOS_NOMINAIS INT NOT NULL,
    QT_VOTOS_VALIDOS INT NOT NULL,
    ST_VOTO_EM_TRANSITO BIT NOT NULL DEFAULT 0,
    CD_SIT_TOT_TURNO SMALLINT NOT NULL,
    DS_SIT_TOT_TURNO VARCHAR(50),
    PRIMARY KEY (SQ_CANDIDATO, CD_ELEICAO, CD_CARGO, CD_MUNICIPIO, NR_TURNO),
    FOREIGN KEY (SQ_CANDIDATO) REFERENCES Dim_Candidato(SQ_CANDIDATO),
    FOREIGN KEY (CD_ELEICAO) REFERENCES Dim_Eleicao(CD_ELEICAO),
    FOREIGN KEY (CD_CARGO) REFERENCES Dim_Cargo(CD_CARGO),
    FOREIGN KEY (CD_MUNICIPIO) REFERENCES Dim_Municipio(CD_MUNICIPIO)
);
GO
```

```

-- Criação das Dimensões
CREATE TABLE Dim_Candidato (
    SQ_CANDIDATO INT PRIMARY KEY,
    NM_CANDIDATO VARCHAR(100) NOT NULL,
    NM_URNA_CANDIDATO VARCHAR(50),
    NM_SOCIAL_CANDIDATO VARCHAR(50),
    NR_CPF_CANDIDATO CHAR(11),
    DS_EMAIL VARCHAR(100),
    DT_NASCIMENTO DATE,
    NR_TITULO_ELEITORAL_CANDIDATO CHAR(12),
    CD_GENERO TINYINT CHECK (CD_GENERO IN (2, 4)),
    CD_GRAU_INSTRUCAO TINYINT,
    CD_ESTADO_CIVIL TINYINT,
    CD_COR_RACA TINYINT,
    CD_OCUPACAO INT,
    DS_OCUPACAO VARCHAR(50),
    SG_PARTIDO CHAR(5),
    SQ_COLIGACAO INT,
    FOREIGN KEY (SG_PARTIDO) REFERENCES Dim_Partido(SG_PARTIDO),
    FOREIGN KEY (SQ_COLIGACAO) REFERENCES Dim_Coligacao(SQ_COLIGACAO)
);
GO

CREATE TABLE Dim_Eleicao (
    CD_ELEICAO INT PRIMARY KEY,
    AA_ELEICAO YEAR NOT NULL,
    DT_ELEICAO DATE,
    CD_TIPO_ELEICAO TINYINT CHECK (CD_TIPO_ELEICAO IN (1, 2, 3)),
    NM_TIPO_ELEICAO VARCHAR(50),
    DS_ELEICAO VARCHAR(100),
    TP_ABRANGENCIA VARCHAR(20) CHECK (TP_ABRANGENCIA IN ('Municipal', 'Estadual',
    'Federal'))
);
GO

CREATE TABLE Dim_Cargo (
    CD_CARGO SMALLINT PRIMARY KEY,
    DS_CARGO VARCHAR(50) NOT NULL
);
GO

CREATE TABLE Dim_Partido (
    SG_PARTIDO CHAR(5) PRIMARY KEY,
    NR_PARTIDO SMALLINT UNIQUE NOT NULL,
    NM_PARTIDO VARCHAR(50) NOT NULL
);
GO

CREATE TABLE Dim_Coligacao (
    SQ_COLIGACAO INT PRIMARY KEY,
    NM_COLIGACAO VARCHAR(100),
    DS_COMPOSICAO_COLIGACAO VARCHAR(200)
);
GO

CREATE TABLE Dim_Municipio (
    CD_MUNICIPIO INT PRIMARY KEY,
    NM_MUNICIPIO VARCHAR(100) NOT NULL,
    SG_UF CHAR(2) NOT NULL
);
GO

CREATE TABLE Dim_Tempo (
    Data DATE PRIMARY KEY,
    Ano AS YEAR(Data),
    Mes AS MONTH(Data),
    Dia AS DAY(Data)
);
GO

```

```
-- Índices para otimização de consultas
CREATE INDEX idx_candidato_nome ON Dim_Candidato (NM_CANDIDATO);
CREATE INDEX idx_municipio_nome ON Dim_Municipio (NM_MUNICIPIO);
CREATE INDEX idx_eleicao_tipo ON Dim_Eleicao (CD_TIPO_ELEICAO);
CREATE INDEX idx_cargo_descricao ON Dim_Cargo (DS_CARGO);

-- Comentários de Documentação
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Fato que armazena a votação dos candidatos por eleição e município', 'SCHEMA', 'dbo', 'TABLE', 'Fato_Votacao';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações dos candidatos', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Candidato';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações das eleições', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Eleicao';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações dos cargos', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Cargo';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações dos partidos', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Partido';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações das coligações', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Coligacao';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações dos municípios', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Municipio';
EXEC sp_addextendedproperty 'MS_Description', 'Tabela Dimensão que armazena informações de tempo', 'SCHEMA', 'dbo', 'TABLE', 'Dim_Tempo';

-- Fim do Script
```