Codebook - Dados Individuais dos Deputados

Dataset: Empirical Support for "Deputies' assets and their legislative activity (2023-2025)" - Dados Individuais

Autor: Pedro Lenhagui Bergamaschi (UFPR)

Data: 03 de Outubro de 2025

Introdução

Este documento detalha a estrutura dos arquivos individuais de cada deputado da amostra (1% mais rico da 57ª legislatura). Os dados estão organizados em diferentes formatos (CSV, JSON, PDF) e provenientes de múltiplas fontes governamentais.

Estrutura de nomenclatura dos arquivos:

- `[NomeDeputado]` = Nome do parlamentar sem acentuação
- `[CNPJ]` = Cadastro Nacional de Pessoa Jurídica (14 dígitos)
- `[nome-empresa-abreviado]` = Nome simplificado da empresa

Arquivo 1: Declaração de Bens - `bens-[nome-deputado]-2022.csv`

Fonte: Tribunal Superior Eleitoral (TSE)

Descrição: Declaração detalhada dos bens do candidato na eleição de 2022, conforme autodeclarado no sistema CandEx.

Variáveis

Nome da Variável	Descrição	Tipo	Exemplo
ANO_ELEICAO	Ano da eleição a que os dados se referem	Numérico	2022
SQ_CANDIDATO	Sequencial único do candidato (chave primária)	Numérico	60001643110
NR_ORDEM_BEM_CANDIDATO	Número sequencial do bem na declaração	Numérico	1, 2, 3...
DS_TIPO_BEM_CANDIDATO	Categoria do bem declarado	Texto	Apartamento, Veículo, Aplicação financeira
DS_BEM_CANDIDATO	Descrição detalhada do bem em texto livre	Texto	UM IMOVEL COMERCIAL SITUADO NA RUA...
VR_BEM_CANDIDATO	Valor do bem em Reais (R$)	Numérico	1262827.00
NM_CANDIDATO	Nome completo do candidato conforme registro no TSE	Texto	EUNICIO LOPES DE OLIVEIRA

Notas:

- Cada linha representa um bem individual
- Os valores estão em reais nominais (não corrigidos pela inflação)
- A descrição é texto livre fornecido pelo candidato

Arquivo 2: Proposições Legislativas - `[deputado]_json-CD-projetos.json`

Fonte: Câmara dos Deputados - API de Dados Abertos

Descrição: Proposições legislativas de autoria do deputado apresentadas durante a 57ª legislatura.

Estrutura do JSON

```
{
  "dados": [
    {
      "id": 2350927,
      "uri": "https://dadosabertos.camara.leg.br/api/v2/proposicoes/2350927",
      "siglaTipo": "PL",
      "codTipo": 139,
      "numero": 1048,
      "ano": 2023,
      "ementa": "Altera a Lei Complementar nº 123..."
    }
  ],
  "links": [...]
}
```

Variáveis em `dados`

Nome da Variável	Descrição	Tipo	Exemplo
`id`	ID único da proposição na Câmara	Numérico	`2350927`
`uri`	URL da API para detalhes da proposição	Texto (URL)	`https://dadosabertos.camara.leg.br/...`
`siglaTipo`	Sigla do tipo de proposição	Texto	`PL, PLP, PEC, PDC, PLV`
`codTipo`	Código numérico do tipo	Numérico	`135, 136, 139, 140, 390`
`numero`	Número da proposição	Numérico	`1048`
`ano`	Ano de apresentação	Numérico	`2023, 2024, 2025`
`ementa`	Resumo do objetivo da proposição	Texto	`Altera a Lei Complementar nº 123...`

Tipos de Proposição:

- `135` = PDC (Projeto de Decreto Legislativo)
- `136` = PEC (Proposta de Emenda à Constituição)
- `139` = PL (Projeto de Lei)
- `140` = PLP (Projeto de Lei Complementar)
- `390` = PLV (Projeto de Lei de Conversão)

Variáveis em `links`

Nome da Variável	Descrição	Tipo
`rel`	Tipo de relação do link	Texto
`href`	URL do endpoint	Texto (URL)

Tipos de `rel`:

- `self` = Própria consulta
- `next` = Próxima página
- `first` = Primeira página
- `last` = Última página

Notas:

- A API retorna resultados paginados (máximo 15 itens por página)
- Para obter todos os projetos, é necessário seguir os links `next` até `last`
- Somente proposições dos tipos especificados foram coletadas

---

Arquivo 3: Rede de Empresas (JSON) - `[deputado]_json-rede-empresas.json`

Fonte: Receita Federal do Brasil (processado para visualização)

Descrição: Estrutura de dados para visualização em grafo das conexões empresariais.

Estrutura do JSON


```
{
  "no": [
    {
      "id": "PF_***063049**-DUDA BRITO RAMOS",
      "descricao": "DUDA BRITO RAMOS",
      "camada": 0,
      "sexo": 2,
      "imagem": "icone-grafo-feminino.png",
      "cor": "orange",
      "posicao": {"x": 448.79, "y": -793.39},
      "pinado": true
    }
  ],
  "ligacao": [
    {
      "origem": "PF_***063049**-DUDA BRITO RAMOS",
      "destino": "PJ_34894844000145",
      "label": "Sócio",
```

```
      "cor": "silver",
      "camada": 0,
      "id": 5
    }
  ]
}
```

Variáveis em `no` (Nós)

Nome da Variável	Descrição	Tipo	Valores Possíveis
`id`	Identificador único do nó	Texto	`PF_[CPF]-[NOME]` ou `PJ_[CNPJ]`
`descricao`	Nome da pessoa ou razão social	Texto	Nome completo
`camada`	Nível de distância do deputado	Numérico	`0` (deputado), `1`, `2`, `3`
`sexo`	Sexo da pessoa física	Numérico	`0` (desconhecido), `1` (masculino), `2` (feminino)
`tipo`	Tipo de nó (apenas para PJ)	Numérico	`0` (empresa comum)
`situacao_ativa`	Se a empresa está ativa (apenas PJ)	Booleano	`true`, `false`
`logradouro`	Endereço da empresa (apenas PJ)	Texto	`AVENIDA ANDRE ARAUJO, 2721`
`logradouro_complemento`	Complemento do endereço (apenas PJ)	Texto	`LOTEAMENTO PARQUE-ALEIXO`
`municipio`	Município (apenas PJ)	Texto	`MANAUS`
`uf`	UF (apenas PJ)	Texto	`AM`
`cod_nat_juridica`	Código da natureza jurídica (apenas PJ)	Texto	`2062`, `4090`
`nome_fantasia`	Nome fantasia (apenas PJ)	Texto	`RAMOS SETE`
`imagem`	Nome do arquivo de ícone para visualização	Texto	`icone-grafo-empresa.png`
`cor`	Cor do nó no grafo	Texto	`orange`, `silver`, vazio
`nota`	Anotações adicionais	Texto	Geralmente vazio
`posicao`	Coordenadas x,y para visualização	Objeto	`{"x": 689.34, "y": -133.51}`
`pinado`	Se a posição está fixada	Booleano	`true`, `false`

Prefixos de ID:

- `PF_` = Pessoa Física
- `PJ_` = Pessoa Jurídica

Variáveis em `ligacao` (Arestas)

Nome da Variável	Descrição	Tipo	Exemplo
`origem`	ID do nó de origem	Texto	`PF_***063049**-DUDA BRITO RAMOS`
`destino`	ID do nó de destino	Texto	`PJ_34894844000145`
`label`	Tipo de relação societária	Texto	`Sócio, Sócio-Administrador, Administrador`
`cor`	Cor da aresta	Texto	`silver`
`camada`	Camada da ligação	Numérico	`0`
`tipoDescricao`	Descrição adicional do tipo	Texto	Geralmente vazio
`id`	ID único da ligação	Numérico	`0, 1, 2...`

Tipos de Relação (`label`):

- `Sócio` = Participação societária sem função administrativa
- `Sócio-Administrador` = Participação com função administrativa
- `Administrador` = Função administrativa sem participação no capital
- `rep-sócio-Pai` = Representante legal (quando sócio é menor)
- `Sócio Menor (Assistido/Representado)` = Sócio menor de idade
- `Sociedade Consorciada` = Empresa membro de consórcio

Notas:

- Camada 0 = Deputado
- Camada 1 = Empresas e pessoas diretamente ligadas ao deputado
- Camada 2 = Empresas e pessoas ligadas à camada 1
- Camada 3 = Empresas e pessoas ligadas à camada 2

Arquivo 4: Visualização de Rede - `[deputado]_visualizacao-rede-empresas.pdf`

Descrição: Diagrama visual das conexões empresariais em formato de grafo.

Conteúdo:

- Representação gráfica dos nós (pessoas físicas e jurídicas)
- Arestas direcionadas indicando relações societárias
- Rótulos nas arestas com tipo de qualificação
- CPFs parcialmente ofuscados (**`*XXXXXX`**)
- CNPJs completos
- Razão social das empresas

Elementos Visuais:

- Círculos/elipses = Nós (pessoas ou empresas)
- Setas = Direção da relação (de sócio para empresa)
- Rótulos nas setas = Qualificação do sócio
- Cor laranja = Deputado analisado
- Cores neutras = Demais pessoas e empresas

Notas:

- Não possui métricas de centralidade de rede
- Posicionamento espacial facilita visualização, mas não indica hierarquia
- Documento estático (não interativo)

Arquivo 5: Contratos Públicos - `Portal-transp_json-[CNPJ]_[nome-empresa].json`

Fonte: Portal da Transparência - API de Dados Abertos

Descrição: Contratos celebrados entre empresas da rede e o governo federal durante a 57ª legislatura (2023-2025).

Estrutura do JSON

Array de objetos, cada um representando um contrato.

Variáveis

Nome da Variável (Chave JSON)	Descrição	Tipo	Exemplo
`id`	ID único do contrato no sistema	Numérico	`668677982`
`numero`	Número do contrato	Texto	`000122022`
`objeto`	Descrição detalhada do objeto contratado	Texto	`Objeto: CONTRATAÇÃO DE EMPRESA PARA...`
`numeroProcesso`	Número do processo administrativo	Texto	`23129008733202149` ou `Sem informação`
`fundamentoLegal`	Fundamentação legal	Texto	Geralmente vazio
`situacaoContrato`	Situação atual do contrato	Texto	`Não se aplica`
`modalidadeCompra`	Modalidade de licitação	Texto	`Tomada de Preços`, `Pregão`, `Dispensa de Licitação`
`dataAssinatura`	Data de assinatura do contrato (AAAA-MM-DD)	Data	`2022-03-28`
`dataPublicacaoDOU`	Data de publicação no DOU	Data	`2022-04-07`
`dataInicioVigencia`	Data de início da vigência	Data	`2022-03-28`
`dataFimVigencia`	Data de término da vigência	Data	`2024-03-28`

valorInicialCompra	Valor inicial contratado em R$	Numérico	2417793.88
valorFinalCompra	Valor final após aditivos em R$	Numérico	562513.09

Subobjetos

compra

Nome da Variável	Descrição	Tipo
numero	Número da licitação	Texto
objeto	Objeto da licitação	Texto
numeroProcesso	Número do processo	Texto
contatoResponsavel	Contato do responsável	Texto

unidadeGestora

Nome da Variável	Descrição	Tipo
codigo	Código da unidade gestora	Texto
nome	Nome da unidade gestora	Texto
descricaoPoder	Poder (Executivo, Legislativo, Judiciário)	Texto

orgaoVinculado

Nome da Variável	Descrição	Tipo
codigoSIAFI	Código SIAFI do órgão	Texto
cnpj	CNPJ do órgão	Texto
sigla	Sigla do órgão	Texto
nome	Nome do órgão	Texto

orgaoMaximo

Nome da Variável	Descrição	Tipo
codigo	Código do órgão máximo	Texto
sigla	Sigla do ministério	Texto
nome	Nome do ministério	Texto

unidadeGestoraCompras

Estrutura idêntica a `unidadeGestora` (pode conter informações ligeiramente diferentes).

fornecedor

Nome da Variável	Descrição	Tipo

`id`	ID do fornecedor no sistema	Numérico
`cpfFormatado`	CPF formatado (quando PF)	Texto
`cnpjFormatado`	CNPJ formatado com pontuação	Texto
`numeroInscricaoSocial`	NIS quando aplicável	Texto
`nome`	Nome ou razão social	Texto
`razaoSocialReceita`	Razão social conforme RFB	Texto
`nomeFantasiaReceita`	Nome fantasia conforme RFB	Texto
`tipo`	Tipo de fornecedor	Texto

Tipos de Fornecedor:

- `Entidades Empresariais Privadas`
- `Pessoa Física`
- `Administração Pública`

Notas:

- Valores monetários em formato decimal com ponto (ex: `1819823.55`)
- Datas no formato ISO 8601 (AAAA-MM-DD)
- Alguns campos podem estar vazios ou conter "Sem informação"
- Contratos podem ter valores inicial e final diferentes devido a aditivos

---

Observações Gerais sobre os Dados

Proteção de Dados Pessoais

De acordo com a Lei Geral de Proteção de Dados (LGPD), os dados fornecidos pela RFB contêm CPFs parcialmente ofuscados no formato `***XXXXXX**`, onde apenas 6 dígitos centrais são visíveis.

Formatação de Valores

- **Datas:** Múltiplos formatos encontrados
  - `DD/MM/AAAA` (TSE, alguns campos RFB)
  - `AAAAMMDD` (data de entrada sociedade)
  - `AAAA-MM-DD` (Portal da Transparência, ISO 8601)
- **Valores Monetários:**
  - TSE: numéricos com ponto decimal (ex: `1262827.00`)
  - RFB: texto com formatação brasileira (ex: `"7.100.000,00"`)
  - Portal Transparência: numéricos com ponto decimal

Codificações de Caracteres

- TSE: Latin1 (ISO-8859-1)
- RFB: UTF-8 (com possíveis caracteres especiais)
- Câmara: UTF-8
- Portal Transparência: UTF-8

Limitações e Avisos

1. **Incompletude:** Nem todos os deputados possuem todos os tipos de arquivo (ex: contratos públicos, caso não tenha nenhum)
2. **Temporalidade:** Dados extraídos entre março e julho de 2025
3. **Dados Autodeclarados:** Patrimônio do TSE é autodeclarado, sujeito a imprecisões
4. **Sigilo Fiscal:** Capital social não necessariamente reflete valor de mercado

---

Metadados de Coleta

Período de Coleta: Março a Julho de 2025

Fontes de Dados:

- TSE: Repositório de Dados Eleitorais
- RFB: Dados Públicos CNPJ
- Câmara dos Deputados: API de Dados Abertos v2
- Portal da Transparência: API de Contratos e Compras

Métodos de Extração:

- APIs REST (Câmara e Portal da Transparência)
- Download de datasets completos (TSE)
- Consultas individuais por CPF/CNPJ (RFB)

Processamento:

- Padronização de nomes para merge
- Agregação de patrimônio por candidato
- Construção de redes até 3º grau
- Geração de visualizações em grafo

Contato e Citação

Para dúvidas sobre os dados ou citação deste trabalho:

Autor: Pedro Lenhagui Bergamaschi
Orcid: https://orcid.org/0000-0002-8834-5470
Instituição: Universidade Federal do Paraná (UFPR)
Programa: Doutorado em Ciência Política
Ano: 2025
E-mail: plberga@gmail.com

Citação sugerida:

> *Bergamaschi, P. L. (2025). Patrimônio de Deputados e sua Atuação Legislativa (2023-2025). Trabalho apresentado no IX Fórum Brasileiro de Pós-Graduação em Ciência Política, São Paulo.*