Despesa do Orçamento Federal (LOA) em CSV e Parquet: dicionário de dados

Esta página descreve os arquivos `loaANO.csv.gz` e `loaANO.parquet` publicados na página de Dados Abertos. Eles trazem a despesa dos Orçamentos Fiscal e da Seguridade Social de um exercício: projeto de lei (PLOA), autógrafo, dotação inicial, dotação atualizada e execução (empenhado, liquidado e pago), com todas as classificações da programação.

Os dois arquivos têm as mesmas linhas e as mesmas colunas. Escolha pelo uso:

Arquivo	Para quem	Formato
`loaANO.csv.gz`	Planilhas e qualquer programa que leia texto	CSV conforme a RFC 4180: separador vírgula, ponto como separador decimal, aspas só quando o campo tem vírgula, aspas ou quebra de linha, codificação UTF-8, primeira linha com os nomes das colunas. Compactado com gzip
`loaANO.parquet`	Python, R, bancos de dados e ferramentas de análise	Apache Parquet, com os tipos de cada coluna (texto, inteiro e decimal), compressão zstd

São gerados junto com o arquivo RDF (`loaANO.zip`) e com os dados do endpoint SPARQL, na mesma execução e a partir da mesma base. Os valores são os mesmos nos quatro formatos.

O que é cada linha

Cada linha é **uma combinação de classificações** da despesa no exercício: esfera, unidade orçamentária, função, subfunção, programa, ação, subtítulo, plano orçamentário, categoria econômica, grupo de natureza de despesa, modalidade de aplicação, elemento de despesa, fonte de recursos, identificador de uso e resultado primário. Não há duas linhas com a mesma combinação.

Os valores de uma linha são a soma dos itens de despesa (`loa:ItemDespesa`) do RDF com essas mesmas classificações. Por isso, a soma de uma coluna de valor no arquivo é igual à soma da mesma propriedade no RDF e no endpoint.

Para obter totais, some os valores e agrupe pelas classificações desejadas. Uma programação costuma ocupar várias linhas.

Colunas

Os códigos são texto e mantêm os zeros à esquerda (`01`, `0001`). Ao abrir o CSV, leia as colunas de código como texto, senão `01000` vira `1000`.

As descrições são as do próprio exercício. Um código cuja descrição mudou ao longo dos anos aparece em cada arquivo com a descrição daquele ano.

Coluna	Tipo no Parquet	Descrição	Exemplo (2025)
exercicio	inteiro (SMALLINT)	Ano do exercício financeiro	2025
esfera_codigo	texto	Esfera orçamentária: 10 = Fiscal, 20 = Seguridade Social	10
esfera_descricao	texto		Orçamento Fiscal
orgao_codigo	texto	Órgão orçamentário da unidade orçamentária	01000
orgao_descricao	texto		Camara dos Deputados
unidade_orcamentaria_codigo	texto	Unidade orçamentária (UO)	01101
unidade_orcamentaria_descricao	texto		Câmara dos Deputados
funcao_codigo	texto	Função	01
funcao_descricao	texto		Legislativa
subfuncao_codigo	texto	Subfunção	031
subfuncao_descricao	texto		Ação Legislativa
programa_codigo	texto	Programa	0034
programa_descricao	texto		Programa de Gestão e Manutenção do Poder Legislativo
acao_codigo	texto	Ação orçamentária	4061
acao_descricao	texto		Processo Legislativo, Fiscalização e Representação Política
subtitulo_codigo	texto	Subtítulo (localizador do gasto). O código só identifica o subtítulo dentro da UO e da ação	5664
subtitulo_descricao	texto		Em Brasília - DF
subtitulo_identificador	texto	Identificador único do subtítulo no exercício (o mesmo do RDF)	0110140615664

Coluna	Tipo no Parquet	Descrição	Exemplo (2025)
plano_orcamentario_codigo	texto	Plano orçamentário (PO). Até 2012 não havia PO: o código é @	0001
plano_orcamentario_descricao	texto		Administração Legislativa
plano_orcamentario_identificador	texto	Identificador único do PO no exercício (o mesmo do RDF). Até 2012, termina em @	011010034406156640001
categoria_economica_codigo	texto	Categoria econômica: 3 = Despesas Correntes, 4 = Despesas de Capital, 9 = Reserva de Contingência	3
categoria_economica_descricao	texto		Despesas Correntes
grupo_natureza_despesa_codigo	texto	Grupo de natureza de despesa (GND)	3
grupo_natureza_despesa_descricao	texto		Outras Despesas Correntes
modalidade_aplicacao_codigo	texto	Modalidade de aplicação	90
modalidade_aplicacao_descricao	texto		Aplicações Diretas
elemento_despesa_codigo	texto	Elemento de despesa. 00 nas linhas de autógrafo e dotação (veja abaixo)	00
elemento_despesa_descricao	texto		A Definir
fonte_recursos_codigo	texto	Fonte de recursos: 3 dígitos até 2022 (ex.: 100), 4 dígitos a partir de 2023 (ex.: 1000)	1000
fonte_recursos_descricao	texto		Recursos Livres da União
identificador_uso_codigo	texto	Identificador de uso (IDUSO)	0

Coluna	Tipo no Parquet	Descrição	Exemplo (2025)
identificador_uso_descricao	texto		Recursos não destinados à contrapartida ou à identificação de despesas com ações e serviços públicos de saúde, ou referentes à manutenção e ao desenvolvimento do ensino
resultado_primario_codigo	texto	Resultado primário (RP): classificação atual da programação	2
resultado_primario_descricao	texto		Primária discricionária, considerada no cálculo do RP
valor_projeto_lei	decimal (18,2)	Projeto de Lei Orçamentária (PLOA) enviado pelo Poder Executivo	0.00
valor_dotacao_inicial	decimal (18,2)	Dotação inicial: valor da Lei Orçamentária Anual (LOA) sancionada.	909662640.00
valor_lei_mais_creditos	decimal (18,2)	Dotação atualizada: LOA mais créditos adicionais	894584460.00
valor_empenhado	decimal (18,2)	Empenhado no exercício	0.00
valor_liquidado	decimal (18,2)	Liquidado no exercício. De 2007 em diante, não inclui as despesas liquidadas por inscrição em restos a pagar não processados no encerramento do exercício (veja os cuidados na análise)	0.00

Coluna	Tipo no Parquet	Descrição	Exemplo (2025)
valor_pago	decimal (18,2)	Pago no exercício	0.00
valor_autografo	decimal (18,2)	Autógrafo: valor aprovado pelo Congresso Nacional e enviado à sanção presidencial. **Vazio de 2000 a 2009**, anos para os quais não há registro do autógrafo	909662640.00

Todos os valores estão em reais (R$), com duas casas decimais. Os valores, exceto o autógrafo de 2000 a 2009, nunca são vazios: a ausência de valor é publicada como 0.00.

Cuidados na análise

Dotação e execução ficam em linhas diferentes

O **autógrafo, a dotação inicial e a dotação atualizada** são registrados no elemento de despesa 00. O **PLOA e a execução** (empenhado, liquidado e pago) ficam no elemento detalhado (14, 30, 39 etc.).

Por isso, quem agrupa ou filtra pelo elemento de despesa vê a dotação e a execução em linhas diferentes. Para comparar dotação com execução, some **sem** o elemento de despesa.

Exercícios de 2000 a 2012

- **Plano orçamentário:** não existia até 2012. Nesses anos, o código é @ em todas as linhas.
- **Autógrafo:** vazio de 2000 a 2009.
- **Reserva de contingência de 2000 e 2001:** a modalidade de aplicação tem o código (descrição “Reserva de Contingência”).

Liquidado e restos a pagar não processados

De 2007 em diante, o liquidado não inclui as despesas liquidadas por inscrição em restos a pagar não processados no encerramento do exercício. Assim, o empenhado menos o liquidado de um exercício encerrado é o valor inscrito em restos a pagar não processados. No exercício corrente, essa diferença é o empenhado ainda a liquidar.

De 2000 a 2006, a base de origem não separa a liquidação por inscrição em restos a pagar: o liquidado desses anos é praticamente igual ao empenhado, e a diferença entre os dois não mostra o valor inscrito.

Descrições vazias

Algumas descrições não existem na base de origem e vêm vazias. O código está sempre preenchido.

Como abrir os arquivos

Planilha (Excel, LibreOffice): descompacte o `.csv.gz` (7-Zip, WinRAR ou o próprio sistema) e importe o CSV pela opção de importar texto (no Excel: *Dados > De Texto/CSV*), com origem **UTF-8**, delimitador **vírgula** e as colunas de código como **texto**. Abrir o arquivo com dois cliques costuma usar o ponto e vírgula como separador e perder os zeros à esquerda.

Python (pandas):

```
import pandas as pd

df = pd.read_parquet("loa2025.parquet")                    # tipos já definidos (precisa do pyarrow)
# ou, pelo CSV, com os códigos como texto:
df = pd.read_csv("loa2025.csv.gz", dtype=str, keep_default_na=False)
```

R:

```
library(arrow)
df <- read_parquet("loa2025.parquet")
```

SQL (DuckDB):

```
SELECT funcao_codigo, funcao_descricao, SUM(valor_dotacao_inicial) AS dotacao_inicial, SUM(valor_pago) AS pago
FROM 'loa2025.parquet'
GROUP BY ALL
ORDER BY dotacao_inicial DESC;
```

Como conferir os arquivos baixados

O arquivo `loaANO.sha256` traz o código SHA-256 de cada arquivo do exercício. Baixe-o junto com os arquivos e, na pasta em que estão, confira o arquivo desejado:

```
grep loa2025.parquet loa2025.sha256 | sha256sum -c        # Linux
grep loa2025.parquet loa2025.sha256 | shasum -a 256 -c    # macOS
```

O resultado deve ser `loa2025.parquet: OK`. Troque o nome para conferir o `.csv.gz` ou o `.zip`. No Windows (PowerShell), calcule o código com `Get-FileHash loa2025.parquet -Algorithm SHA256` e compare com a linha do arquivo no `loa2025.sha256`.

O `.sha256` também lista o `loaANO.nt.gz`, o mesmo RDF do `.zip` compactado em gzip, disponível

no mesmo endereço, para ferramentas que leem o RDF sem descompactar.

O exercício corrente é gerado novamente todos os dias. Use sempre o `.sha256` baixado no mesmo dia dos arquivos.

Correspondência com o RDF

Coluna	Propriedade do item no RDF
`esfera_codigo` … `resultado_primario_codigo`	`loa:codigo` da classificação ligada ao item por `loa:temEsfera`, `loa:temUnidadeOrcamentaria`, `loa:temFuncao`, `loa:temSubfuncao`, `loa:temPrograma`, `loa:temAcao`, `loa:temSubtitulo`, `loa:temPlanoOrcamentario`, `loa:temCategoriaEconomica`, `loa:temGND`, `loa:temModalidadeAplicacao`, `loa:temElementoDespesa`, `loa:temFonteRecursos`, `loa:temIdentificadorUso` e `loa:temResultadoPrimario`; o órgão, pela UO (`loa:temOrgao`)
colunas `_descricao`	`rdfs:label` da classificação
`subtitulo_identificador`, `plano_orcamentario_identificador`	`loa:identificador` do subtítulo e do PO
`valor_projeto_lei`	`loa:valorProjetoLei`
`valor_autografo`	`loa:valorAutografo`
`valor_dotacao_inicial`	`loa:valorDotacaoInicial`
`valor_lei_mais_creditos`	`loa:valorLeiMaisCredito`
`valor_empenhado`	`loa:valorEmpenhado`
`valor_liquidado`	`loa:valorLiquidado`
`valor_pago`	`loa:valorPago`

Mais detalhes sobre o RDF e exemplos de consulta estão em Scripts SPARQL de exemplo.

Atualização

O exercício corrente é atualizado de segunda a sábado, com a última carga dos dados de execução. Os demais exercícios só mudam quando há correção ou mudança de regra, registrada no histórico abaixo.

Histórico de alterações

17/11/2026: todos os exercícios de 2000 a 2025 foram gerados novamente, com as mesmas regras do exercício corrente, e passaram a ser publicados também em CSV e Parquet, com o arquivo SHA-256. O que mudou em relação aos arquivos e ao endpoint anteriores:

- **Liquidado:** a mesma regra em todos os exercícios, sem as despesas liquidadas por inscrição em restos a pagar não processados, onde a base de origem as separa (de 2007 em diante). Mudam os valores de 2012 a 2014 nos arquivos e de 2007 a 2014 no endpoint SPARQL, que antes incluíam essas despesas. Com isso, o empenhado menos o liquidado desses exercícios deixa de ser zero e passa a mostrar o valor inscrito em restos a pagar não processados (veja os cuidados na análise).
- **Autógrafo:** passou a ser publicado também de 2010 a 2015 (antes, só de 2016 em diante).
- **Valores zerados:** quando uma programação não tem um dos valores, ele é publicado como zero. Antes, alguns itens dos exercícios antigos não tinham todas as propriedades de valor, e consultas que exigem as sete propriedades deixavam esses itens de fora.
- **Reserva de contingência de 2000 a 2002:** classificada no grupo de natureza de despesa 9 (Reserva de Contingência), como no Painel do Orçamento. Antes, parte dela aparecia num grupo 0, inexistente.
- **Endereços (URIs) dos planos orçamentários:** o mesmo formato em todos os exercícios, com o caractere @ codificado como `%40`.
- **Endpoint de 2014:** saíram 157.450 itens com todos os valores zerados que não estavam no arquivo.

16/11/2026: início da publicação do exercício de 2026 em CSV e Parquet, com o arquivo SHA-256.

From:
http://orcamento.dados.gov.br/siopdoc/ - **Manuais do SIOP**

Permanent link:
http://orcamento.dados.gov.br/siopdoc/doku.php/acesso_publico:dados_abertos_loa

Last update: **2026/11/17 17:24**