

Esta documentação é um guia para compreender a harmonização da base Microdados de Despesas de Entes Subnacionais (MiDES), descrita em Dahis, Ricca, Scot, Sales e Nascimento (2023). O documento contém informações sobre os principais arquivos e fontes usadas na construção das tabelas de compras públicas e execução orçamentária, bem como as variáveis usadas para identificação única de cada tabela. Além disso, listamos questões em aberto e problemas encontrados ao longo do projeto.

## Licitações

### 1. Licitação

Os dados dessa tabela estão ao nível da licitação, identificados unicamente pela variável *id\_licitacao\_bd*. Constam nessa tabela dados de licitações competitivas e não competitivas<sup>1</sup> para todos os municípios dos seguintes estados: CE, PE, PR, MG, RS e PB.


#### 1.1. Dados utilizados:

UF	Arquivos	Observações	Fonte
CE	licitacoes	Informações gerais sobre licitações competitivas e dispensas/inexigibilidade.	<a href="https://api-dados-abertos.tce.ce.gov.br/docs/">https://api-dados-abertos.tce.ce.gov.br/docs/</a>
PE	licitacoesdetalhes unidadesjurisdicionadas	Contém informação sobre as licitações e participantes. Excluímos as variáveis relacionadas aos participantes e também as linhas duplicadas no conjunto <i>id_municipio</i> , <i>id_unidade_gestora</i> e <i>ano</i> para que a tabela fique no nível da licitação. Já o arquivo “unidadesjurisdicionadas” fornece a correspondência entre unidades gestoras e códigos TCE dos municípios.	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos/Exemplo!listar">https://sistemas.tce.pe.gov.br/DadosAbertos/Exemplo!listar</a>
MG	licitacao.licitacao licitacao.dispensa licitacao.recLicitacao licitacao.recDispensa licitacao.homologLicitacao licitacao.fornDispensa	No primeiro, encontram-se as licitações competitivas, já no segundo obtemos as informações sobre dispensa e inexigibilidade. No terceiro e no quarto obtemos informações sobre o valor estimado ( <i>valor_orcamento</i> ) da licitação. Por fim, nos dois últimos arquivos conseguimos extrair preço e valor unitário dos itens licitados. Ao agregar	<a href="https://dadosabertos.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.xhtml">https://dadosabertos.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.xhtml</a>

<sup>1</sup> Dispensa e Inexigibilidade

UF	Arquivos	Observações	Fonte
		o valor total por <i>id_licitacao_bd</i> , obtemos o valor total homologado por licitação ( <i>valor</i> ).	
PR	Licitacao LicitacaoVencedor	“Licitacao” possui informações gerais sobre licitações competitivas e dispensas/inexigibilidade. Já “LicitacaoVencedor” possui informação sobre preço e valor unitário dos itens licitados. Ao agregar o valor total por <i>id_licitacao_bd</i> , obtemos o valor total homologado por licitação ( <i>valor</i> ).	<a href="https://servicos.tce.pr.gov.br/TC-EPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado">https://servicos.tce.pr.gov.br/TC-EPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado</a>
RS	licitacao orgaos_auditados	Licitações por órgão. Necessária correspondência com o arquivo “orgaos_auditados” para obter a relação entre os órgãos e municípios.	<a href="http://dados.tce.rs.gov.br/group/licitacoes">http://dados.tce.rs.gov.br/group/licitacoes</a>
PB	licitacoes_propostas	Contém informações sobre as licitações, participantes e propostas. Excluímos as variáveis relacionadas aos participantes e também as linhas duplicadas no conjunto <i>id_municipio</i> , <i>id_unidade_gestora</i> , <i>ano</i> , para deixar a tabela no nível da licitação.	<a href="https://dados.tce.pb.gov.br">https://dados.tce.pb.gov.br</a>

## 1.2. Tabela de Arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em:  [arquitetura] `procurement_brazil.licitacao`

A variável ‘*id\_licitacao\_bd*’ identifica unicamente as observações e foi construída da seguinte forma:

UF	<i>id_licitacao</i>	<i>id_licitacao_bd</i>
CE	<i>numero_licitacao</i>	<i>id_licitacao</i> , <i>id_municipio</i> , <i>ano</i> , <i>sigla_uf</i>

MG	seq_licitacao	id_licitacao (ou id_dispena <sup>2</sup> ), id_unidade_gestora, sigla_uf
PB	protocolo_licitacao	id_licitacao, id_unidade_gestora, sigla_uf
PE	codigopl	id_licitacao, id_unidade_gestora, sigla_uf
PR	idLicitacao	id_licitacao, id_unidade_gestora, sigla_uf
RS	nr_licitacao	id_licitacao, ano, modalidade, orgao, sigla_uf

### 1.3. Questões em Aberto

- MG: Inconsistências entre 'valor\_total' e 'valor\_orcamento'. Explicação e exemplos em: [valor\\_licitacao](#)
- MG: Belo Horizonte parece ter dados incompletos em 2020 e 2021. Número reduzido de observações e valor total das licitações muito baixo.
- PE: número reduzido de observações em 2012
- PR: Outliers em valor\_orcamento. Valores acima do percentil 99.99 substituídos por missing (736 observações). Além disso, a variável 'valor' foi obtida a partir das variáveis 'quantidade' e 'valor vencedor', presentes nos arquivos de itens. Assim como acontece em MG, a variável 'valor\_vencedor' apresenta inconsistência, sendo imputada tanto como valor unitário do item quanto como valor total - já multiplicado pela quantidade adquirida. Para obter uma proxy do valor da licitação, consideramos o 'valor\_vencedor' como valor unitário de cada item, multiplicamos pela quantidade e agregamos ao nível da licitação.
- RS: presença de órgãos municipais e estaduais. Alguns órgãos da tabela de licitação não possuem municípios correspondentes ao fazer o merge com o arquivo *orgaos\_auditados*. Em função disso, 9671 observações foram excluídas.

Exemplos de órgãos sem correspondência:

CENTRAL DE LICITAÇÕES DO ESTADO

SES - SECRETARIA DA SAÚDE DO RS

SECRETARIA DO MEIO AMBIENTE E INFRAESTRUTURA

SEDUC - SECRETARIA DA EDUCAÇÃO DO RS

---

<sup>2</sup> No caso de MG, temos duas variáveis chaves, 'id\_licitacao' e 'id\_dispena'. Utilizamos a variável 'id\_licitacao' para criar o 'id\_licitacao\_bd' no caso das licitações competitivas e 'id\_dispena' para criar o 'id\_licitacao\_bd' nos casos de dispensa e inexigibilidade.

SEFAZ - SECRETARIA DA FAZENDA DO RS  
SSP - SECRETARIA DA SEGURANÇA PÚBLICA

- CE: casos em que o 'id\_licitacao' difere apenas pelo prefixo. Por exemplo:

FMS04.03.002/19

PMB04.03.002/19

FMA04.03.002/19

FME04.03.002/19

Mesmo id do município, descrição do objeto licitado e valor orçado. Aparentemente, parece uma única licitação direcionada à unidade gestoras diferentes ([Fonte](#)).

- CE: id\_licitacao\_bd não é único mesmo ao adicionar id\_municipio e ano. Ao fazer essa construção, ainda obtemos 0.87% de observações com id duplicados. Parte dessas observações são duplicadas em todas as variáveis relevantes (tais como município, modalidade, preço, data do edital), e diferem apenas em variáveis que não estamos usando na construção da tabela de arquitetura de licitação (ex: data de autuação, comissão responsável). Nesses casos, temos razões para acreditar que se trata da mesma licitação e excluimos as duplicadas.

#### 1.4. FAQs

- O que diferencia a variável valor\_orcamento da variável valor?  
O orçamento é uma estimativa feita ao abrir a licitação, enquanto a variável valor é resultado do valor homologado aos fornecedores. Por vezes, quando não há informação sobre este último, temos que agregar os valores unitários da tabela de itens para obter o valor total de cada licitação.
- O que é a variável valor corrigido?  
É o mínimo entre valor e valor\_orcamento.
- O que significa id\_licitacao\_bd missing?  
Nos casos em que id\_licitacao\_bd é missing, não conseguimos identificar unicamente a licitação.  
Ex: Após a construção do id\_licitacao\_bd no CE, restam 0.87% das observações com id duplicado. Isso significa que, para o mesmo id, há duas ou mais linhas que diferem em alguma variável relevante (município, modalidade, preço, data do edital). Esses ids foram substituídos por missing, já que a duplicidade de informações para um único id não nos permite identificar a licitação unicamente.


## 2. Licitação Item

Dados estão ao nível do item, identificados unicamente pela variável 'id\_item'. Constatamos nessa base, dados dos itens licitados para todos os municípios dos seguintes estados: CE, PR, MG e RS.

### 2.1. Dados utilizados:

UF	Arquivos	Observações	Fonte
CE	itens_licitacoes licitantes	Realizamos o merge entre os dois arquivos. O primeiro contém informação sobre os itens licitados e o segundo sobre os vencedores.	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos/ExemploIstar">https://sistemas.tce.pe.gov.br/DadosAbertos/ExemploIstar</a>
MG	licitacao.itemLicitacao licitacao.cotacaoLicitacao licitacao.refLicitacao licitacao.homologLicitacao licitacao.itemDispensa licitacao.cotDispensa licitacao.fornDispensa	Merge dos itens licitados e dispensados, com suas respectivas cotações ou preços de referência. Obtemos também os vencedores, a quantidade licitada e o valor unitário (homologLicitacao e fornDispensa).	<a href="https://dadosabertos.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.shtml">https://dadosabertos.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.shtml</a>
PR	LicitacaoVencedor	O arquivo já reporta os vencedores ao nível do item.	<a href="https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacao/Dados/DadosConsulta/Consolidado">https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacao/Dados/DadosConsulta/Consolidado</a>
RS	item licitacoes pessoas	Realizamos o merge entre os três arquivos. O primeiro contém informação sobre os itens licitados e cpf/cnpj dos vencedores, exceto para as modalidades dispensa, inexigibilidade, chamamento público, chamada pública e registro de preços. O segundo, por sua vez, possui os dados de cpf/cnps não contidos em <i>item</i> . O terceiro contém a razão social de todos os participantes nas licitações.	<a href="http://dados.tce.rs.gov.br/group/licitacoes">http://dados.tce.rs.gov.br/group/licitacoes</a>

## 2.2. Tabela de Arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em:  [arquitetura] procurement\_brazil.licitacao\_item

A variável 'id\_item\_bd' identifica unicamente as observações. Abaixo estão as variáveis chaves:

UF	id_item	id_item_bd
MG	já existente nos arquivos originais	id_item, id_unidade_gestora, sigla_uf <sup>3</sup>
PR	numero, numero_lote, id_licitacao, id_municipio	id_item, sigla_uf
CE	numero, id_municipio, id_licitacao, ano	id_item, sigla_uf
RS	numero, numero_lote, id_licitacao, ano, modalidade, orgao	id_item, sigla_uf

## 2.3 Questões em aberto

- PR: problemas ao converter xml para csv:  
2013\_411915  
2015\_411370  
2016\_411535  
2017\_411710  
2018\_412627  
2021\_410140  
Csv vazio: 2021\_410350
- Variável vlLicitacaoVencedorLicitacao em alguns casos representa o valor unitário do item e em outros casos representa o valor total do item (p x q). Por isso, optei por renomeá-la como valor\_vencedor ao invés de valor\_unitario ou valor\_total.
- Inconsistências na variável valor\_unitário em MG. Alguns valores parecem ser os valores totais. Explicação e exemplos em: [valor\\_licitacao](#)

---

<sup>3</sup> O id de item original de MG não identifica unicamente (aproximadamente 5% dos ids são duplicados). Isso ocorre, principalmente, porque alguns itens dos arquivos de licitação possuem o mesmo número de itens dos arquivos de dispensa. Criando um id item que seja id\_item + id\_unidade gestora resolvemos o problema (cai para aproximadamente 0,5%). As observações duplicadas restantes mostram o mesmo item (mesmo id e mesma descrição) com mais de um vencedor.

## 2.4 FAQs

- Por que PE e PB não aparecem nesta tabela?  
Não constam nessa tabela porque não disponibilizam os arquivos com itens. Apenas disponibilizam os dados agregados por licitação.
- Por que em alguns casos id\_item aparece como missing ou duplicado?  
Substituímos id\_item por missing nos casos em que um mesmo id aparece com diferentes descrições ou mesma descrição mas diferentes quantidades ou valores. Só permitimos que os ids sejam duplicados quando o mesmo item é fornecido por diferentes fornecedores.

Ex: No caso do CE, 0.70% das observações são duplicadas, mas grande parte ocorre por conta de id\_licitacao\_bd missing, já que essa variável é usada para construir o id do item. Considerando os casos em que id\_licitacao\_bd não é missing, o percentual de duplicadas cai para 0.11%. Substituímos id\_item por missing os casos com mesmo id, mas diferentes descrições ou mesma descrição mas diferentes quantidades ou valores. Só permitimos que os ids sejam duplicados quando o mesmo item é fornecido por diferentes fornecedores.

## 3. Licitação Participante

Dados estão ao nível do participante por licitação. Constam nessa base, dados dos participantes (habilitados, vencedores e não vencedores) para todos os municípios dos seguintes estados: CE, PE, PR, MG, RS, PB.

### 3.1 Dados utilizados:

UF	Arquivos	Observações	Fonte
CE	licitantes	Excluimos observações duplicadas no grupo <i>id_licitacao,id_municipio,ano</i> . Ou seja, participantes que possam ter vencido mais de um item na mesma licitação.	<a href="https://api.tce.ce.gov.br/">https://api.tce.ce.gov.br/</a>
PE	licitacoesdetalhes unidadesjurisdicionadas	O arquivo “licitacoesdetalhes” já fornece a informação dos participantes no nível da licitação. Já o arquivo “unidadesjurisdicionadas” fornece a correspondência entre unidades gestoras e códigos TCE dos municípios.	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos/ExemploIllustrar">https://sistemas.tce.pe.gov.br/DadosAbertos/ExemploIllustrar</a>

MG	licitacao.habLicitacao licitacao_item	O primeiro arquivo lista os participantes habilitados por licitação. Conseguimos saber quais desses são vencedores dando merge com a tabela licitacao item.	<a href="https://dados.bertolotto.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.xhtml">https://dados.bertolotto.tce.mg.gov.br/view/xhtml/paginas/downloadArquivos.xhtml</a>
PR	LicitacaoParticipante LicitacaoVencedor	O primeiro arquivo lista os participantes classificados, desclassificados, habilitados e não habilitados por licitação. Merge com a tabela licitacao item (construída a partir de LicitacaoVencedor) para obter vencedores.	<a href="https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado">https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado</a>
RS	licitante licitacao item pessoas	Merge entre os quatro arquivos. O primeiro arquivo lista os participantes (exceto para modalidades dispensa, inexigibilidade e registro de preços). Já pelos arquivos de licitacao e itens conseguimos identificar os vencedores de todas as modalidades. Por fim, o arquivo de pessoas é utilizado para obter a razão social.	<a href="http://dados.tce.rs.gov.br/group/licitacoes">http://dados.tce.rs.gov.br/group/licitacoes</a>
PB	Licitacoes_Propostas	Contém informação sobre as licitações, participantes e propostas. Aqui utilizamos apenas as variáveis relacionadas aos participantes.	<a href="https://dados.tce.pb.gov.br">https://dados.tce.pb.gov.br</a>

### 3.2 Tabela de Arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em: [\[arquitetura\] procurement\\_brazil.licitacao\\_participante](#)

### 3.3 Questões em aberto

- Identificação do tipo de participante (pessoa física ou jurídica) pelo número de dígitos do documento. Variável construída para PE e MG. Ocorrência de alguns casos em que o documento não tem 14 (CNPJ) ou 11 dígitos (CPF).
- MG: 1.1 milhão de vencedores e só 126 mil de não vencedores. Esses números significam que tem muita licitação com concorrência zero? Municípios reportam bem vencedores, mas não estão reportando bem habilitados? Grande parte da discrepância é dada pelas modalidades dispensa/inexigibilidade (~215 mil). Nas



modalidades convite, concorrência, tomada de preço e concurso a quantidade de vencedores e não vencedores está balanceada. Em pregão presencial, pregão eletrônico e leilão há uma grande diferença. Para visualizar essa estatística, utilize a query:

```
SELECT COUNT(a.documento), a.vencedor, b.modalidade
FROM `basedosdados.world_wb_mides.licitacao_participante` a
LEFT JOIN `basedosdados.world_wb_mides.licitacao` b
ON a.id_licitacao_bd = b.id_licitacao_bd
WHERE a.sigla_uf = 'MG'
GROUP BY a.vencedor, b.modalidade
ORDER BY modalidade
```

- MG: 1.2M obs, mas usando documento e razão social só tem 330.747 participantes únicos. Alguns participantes aparecem mais de 3,000 vezes na base. Exemplo: BIOHOSP PRODUTOS HOSPITALARES LTDA, ALFALAGOS LTDA.

### 3.4 FAQs

- Por que, em MG, algumas licitações não apresentam informações sobre habilitados? As licitações que não apresentam informação sobre habilitados são aquelas que ocorreram via dispensa ou inexigibilidade. Nesse caso, sabemos apenas quem são os vencedores.

## Despesas

Entende-se por despesas da Execução Orçamentária, os dados de empenhos, liquidações e pagamentos.

### 1. Empenho

Dados estão ao nível do empenho, identificados unicamente pela variável 'id\_empenho\_bd'. Constam nessa base, dados dos empenhos segundo órgãos/unidades gestoras, função, subfunção, programa/ação e a descrição da despesa. Os dados contemplam os municípios dos seguintes estados: CE, MG, PB, PE, PR, RS e SP (exclusive a capital).

#### 1.1. Dados utilizados

UF	Arquivos	Observações	Fonte
CE	Notas_Empenho Anulações_Empenho	Contém informação sobre o município, órgão e unidade gestora. Traz o número da licitação, embora o número de observações seja relativamente pequeno dentro o total. Identifica o empenho por um número não único e informa o valor empenhado	<a href="https://api.tce.ce.gov.br/">https://api.tce.ce.gov.br/</a>
PE	DespesasMunicipais EmpenhoResumo	Nos dados de empenho com informações sobre a unidade gestora e orçamentária do empenho, função, subfunção, programa e elemento da despesa, os empenhos aparecem uma única vez com o valor final identificados por um ID e número do empenho. Já nos arquivos de EmpenhoResumo via ID é possível reconhecer (para alguns casos) as etapas do empenho, todavia apresenta-se com IDs diferentes impossibilitando a conexão entre as bases	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos">https://sistemas.tce.pe.gov.br/DadosAbertos</a>
MG	Despesa	Por um ID único, apresenta a origem (município, órgão e unidade gestora), além das características (função, subfunção, ação, programa, etc) e os valores. Em relação a este último, há a demonstração do valor empenhado, anulado e/ou reforçado	<a href="https://dadosabertos.tce.mg.gov.br/">https://dadosabertos.tce.mg.gov.br/</a>
PR	Empenho	Identificados por um id e um número, os empenhos tem informação de origem e as demais características. A tabela ainda mostra de forma consolidada o valor do empenho, o	<a href="https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Rela">https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Rela</a>

		valor de anulação, bem como os valores liquidados, pagos e os de estorno. Os dados de id_municipio nos dados originais constam apenas 6 dígitos, sendo transformado para 7 no tratamento.	con/Dados/DadosConsulta/Consolidado
RS	Despesa	Os dados originais mostram em um único arquivo toda a movimentação da execução, isto é, empenho (E), liquidação (L) e pagamento (P) em um só <i>dataframe</i> através da variável <i>tipo_operacao</i> . Identificados sob um número, as informações que caracterizam o empenho se repetem, contudo, a coluna “valor_empenho” mostra o valor empenhado, e caso haja anulação, o valor é negativo.	<a href="http://dados.tce.rs.gov.br/">http://dados.tce.rs.gov.br/</a>
PB	TCE-PB-SAGRES-Empenhos_Esfera_Municipal.txt	Os dados foram particionados por ano. Informa sobre a unidade gestora e orçamentária, mas sem identificação exclusiva do município. Além disso, cada empenho, de forma única, é conectado aos dados de estorno demonstrado através da variável <i>valor_anulacao</i> . É posto aqui e não em outras bases, porque não se identifica o momento do estorno (pós pagamento, por exemplo) e a base de liquidação e pagamento, comumente possui mais de 1x o id_empenho_bd ocorrendo em um merge m:1. Apresenta muita das informações via texto, é o caso da natureza_despesa, programa e ação, por exemplo.	<a href="https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb">https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb</a>
SP	Despesa	Os empenhos são identificados pela variável <i>tp_despesa</i> = ‘Empenhado’ e o número_empenho de um dado órgão e município. Os demais dados tratam das características do empenho, sem informação da modalidade, isto é, se o empenho é ordinário, global ou estimativo.	<a href="https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados">https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados</a>

## 1.2. Tabela de arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em: [\[arquitetura\] procurement brazil.empenho](#)

A variável 'id\_empenho\_bd' identifica unicamente as observações. Abaixo estão as variáveis chaves:

UF	id_empenho_bd
CE	numero_empenho, orgao, id_unidade_gestora, id_municipio, mes_empenho, 2 últimos dígitos do ano [20AA]
MG	id_empenho, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
PB	numero_empenho, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
PE	Não é possível criar um id único
PR	id_empenho, id_municipio
RS	numero_empenho, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
SP	numero_empenho [exclusive os 5 últimos dígitos], orgao, id_municipio, 2 últimos dígitos do ano [20AA]

Algumas bases como as de MG e PR já fornecem um id único para o empenho que o acompanha em toda a execução orçamentária. A construção de um id imputando mais elementos como código IBGE do município ou os últimos algarismos do ano consiste para garantir, que entre anos, a unicidade se mantenha.

### 1.3. Questões em Aberto

- No caso de PE, não foi possível criar a variável id\_empenho\_bd que identifica cada linha unicamente. O id do empenho presente no arquivo de empenho não corresponde ao id do empenho presente nos arquivos de liquidação, pagamento e resumo.
- Aprimorar as informações de licitacao (id\_licitacao\_bd, id\_licitacao, modalidade\_licitacao). Investigar como os empenhos que tenham informação de licitação podem se conectar com as bases de licitação.

### 1.4. FAQs

- Por que em alguns casos o identificador id\_empenho\_bd aparece como missing? Isso ocorre quando não é possível identificar unicamente a observação com as variáveis utilizadas para gerar o id\_empenho\_bd. Em alguns casos, o mesmo número do empenho, com a mesma data, gerado pelo mesmo órgão e/ou unidade gestora, apresenta informações destoantes como função, subfunção, programa ou ação, o que gera incerteza na determinação de um empenho com apenas uma finalidade (como é recorrente em outras observações e estados). No caso de imprecisão em determinar a unicidade da observação com a combinação de variáveis na criação do id\_empenho\_bd, definimos como missing.

## 2. Liquidação

Dados estão ao nível do empenho, identificados unicamente pela variável 'id\_liquidacao\_bd'. Constam nessa base, dados dos empenhos liquidados segundo órgãos/unidades gestoras, função, subfunção, programa/ação e a descrição da despesa. Os dados contemplam os municípios dos seguintes estados: CE, PE, PR, MG, SP (exclusive a capital), RS e PB.

### 2.1. Dados utilizados

UF	Arquivos	Observações	Fonte
CE	Liquidacoes Estornos_liquidacao	Identifica via numero_empenho a quantidade de vezes em que o mesmo foi liquidado. Traz informação ainda do responsável pela liquidacao.	<a href="https://api.tce.ce.gov.br/">https://api.tce.ce.gov.br/</a>
PE	EmpenhoLiquidacao	Os dados mostram informações sobre liquidação, estornos, cancelamentos e anulações. A limitação é que o id_empenho aqui adotado identifica a mesma despesa de maneiras diferentes não havendo possibilidade de conexão entre as operações. Não há informações sobre o responsável pela liquidação, apenas o número de vezes em que o empenho foi liquidado.	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos">https://sistemas.tce.pe.gov.br/DadosAbertos</a>
MG	Despesa.liquidacao Rsp	Por um ID único (id_liquidacao), apresenta informações de conexão com o id_empenho, o responsável pela liquidação e os valores liquidado, anulado e final. Complementa-se ainda com as informações de restos a pagar. Nos dados originais não consta id_município, adicionado durante a leitura dos dados para posterior subida pro BQ	<a href="https://dadosabertos.tce.mg.gov.br/">https://dadosabertos.tce.mg.gov.br/</a>
PR	EmpenhoLiquidacao EmpenhoLiquidacaoEstorno	Identificados por um id (id_liquidacao) e um número (nrLiquidacao), a base mostra as informações de identificação do empenho, do responsável e os valores inicial, estornado e final. O identificador não conecta perfeitamente todas as movimentações da transação.	<a href="https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado">https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacon/Dados/DadosConsulta/Consolidado</a>
RS	Despesas	Filtrados pela variável <i>tipo_operacao</i> = L a base replica as informações em uma outra linha. A diferença entre empenho e liquidação é que a análise	<a href="http://dados.tce.rs.gov.br/">http://dados.tce.rs.gov.br/</a>

		aqui traz informação no campo <i>numero_liquidacao</i> e <i>valor_liquidacao</i>	
PB	TCE-PB-SAGRES-Liquidacoes_Esfera_Municipal.txt TCE-PB-SAGRES-Estornos_Esfera_Municipal.txt	O arquivo obtido até aqui traz no formato .txt os registros de 2008 a 2021. Os dados foram particionados por ano. A base precisa de join com o arquivo auxiliar de id_unidade_gestora que associa cada uma das unidades aos seus respectivos municípios.	<a href="https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb">https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb</a>
SP	Despesa	Os dados de empenho se repetem na tabela de liquidação, distintos apenas pela variável <i>tp_despesa</i> = 'Valor Liquidado'. O <i>id_empenho_bd</i> é replicável pela base se manter igual, mas a criação de uma chave primária para liquidação é impossível porque não haveria equivalente na base de pagamentos	<a href="https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados">https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados</a>

## 2.2. Tabela de arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em: [\[arquitetura\] procurement\\_brazil.liquidacao](#)

A variável 'id\_liquidacao\_bd' identifica unicamente as observações. Abaixo estão as variáveis chaves:

UF	id_liquidacao_bd
CE	Não é possível criar um id único
MG	id_liquidacao, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
PB	Não é possível criar um id único
PE	Não é possível criar um id único
PR	id_liquidacao, id_municipio
RS	numero_liquidacao, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
SP	numero_empenho [exclusive 5 últimos dígitos], orgao, id_municipio, identificador_credor, 2 últimos dígitos do ano [20AA]

## 2.3. Questões em Aberto

- Os dados de restos a pagar em PE, RS, PB, PR, CE e SP.
- Inclusão do CNPJ para calcular delay-payment no nível da firma.

## 2.4. FAQs

- Por que alguns estados não possuem id\_liquidacao\_bd?  
O Estado não apresenta informação sobre liquidação ou apresenta informação sobre liquidação mas não seria possível construir essa variável na base de pagamentos.
- Por que em alguns casos o identificador id\_liquidacao\_bd aparece como missing?  
Isso ocorre quando não é possível identificar unicamente a observação com as variáveis utilizadas para gerar o id\_liquidacao\_bd. Alguns estados (SP e RS, até então) que apresentam os dados no formato long podem conter informações diferentes para um mesmo empenho, ainda que não expostas aqui pela arquitetura. Dada a impossibilidade de identificação única, substituímos os identificadores criados por missing.

## 3. Pagamento

Dados estão ao nível do empenho, identificados unicamente pela variável 'id\_liquidacao\_bd'. Constam nessa base, dados dos empenhos liquidados segundo órgãos/unidades gestoras, função, subfunção, programa/ação e a descrição da despesa. Os dados contemplam os municípios dos seguintes estados: CE, PE, PR, MG, SP (exclusive a capital), RS e PB.

### 3.1. Dados utilizados

UF	Arquivos	Observações	Fonte
CE	Notas_Pagamento Estorno_Pagamentos	Contém informação sobre o município, órgão e unidade gestora. Traz o número da licitação, embora o número de observações seja relativamente pequeno dentro o total de observações. Identifica o empenho por um número não único e informa o valor empenhado	<a href="https://api.tce.ce.gov.br/">https://api.tce.ce.gov.br/</a>
PE	EmpenhoPagamento	Nos dados de empenho com informações sobre a unidade gestora e orçamentária do empenho, função, subfunção, programa e elemento da despesa, os empenhos aparecem uma única vez com o valor final identificados por um ID e número do empenho. Já nos arquivos de	<a href="https://sistemas.tce.pe.gov.br/DadosAbertos">https://sistemas.tce.pe.gov.br/DadosAbertos</a>

		EmpenhoResumo via ID é possível reconhecer as etapas do empenho, todavia apresenta-se com IDs diferentes impossibilitando a conexão entre as bases	
MG	Despesa.pagamento Rsp	Por um ID único (id_pagamento), apresenta informações de conexão com o id_empenho, id_liquidação e informações referente ao tipo de pagamento, fonte do recurso, valores retidos na fonte, anulado e pago. A base possui ainda inconsistências de equivalência em relação aos nomes e as observações prestadas.	<a href="https://dados.abertos.tce.mg.gov.br/">https://dados.abertos.tce.mg.gov.br/</a>
PR	EmpenhoPagamento EmpenhoPagamentoEstorno	Identificados por um id e um número, os empenhos tem informação de origem e as demais características. A tabela ainda mostra de forma consolidada o valor do empenho, o valor de anulação, bem como os valores liquidados, pagos e os de estorno	<a href="https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacao/Dados/DadosConsulta/Consolidado">https://servicos.tce.pr.gov.br/TCEPR/Tribunal/Relacao/Dados/DadosConsulta/Consolidado</a>
RS	Despesas	Filtrados pela variável <i>tipo_operacao = P</i> a base replica as informações em uma outra linha. A diferença entre empenho, liquidação e pagamento é que a análise aqui traz informação no campo <i>numero_pagamento</i> e <i>valor_pagamento</i>	<a href="http://dados.tce.rs.gov.br/">http://dados.tce.rs.gov.br/</a>
PB	TCE-PB-SAGRES-Pagamentos_Esfera_Municipal.txt	O arquivo obtido até aqui traz no formato .txt os registros de 2003 a 2021. Os dados foram particionados por ano. Traz informação sobre o numero da parcela de pagamento, o credor e os valores de retenção e pago.	<a href="https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb">https://zeosever.pb.gov.br/portaltcepb/tcepb/servicos/dados-abertos-do-sagres-tce-pb</a>
SP	Despesa	Os dados originais são identificados via número e uma coluna descreve a etapa da execução orçamentária, isto é, se o valor é Empenhado, Liquidado ou Pago naquela data. Para criar um id_pagamento_bd, como este não precisa se conectar a nenhuma outra base, é possível criá-lo utilizando do id_despesa_detalhe. É um id nos dados originais que identifica unicamente cada observação	<a href="https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados">https://transparencia.tce.sp.gov.br/sites/default/files/conjunto-dados</a>



### 3.2. Tabela de arquitetura

Na tabela de arquitetura é possível ver detalhadamente todas as variáveis criadas e renomeadas a partir das variáveis originais. A tabela de arquitetura pode ser encontrada em: [\[arquitetura\] procuremente\\_brazil.pagamento](#)

A variável 'id\_pagamento\_bd' identifica unicamente as observações. Abaixo estão as variáveis chaves:

UF	id_pagamento_bd
CE	numero_empenho, numero_nota_pagamento, orgao, id_unidade_gestora, id_municipio,, 2 últimos dígitos do ano [20AA]
MG	id_pagamento, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
PB	numero_pagamento, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
PE	Não é possível criar um id único
PR	id_pagamento, id_municipio
RS	numero_pagamento, orgao, id_municipio, 2 últimos dígitos do ano [20AA]
SP	numero_empenho [exclusive 5 últimos dígitos], orgao, id_municipio, identificador_credor, 2 últimos dígitos do ano [20AA]

### 3.3. Questões em Aberto

- Os dados de restos a pagar em PE, RS, PB, PR, CE e SP.
- Analisar observações em que o mesmo id\_empenho\_bd aparece com dois fornecedores diferentes.
- Informações sobre a fonte das receitas para os empenhos. Apesar de não ser comum a todos os estados, é preciso padronizar dado a diversidade oriunda dos dados originais

### 3.4. FAQs

- Por que em alguns casos o identificador id\_pagamento\_bd aparece como missing? Isso ocorre quando não é possível identificar unicamente a observação com as variáveis utilizadas para gerar o id\_pagamento\_bd. Alguns estados (SP e RS, até então) que apresentam os dados no formato long podem conter informações diferentes para um mesmo empenho, ainda que não expostas pela arquitetura final. Dada a impossibilidade de identificação única, substituímos os identificadores criados por missing.