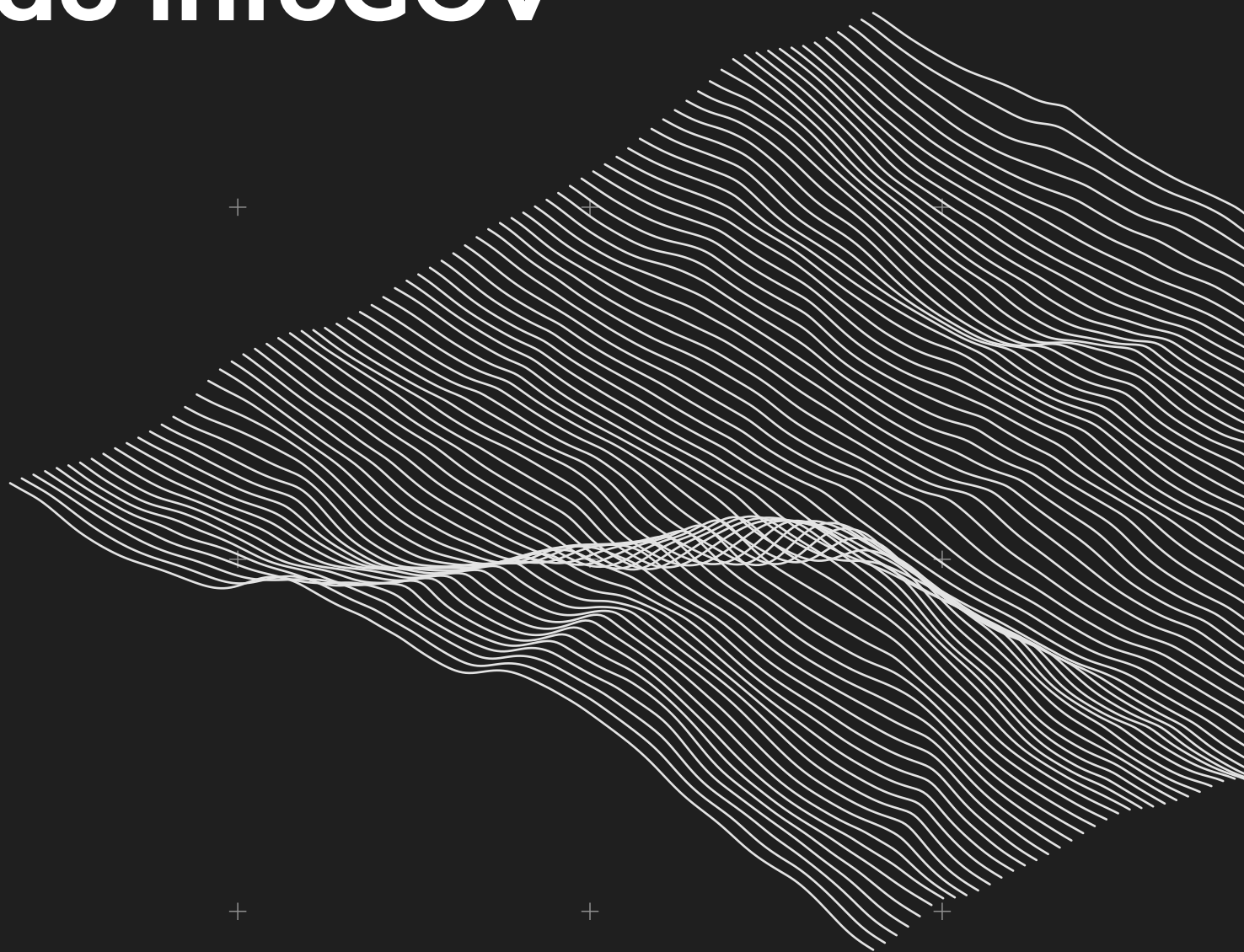


Dicionário de
Dados das Bases
do infoGOV



Os dados do infoGOV

Os dados que fazem parte do InfoGOV são relativos aos sistemas estruturantes da Administração Pública Federal. Isso é, são sistemas que fornecem dados administrativos que precisam ser tratados para se transformarem em dados analíticos. O intuito do InfoGOV é ler o Estado brasileiro de forma analítica, sendo que as bases foram construídas para atingir esse objetivo. Os dados fornecidos aqui não visam subsidiar processos de auditoria, ou de qualquer outra forma de controle pormenorizado do funcionamento da máquina pública.

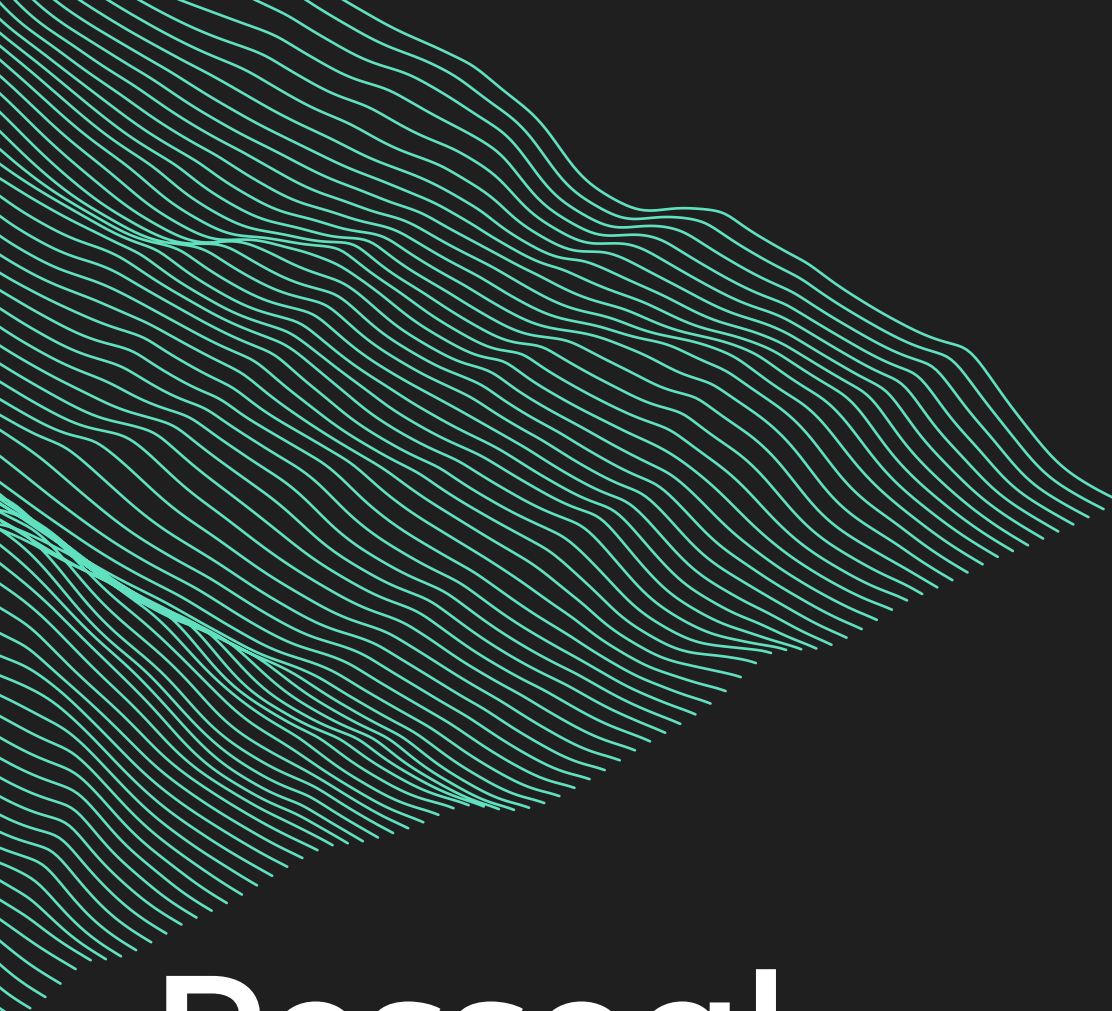
O processo de transformação dos dados é explicado brevemente abaixo, para cada um dos sistemas que compõem o InfoGOV: SIAPE, SICONV e SIASG.

Além dos sistemas estruturantes, há ainda a base de dados do RegbBr, uma iniciativa da Enap que visou quantificar e analisar todo o regramento brasileiro: leis, portarias, resoluções, instruções normativas, entre outros. Também consta no banco do InfoGOV as bases de dados construídas para a análise da Lei 12.990/2014 que instituiu as cotas no serviço público federal.

Esperamos que o(a) pesquisador(a) que tenha acesso a esse documento, possa compreender o universo dos dados constantes do InfoGOV e ter insights analíticos capazes de subsidiar estudos inovadores sobre a Administração Pública Federal brasileira.

Façam bom uso!

Equipe CGDADOS



Pessoal

(SIAPÉ)

A base do SIAPE possui uma atualização trimestral das informações. Os dados são baixados via DW (Datawarehouse), com senha e login fornecidos pelo Ministério da Economia (ME).

Para entender melhor o tratamento é importante destacar a seguinte informação: pessoas podem ter mais de um vínculo (contrato) com a administração pública. Às vezes essas pessoas podem estar cedidas de um órgão para outro (chamamos de relação cedido/requisitado). Às vezes essas pessoas podem ter mais de um vínculo de forma independente (conforme alguns casos previstos na Lei 8.112/90). Dessa forma, optou-se por unificar os registros dos servidores quando ocorre a relação “cedido/requisitado”, ou quando o servidor ocupa um função/cargo comissionado. Esse entendimento se deve ao fato de que a pessoa exerce as atividades laborais apenas em 1 local/órgão, e não em 2, como proveniente da base de dados extraída do DW .

As rotinas de tratamento dos dados seguem os seguintes passos:

- 1.** Pré-processamento consiste em:
 - a.** remoção de acentos e caracteres indesejados
 - b.** remoção de whitespace (espaço, tab, enter, etc) em excesso
 - c.** remoção de alguns valores inúteis como ‘S/ORGÃO SUPERIOR’ que é trocado por vazio e, por fim, 4) upload dos dados baixados do DW em uma tabela temporária no banco de dados.

- 2.** Processamento de dados
 - a.** Cria-se uma lista de CPFs existentes na tabela temporária do final do pré-processamento e separam-se blocos contendo 5 mil CPFs para serem trabalhados por vez. Note que registros com CPF vazio são descartados automaticamente e nós adicionamos uma proibição de CPF com valor zero.
 - b.** Dado um bloco de 5 mil CPFs, se puxam as informações da tabela temporária em que o CPF consta na lista. Isso gera um pouco mais de 5 mil linhas por conta do efeito “cedido/requisitado”.
 - c.** Para cada CPF da lista:
 - i.** Cria-se um conjunto de informações com todos os vínculos com o mesmo CPF.

ii. Se o vínculo constar na “Fundação Roquete Pinto” ou no “Hospital Nossa Senhora da Conceição”, troca-se o órgão superior para ‘CEDIDO’.

iii. Se for Março de 2018 e o vínculo constar em um órgão subordinado ao órgão superior Ministério da Saúde (Fiocruz, Funasa, Anvisa, Agência Nacional de Saúde ou o próprio Ministério da Saúde), forçamos o órgão superior para “MS” para corrigir um defeito na base que atribui essas pessoas erroneamente à números estranhos.

iv. Se o título de eleitor vier com valor zero, trocamos por vazio

v. No caso da existência de informação nas colunas de “destino” (‘ORGAO SUP DESTINO’, ‘ORGAO DESTINO’, ‘FUNCAO DESTINO’ e ‘NIVEL FUNCAO DESTINO’), damos preferência às mesmas sobre as análogas de origem.

vi. Recodificamos vários valores (‘MES’, ‘ORGAO SUPERIOR’, ‘COR ORIGEM ETNICA’, ‘UF RESID’, ‘JORNADA CARGO’, ‘JORNADA TRABALHO’, ‘ESCOLARIDADE’, ‘CARGO’) para diminuir a quantidade de categorias.

vii. Adicionamos informações organizacionais com base em uma tabela externa ao Siape produzida pela própria Enap, feita para melhorar a qualidade da informação organizacional. Afeta as colunas que têm ‘3.0’ no nome, como ‘COD SIORG 3.0’

viii. Criamos algumas variáveis novas que são reclassificações de antigas: ‘ING SPUB 3.0’, ‘SIT VINC 3.0’, ‘FUNCAO 3.0’ e ‘NIVEL FUNCAO 3.0’, ‘idAnon’

ix. Adicionamos informação partidária e de classificação de partido (coligação do governo ou não) com base em uma tabela antigamente disponível no TSE que liga os títulos de eleitores com informação partidária.

x. Unificamos vínculos que atendem ao par cedido/requisitado.

Sobre esse último ponto, pode-se agora explicar o que o Infogov chama de regras de unificação. Como foi colocado anteriormente, contam-se locais de trabalho ocupados no serviço público. Dessa forma, existem pessoas com mais de um contrato/vínculo com a Administração Pública, sem que necessariamente isso caracterize uma relação cedido/requisitado.

A base do Infogov recebe como entrada todos esses contratos/vínculos com a administração. Esses vínculos são então ordenados pelo CPF, de tal forma que todos os vínculos que uma determinada pessoa tem estejam contínuos na base original. Esses vínculos são então analisados em forma de bloco.

Para caracterizar que uma pessoa está cedida de órgão A para órgão B, utiliza-se a seguinte regra:

1. Se algum dos vínculos contém, no campo SITUAÇÃO VÍNCULO, algum dos valores: ('CEDIDO ART93 8112', 'CEDIDO', 'CEDIDO SUS/LEI 8270', 'CLT ANS JUD. CEDIDO', 'CLT ANS -DEC 6657/08', 'CLT ANS DEC JUDICIAL'), o bloco ganha a marcação de 'cedido'

2. Se algum dos vínculos contém, no campo SITUAÇÃO VÍNCULO, algum dos valores: ('REQ. MILITAR', 'REQ.DE OUTROS ORGAOS', 'REQUISITADO', 'REQUISITADO MILITAR', 'ATIVO PERMANENTE'), o bloco ganha a marcação de 'requisitado'

Caso haja um vínculo marcado como 'cedido' e um vínculo marcado como 'requisitado', caracteriza-se então uma relação de cessão de um servidor de um determinado órgão para outro. Dessa forma, os vínculos são unificados, isso é, se tornam um vínculo único constituído de informações vindas dos dois vínculos, segundo as regras de formação descritas a seguir:

1. Primeiramente, o vínculo unificado recebe todas as informações do vínculo requisitado;

2. Posteriormente, as seguintes informações são trazidas do vínculo cedido:

- a. ANO MES OPÇÃO FUNPRESP
- b. TEMPO SERV APOS
- c. INGRESSO
- d. CARGO
- e. COD CARGO
- f. ESC CARGO
- g. NÍVEL CARGO
- h. GRUPO CARGO
- i. DATA INGRESSO CARGO

3. As remunerações, rendimentos e descontos dos dois vínculos são somadas.

Caso haja mais de um vínculo marcado como cedido ou mais de um marcado como requisitado, os últimos vínculos marcados com cada flag são selecionados para a unificação. Os vínculos marcados como ‘cedido’ e ‘requisitado’ são então removidos do bloco.

A questão de dados partidários era feita com base em informações disponibilizadas pelo TSE (<http://www.dados.gov.br/dataset/dataset/filiados-partidos-politicos>). De 2021 em diante a base deixou de ser disponibilizada e com isso as informações partidárias perdem bastante qualidade dessa data em diante.

Index	Nome Variável	Tipo	Comentário
1	MES	Número	Mês/Ano ao qual o registro se refere, na forma aaaamm.
2	ID_ANON	Número	Id de cada pessoa, anonimizado.
3	NOME SERVIDOR	Texto	Nome do servidor (não disponível no ambiente de pesquisa).
4	SEXO	Texto	Sexo do servidor. Assume os valores: <ul style="list-style-type: none"> ■ MAS ■ FEM
5	DATA NASCIMENTO SERVIDOR	Texto	Data de nascimento do servidor, no formato dd/mm/aaaa (não disponível no ambiente de pesquisa).
6	IDADE	Número	Idade do servidor.
7	CPF SERVIDOR	Número	CPF do servidor (não disponível no ambiente de pesquisa).
8	TITULO ELEITOR	Número	Título de eleitor do servidor (não disponível no ambiente de pesquisa).
9	ESCOLARIDADE	Texto	Escolaridade do Servidor. Pode assumir os valores: <ul style="list-style-type: none"> ■ ATE FUNDAMENTAL ■ ENSINO MEDIO ■ SUPERIOR ■ POS GRADUACAO ■ MESTRADO ■ DOUTORADO
10	COR ORIGEM ETNICA	Texto	Raça ou cor autodeclarada do servidor. Pode assumir os valores: <ul style="list-style-type: none"> ■ AMARELA ■ BRANCA ■ INDIGENA ■ NAO INFORMADO ■ NEGRA
11	UF RESID	Texto	Sigla da UF de residência do servidor.
12	PAIS	Texto	País de naturalidade do servidor. Por extenso.
13	COD DEFICIENCIA	Número	Código da deficiência do servidor.
14	NOME DEFICIENCIA	Texto	Descrição da deficiência do servidor.

Index	Nome Variável	Tipo	Comentário
15	UF NATURALIDADE	Texto	Sigla da UF de naturalidade do servidor.
16	ESTADO CIVIL	Texto	Estado civil do servidor, podendo assumir os valores: <ul style="list-style-type: none"> ▪ CASADO ▪ DIVORCIADO ▪ SEP.JUDIC. ▪ SOLTEIRO ▪ VIUVO
17	HABILITACAO PROFISSIONAL	Texto	Habilitação profissional (formação) do servidor, por extenso.
18	QTD VINCULOS	Número	Quantidade de vínculos com a administração pública que o indivíduo tem dentro do SIAPE.
19	VINCULO INFOGOV	Texto	Booleano. Indicador se aquele servidor teve dois vínculos unificados segundo as regras vigentes no InfoGOV.
20	ID SERVIDOR	Texto	Id do servidor junto ao SIAPE.

Index	Nome Variável	Tipo	Comentário
21	VINCULO SERVIDOR	Texto	Id de cada vínculo de cada servidor junto ao SIAPE.
22	ORGAO SUPERIOR	Número	<p>Id de cada ministério existente na história do SIAPE, ministério esse onde o servidor está lotado. Pode assumir os valores:</p> <ul style="list-style-type: none"> ■ 1,Ministério da Fazenda ■ 2,Presidência da República ■ 3,Ministério da Defesa ■ 4,Ministério da Agricultura, Pecuária e Abastecimento ■ 5,Ministério da Educação ■ 6,Ministério do Planejamento ■ 7,Controladoria-Geral da União ■ 8,Governo do Distrito Federal ■ 9,Ministério das Comunicações ■ 10,Ministério da Cidadania ■ 11,Ministério das Cidades ■ 12,Ministério da Ciência, Tecnologia e Inovação ■ 13,Ministério da Ciência, Tecnologia, Inovação e Comunicações ■ 14,Ministério do Desenvolvimento Agrário ■ 15,Ministério dos Direitos Humanos ■ 16,Ministério da Indústria, Comércio Exterior e Serviços ■ 17,Ministério do Desenvolvimento Regional ■ 18,Ministério do Desenvolvimento Social ■ 19,Ministério do Esporte ■ 20,Ministério da Economia ■ 21,Ministério da Integração Nacional ■ 22,Ministério da Cultura ■ 23,Ministério da Infraestrutura ■ 24,Ministério da Justiça ■ 25,Ministério do Meio Ambiente ■ 26,Ministério de Minas e Energia ■ 27,Ministério da Pesca e Aquicultura ■ 28,Ministério da Previdência Social ■ 29,Ministério das Relações Exteriores ■ 30,Ministério da Saúde ■ 31,Ministério dos Transportes ■ 32,Ministério do Trabalho ■ 33,Ministério da Transparência, Fiscalização e Controladoria-Geral da União ■ 34,Ministério do Turismo ■ 35,Ministério da Mulher, da Família e dos Direitos Humanos ■ 36,Banco Central do Brasil
23	ORGAO	Texto	Órgão onde o servidor está lotado.

Index	Nome Variável	Tipo	Comentário
24	COD ORGAO	Número	Código do órgão onde o servidor está lotado, de acordo com o SIORG.
25	ORGAO SUP ORIGEM	Texto	Ministério de onde o servidor é originário, em caso do mesmo ser originário de outra repartição.
26	ORGAO ORIGEM	Texto	Órgão de onde o servidor é originário, em caso do mesmo ser originário de outra repartição.
27	APOSENTADORIA	Texto	Texto: Descrição da razão legal para a aposentadoria do servidor, se aposentado.
28	COD APOSENTADORIA	Número	Código relativo ao campo 27.
29	ANO PREVISAO APOSENTADORIA	Número	Ano previsto para a aposentadoria do servidor segundo o SIAPE.
30	ANO APOSENTADORIA	Número	Ano da aposentadoria efetiva do servidor, se aposentado.
31	MES APOSENTADORIA	Número	Mês da aposentadoria efetiva do servidor, se aposentado.
32	DIA APOSENTADORIA	Número	Dia da aposentadoria efetiva do servidor, se aposentado.
33	OCOR APOSENTADORIA	Data	Data de ocorrência da aposentadoria do servidor.
34	ANO MES OPCAO FUNPRESP	Texto	Ano/Mês da opção do servidor pelo Funpresp, na forma mm/aaaa
35	TEMPO SERV APOS	Texto	Tempo de serviço já cumprido para a aposentadoria.
36	DIA NOMEACAO FUNCAO	Texto	Dia de nomeação para a função, se o servidor está ocupando alguma. Na forma dd/mm/aaaa.
37	INGRESSO	Texto	Forma de ingresso do servidor no serviço público ou na função que o mesmo está ocupando.
38	COD INGRESSO	Número	Código relativo ao campo 37.
39	ANO ING SPUB	Número	Ano de ingresso do servidor no serviço público.
40	COD INGR SPUB	Número	Código relativo à forma de ingresso do servidor no serviço público.
41	INGRESSO SPUB	Texto	Forma de ingresso do servidor no serviço público.
42	ATIVIDADE FUNCAO	Texto	Função que o servidor está ocupando no serviço público.
43	CARGO	Texto	Cargo/Carreira do servidor no serviço público.
44	NIVEL CARGO	Texto	Nível do cargo do servidor dentro da carreira.
45	ESC CARGO	Texto	Indica a escolaridade relativa ao cargo do servidor. Pode assumir os valores: <ul style="list-style-type: none"> ■ NA (nível auxiliar) ■ NI (nível intermediário) ■ NM (nível médio) ■ NS (nível superior)
46	COD CARGO	Número	Código relativo ao cargo que o servidor ocupa.

Index	Nome Variável	Tipo	Comentário
47	GRUPO CARGO	Texto	Código do grupo no qual se encontra o cargo que o servidor ocupa.
48	DIA OCOR INGR ORGAO EV	Data	Data. Valores iniciaram em 2007/12. Se encontra no módulo DATA OCORRÊNCIA (ESTADO DE VINCULO) / DATA OCOR INGR ORGAO EST VINC. Provável que seja a data em que o estado de vínculo do servidor está de fato relacionada ao órgão.
49	DATA INGRESSO CARGO	Data	Data. Data de ingresso do servidor no cargo/ carreira.
50	NACIONALIDADE	Texto	Nacionalidade do servidor. Pode assumir os valores: <ul style="list-style-type: none"> ■ BRASILEIRO NATO ■ BRASILEIRO NATURALIZADO ■ EQUIPARADO ■ ESTRANGEIRO ■ NACIONALIDADE ■ NAO INFORMADO
51	FLAG FUNCAO	Número	Flag que indica o tipo de função que o servidor está ocupando. Pode assumir os valores: <ul style="list-style-type: none"> ■ 0; sem função considerada ■ 1; DAS, NES, FCI, FDI, FCD, FPR, FPE ■ 2; FDI, FCT, FC, FGE, FGR, FIN, FND
52	FLAG COMISSIONADO	Booleano	Booleano. Flag que indica se a pessoa é comissionada "pura", ou seja, originária de fora da máquina e sem vínculo com esta. Seguem valores: <ul style="list-style-type: none"> ■ 0 - Pessoa com vínculo com o serviço público; ■ 1 - Pessoa sem vínculo com o serviço público;
53	COD NACIONALIDADE	Número	Código relativo ao campo 50
54	FUNCAO	Texto	Função que o servidor está ocupando no momento.
55	NIVEL FUNCAO	Texto	Nível da função que o servidor está ocupando no momento.
56	REG JUR	Texto	Regime jurídico ao qual o servidor atualmente está sujeito.
57	NATUREZA JURIDICA	Texto	Natureza jurídica do órgão no qual o servidor está lotado. Pode assumir os valores: <ul style="list-style-type: none"> ■ ADMINISTR PUBLICA ■ AUTARQUIA FEDERAL ■ FUNDACAO FEDERAL ■ PODER EXECUTIVO ESTADUAL ■ PODER EXECUTIVO FEDERAL

Index	Nome Variável	Tipo	Comentário
58	GRUPO SITUAÇÃO VÍNCULO	Texto	Grupo no qual a situação do vínculo do servidor com o serviço público do servidor se encontra. Pode assumir os valores: <ul style="list-style-type: none"> ■ ATIVO; ■ APOSENTADO ■ INSTITUIDOR PENSÃO
59	SITUACAO VINCULO	Texto	Situação do vínculo do servidor com o serviço público.
60	SITUACAO FUNCIONAL	Texto	Código alfanumérico que indica em qual situação funcional, dentro daquele regime jurídico (56), o servidor se encontra.
61	JORNADA CARGO	Número	Quantidade de horas trabalhadas pelo servidor, semanalmente, naquele cargo.
62	JORNADA TRABALHO	Número	Quantidade de horas trabalhadas pelo servidor, semanalmente, naquele vínculo.
63	UF DA ORGANIZACAO	Texto	UF da organização na qual o servidor trabalha.
64	REGIAO	Texto	Região da UF do campo 63. Pode assumir os seguintes valores:
65	UORG	Texto	Sigla da unidade organizacional na qual o servidor está lotado.
66	COD SIORG	Número	Código, no SIORG, do existente no campo 65.
67	COD UORG	Número	Código do existente no campo 65.
68	NOME UORG	Texto	Nome da unidade organizacional na qual o servidor está lotado.
69	GRUPO UORG	Texto	Grupo da unidade organizacional na qual o servidor está lotado.
70	UPAG	Texto	Unidade de pagamento a qual o servidor está vinculado.
71	NOME MUNICIPIO UPAG	Texto	Nome do município relativo àquela UPAG.
72	COD UPAG	Número	Código relativo ao descrito no campo 70.
73	REMUN	Número	Remuneração do servidor naquele mês, com duas casas decimais.
74	RENDIM	Número	Rendimento (Remuneração + adicionais) do servidor naquele mês, com duas casas decimais.
75	DESCONTO	Número	Descontos na folha do servidor naquele mês, com duas casas decimais.
76	REMUN IPCA	Número	Remuneração do servidor naquele mês, atualizada pelo IPCA até a data mais recente na base, com duas casas decimais.
77	COD SIORG 3.0	Número	Número.Código relativo ao campo 65, após trabalho realizado pela ENAP. (não disponível no ambiente de pesquisa)

Index	Nome Variável	Tipo	Comentário
78	AREA GOVERNO 3.0	Texto	Classificação, de acordo com estudo realizado pela ENAP, de em que área cada órgão se encontra dentro da administração pública federal. Pode assumir os valores: <ul style="list-style-type: none"> ■ Economia e Meio Ambiente ■ Estado ■ Fora do Estudo ■ Infraestrutura ■ Social
79	NAT ORGAO 3.0	Texto	Classificação, de acordo com estudo realizado pela ENAP, de qual seria a natureza em que cada órgão se encontra dentro da administração pública federal. Pode assumir os valores: <ul style="list-style-type: none"> ■ Administração Direta ■ Agência Reguladora ■ Escola de Governo/Pesquisa ■ Fora do Estudo ■ Instituto Federal de Ensino ■ Outras Autarquias/Fundações ■ Superintendências ■ Universidades Federais
80	NOME SIORG 3.0	Número	Código relativo ao nome, no SIORG, do conteúdo no campo 66, após trabalho realizado pela ENAP. (não disponível no ambiente de pesquisa)
81	NAT JUR 3.0	Texto	Classificação, de acordo com estudo realizado pela ENAP, de qual seria a natureza jurídica em que cada órgão se encontra dentro da administração pública federal. Pode assumir os valores: <ul style="list-style-type: none"> ■ Administração Direta ■ Autarquia ■ Autarquia Federal ■ Empresa Pública ■ Entidade ■ Fundação Pública ■ Legal Status ■ NA ■ Sociedade de Economia Mista
82	COD UORG 3.0	Número	Código relativo ao campo 65, atualizado ao longo do tempo, após trabalho realizado pela ENAP. (não disponível no ambiente de pesquisa)
83	COD HIER 3.0	Texto	Código de onde determinada UORG se encontra dentro da estrutura da administração pública. (SIORG), em formato de árvore.
84	NIVEL 3.0	Número	Nível relativo ao nível dentro da "árvore" hierárquica estabelecida pelo estudo relativo ao item 83.
85	ING SPUB 3.0	Texto	O campo é o resultado de tratamento e limpeza realizados pela ENAP, no campo 41.
86	SIT VINC 3.0	Texto	O campo é o resultado de tratamento e limpeza realizados pela ENAP, no campo 59.

Index	Nome Variável	Tipo	Comentário
87	FUNCAO 3.0	Texto	O campo é o resultado de tratamento e limpeza realizados pela ENAP, no campo 54.
88	NIVEL FUNCAO 3.0	Texto	O campo é o resultado de tratamento e limpeza realizados pela ENAP, no campo 55.
89	FILIADO	Booleano	Booleano. O campo indica se o servidor está filiado a algum partido político na data. Assume valores: <ul style="list-style-type: none"> ▪ 0; não filiado ▪ 1; filiado
90	POSICAO	Booleano	Booleano. O campo indica se o partido ao qual o servidor está filiado representa oposição ou está junto com o governo. Dados até o final do Governo Dilma vem de https://repositorio.unb.br/handle/10482/23942 . Após, composição da ENAP. O Código é: <ul style="list-style-type: none"> ▪ 0: filiado a partido fora da coalizão no poder; ▪ 1: filiado a partido dentro da coalizão no poder;
91	PARTIDO	Texto	Sigla do partido ao qual o servidor está filiado naquele mês/ano do registro, de acordo com o TSE.



Cotas

(Avaliação da Lei
12.990/2014)

Para a avaliação da Lei 12.990/2014 que instituiu as cotas para negros no serviço público federal, foram construídas duas tabelas a partir de informações extraídas do Diário Oficial da União (DOU) e de editais das bancas avaliadoras.

A tabela “Cotas_infogov”¹ se refere àqueles servidores que se inscreveram como cotistas, foram aprovados no concurso público e que depois foram encontrados no SIAPE. É importante ressaltar que essa tabela traz um batimento entre aprovados em concursos e servidores que estavam no SIAPE após a data de divulgação do edital do concurso. Isso é, não são servidores nomeados em vagas reservadas para negros.

A tabela “Cotas_professor_magsuperior” se refere às nomeações encontradas no DOU para Professores do Magistério Superior. Nesse caso, são servidores que foram nomeados em vagas reservadas para negros.

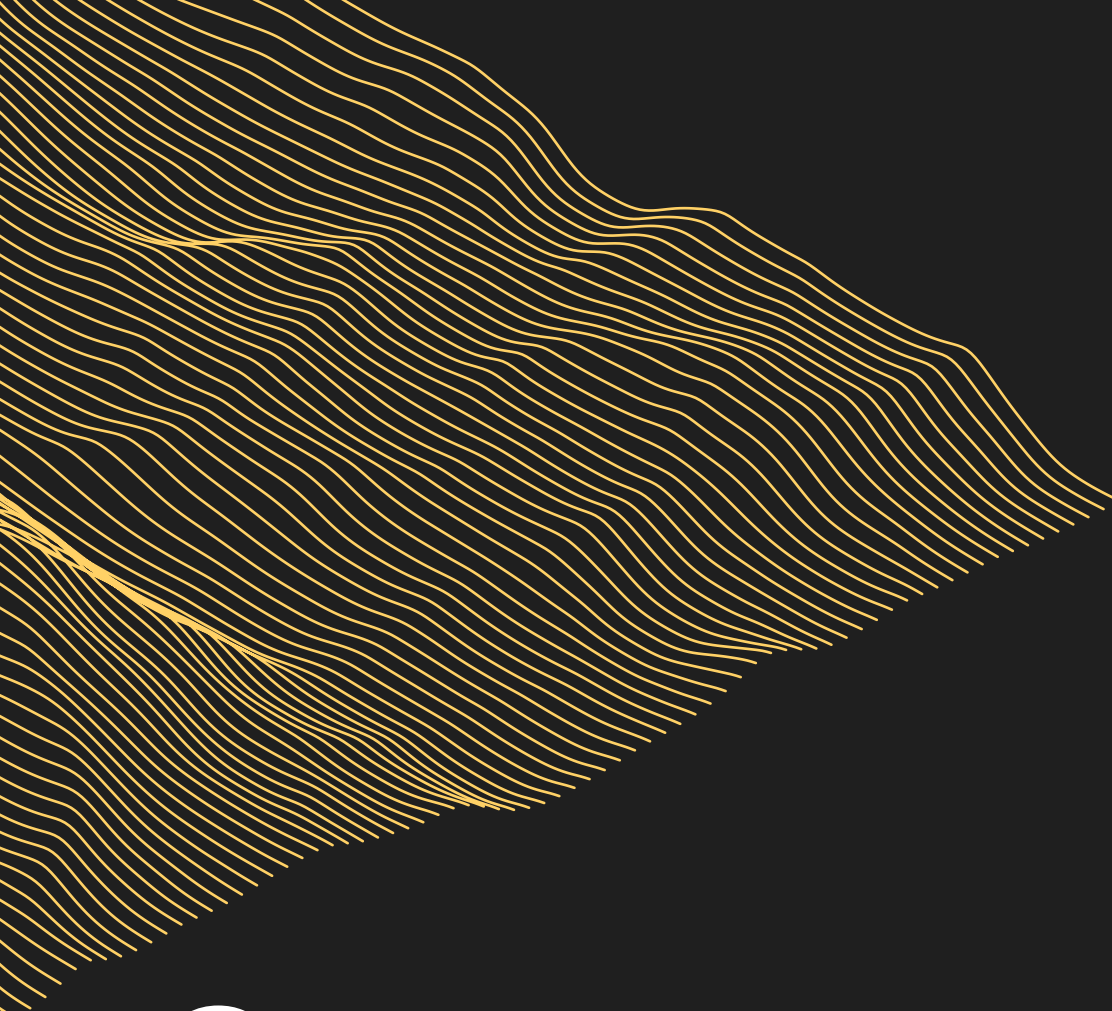
Cotas_infogov

Index	Nome Variável	Tipo	Comentário
1	MES	Número	Mês/Ano ao qual o registro se refere, na forma aaaamm
2	ID_ANON	Número	Id de cada pessoa, anonimizado.
3	MÊS DO EDITAL DE ABERTURA	Texto	Mês que o concurso teve o edital divulgado
4	ANO DO CONCURSO	Número	Ano que foi divulgado o edital do concurso
5	ÓRGÃO SUPERIOR CONCURSO	Texto	Órgão superior que realizou/demandou o concurso
6	ÓRGÃO/ENTIDADE CONCURSO	Texto	Órgão que realizou/demandou o concurso
7	TIPO DE VAGA	Texto	<ul style="list-style-type: none"> ■ Ampla concorrência ■ Cotas
8	TIPO DE VERIFICAÇÃO	Texto	Refere-se ao tipo de verificação que foi feita para atestar que o candidato poderia usufruir das cotas

Cotas_professor_magsuperior

Index	Nome Variável	Tipo	Comentário
1	ID_ANON	Número	Id de cada pessoa, anonimizado.
2	CLASSIFICAÇÃO	Texto	Nomeado em vaga de cotas

¹ Para saber mais sobre a avaliação da Lei 12.990/2014 acessar o relatório produzido pela Enap: [Relatório 4 de 5.pdf \(enap.gov.br\)](#)



Compras (SIASG)

A motivação para o estudo dos dados do SIASG partiu de uma solicitação do Ministério da Economia para que a ENAP contribuísse na construção de um relatório contemplando as compras governamentais feitas a partir do sistema. O SIASG faz parte de um complexo de sistemas governamentais relacionados à compras, com maior foco no executivo federal, com um range de objetos cotados variando desde itens de serviços gerais até o setor de armamentos das forças armadas.

A base do SIASG é acessada de maneira similar à do SIAPE, com um usuário e senha de acesso ao DW de extração de dados. Os dados são extraídos por mês de homologação da compra, para evitar o problema de limite de quantidade de registros por extração.

A unidade mínima possível na base original do SIASG é o “lance de compra”. Ou seja, para cada item a ser comprado (por exemplo, lâmpada) diversos fornecedores podem oferecer lances para ganhar a competição e fazer a entrega para o ente que está fazendo a compra ou contratação de serviço.

Em nosso banco, nossa unidade mínima é o item comprado. Essa escolha facilita o estudo, já que o foco da análise são os gastos despendidos pela administração pública, e não cotações de um determinado item no mercado aberto.

Foram realizados tratamentos no SIASG tanto na categoria de materiais quanto na de serviços, para auxiliar o melhor agrupamento dos itens para análise.

Exemplos em materiais:

Categoria	Categoria Nova
ACESSÓRIOS DE MOTORES	MOTORES
AERONAVES E SEUS COMPONENTES ESTRUTURAIS	AERONAUTICA
ANIMAIS VIVOS	ANIMAIS
APARELHOS E ACESSÓRIOS PARA TREINAMENTO	TREINAMENTOS
ARMAMENTO	ARMAS
ARTIGOS DE HIGIENE	MATERIAIS
BARRAS, CHAPAS E PERFILADOS METÁLICOS	ESTRUTURAS
BOMBAS E COMPRESSORES	MOTORES
CANOS, TUBOS, MANGUEIRAS E ACESSÓRIOS	ESTRUTURAS
COMBUSTÍVEIS, LUBRIFICANTES, ÓLEOS E CERAS	PETROQUIMICOS

E em serviços:

Categoria	Categoria Nova
SERVIÇOS AUXILIARES DA INTERMEDIÇÃO FINANCEIRA EXCETO OS DESEGUROS E DE PENSÕES	FINANCEIROS
SERVIÇOS AUXILIARES DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO (TIC)	TECNOLOGIA
SERVIÇOS AUXILIARES E OUTROS SERVIÇOS DE TRANSPORTE AUXILIAR	LOGISTICA
SERVIÇOS BANCÁRIO DE INVESTIMENTO	FINANCEIROS
SERVIÇOS CIENTÍFICOS E OUTROS SERVIÇOS TÉCNICOS	CONSULTORIA
SERVIÇOS DE AGÊNCIA DE VIAGENS, OPERADORAS DE TURISMO E GUIATURÍSTICO	TURISMO
SERVIÇOS DE AGÊNCIAS DE EMPREGOS E FORNECIMENTO DE PESSOAL	PESSOAL
SERVIÇOS DE AGÊNCIAS DE NOTÍCIAS	IMAGEM
SERVIÇOS DE ALOJAMENTO	ALOJAMENTOS
SERVIÇOS DE ANÁLISE DA DADOS E INDICADORES DE TIC	TECNOLOGIA

Outros tratamentos realizados foram:

- A retirada de elementos de compras realizadas pelos outros Poderes da República ou de entidades paraestatais (Conselhos, MPU, etc);
- O cálculo da quantidade de dias passados entre a publicação do edital no DOU e o mês de homologação da compra;
- A retirada de valores considerados erros/outliers (pregões com mais de R\$ 1bi de valor) e o pregão '16010105000082009', considerado suspeito em termos de volume de vendas e conjuntura específica.

Dessa forma, construiu-se uma base relativamente enxuta, limpa e devidamente preparada para a realização de todos os estudos desejados, desde a sua concepção dentro da ENAP até o presente dia.

Para a utilização no InfoGOV, as análises dos itens comprados ficaram detalhadas. Do ponto de vista da plataforma, não se buscou disponibilizar a informação, por exemplo, de que determinado ente comprou uma quantidade de lâmpadas a um preço individual estipulado. A informação de interesse é saber o quanto se gasta em máquinas, energia, construção, equipamento, etc.

Assim, para o InfoGOV, optou-se por simplificar e resumir a base de dados para tratar de compras de um ponto de vista macro, não se prendendo à composição de cada item individual, mas nas categorias para comparar diferentes entes do governo federal.

Index	Nome Variável	Tipo	Descrição
1	id_compra	Texto	Id do item. Cada registro terá um identificador único deste.
2	id_pregao	Texto	Identificador de pregão, várias compras podem ser feitas no mesmo pregão
3	trimestre	Texto	Trimestre da compra, escrito da forma "1/2007", "2/2007", "3/2007", "4/2007", "1/2008" em diante
4	ano_publicacao_dou	Texto	O ano de publicação no dou
5	mes_publicacao_dou	Texto	O número do mês da publicação no dou
6	data_publicacao_dou	Texto	A data completa da publicação no dou escrita como "aaaammdd", 1 de janeiro de 2007 vira 20070201
7	ano_homologacao_compra	Texto	Ano de homologação da compra
8	mes_homologacao_compra	Texto	O número do mês da homologação da compra (1-12)
9	data_homologacao_compra	Texto	A data completa da homologação escrita como "aaaammdd", 1 de fevereiro de 2007 vira 20070201
10	uf_compra	Texto	Estado do ente que realiza a compra
11	regiao_compradora	Texto	Região do ente que realiza a compra
12	uf_fornecedor	Texto	Estado do ente que entrega a compra
13	regiao_fornecedora	Texto	Região do ente que entrega a compra
14	intervalo	Texto	Quantidade de dias entre a publicação no dou e a homologação da compra
15	org_sup	Texto	Ministério que administra o comprador
16	tipo_adm_uresp_compra	Texto	Tipo de administração do ente que realiza a compra
17	tipo	Texto	Material/Serviço
18	grupo	Texto	Agrupamento feito pela Enap de vários tipos de itens
19	forma_compra	Texto	Sistema utilizado para realizar a compra. Normalmente, SISPP ou SISRP

Index	Nome Variável	Tipo	Descrição
20	forma_realizacao	Texto	Forma de realização da compra. Quase todos são eletrônicos
21	nat_jur_fornecedor	Texto	Natureza jurídica do ente que entrega a compra
22	porte	Texto	Porte do ente que entrega a compra (micro empresa, pequena empresa, etc)
23	pessoa	Texto	Variável posteriormente verificada como inútil. Todos os registros tem o mesmo valor, 'pessoa jurídica'
24	qtd_total_homologada	Numérico	Quantidade de itens comprados
25	valor_total_homologado	Numérico	Valor da compra (quantidade x preço individual)



Transferências Voluntárias (SICONV)

Diferentemente do Siape e Siasg, os dados do SICONV são abertos, fazendo que a maior contribuição da Enap para o uso dessa base, não seja especificamente o acesso aos dados, mas a facilitação desse acesso e na adição de algumas variáveis adicionais.

A facilitação de acesso é relevante, pois os dados do SICONV vem em um arquivo zip com 39 arquivos distintos. A sintetização desse grande conjunto de informações separadas em arquivos separados em uma tabela de fácil acesso foi considerado um objeto de grande valia e utilidade para a produção de conhecimento sobre o tema.

A simplificação do acesso foi feita utilizando informação de 7 das tabelas disponibilizadas pelo Siconv: “convenio”, “emenda”, “historico_situacao”, “pagamento”, “programa”, “programa_proposta” e “proposta”.

Nesse contexto de transferências, a Enap definiu três unidades de análise: os instrumentos, as propostas e os proponentes. Os proponentes são aqueles que solicitam recursos ao governo federal (geralmente municípios, embora existam outros casos). As propostas são as solicitações dos proponentes e as propostas aceitas e passadas pelo trâmite burocrático se tornam os instrumentos.

Para a utilização prática, as propostas e instrumentos ficam em uma tabela de mesma estrutura, a diferença sendo apenas um filtro para separar as tabelas que têm informação na coluna “ano_instrumento” das que não chegaram ao estágio de instrumento celebrado.

Uma adição de conhecimento importante foi a classificação da variável texto “descricao_proposta”. Por ser uma variável aberta e sem limites de caracteres, havia grande dificuldade no Ministério da Economia de extrair informação útil desse campo. Nesse contexto, a Enap aplicou algoritmo de NLP para uma classificação não supervisionada dos textos referidos gerando classificação nas seguintes variáveis:

- classificacao_nivell: o tema do instrumento. Os temas são: “saúde”, “segurança e cidadania”, “infraestrutura urbana”, “educação”, “turismo, cultura e esportes” e “Agricultura”

- classificacao_nivel2: dentro do tema, qual a forma de entrega do benefício à população. As opções são: “obras”, “equipamentos”, “capacitação”, “eventos e projetos”
- classificacao_nivel3: dentro do universo de “obra de infraestrutura” e “obra de turismo, cultura e esportes”, qual o tipo de obra.

Instrumentos / Propostas

Index	Nome Variável	Tipo	Comentário
1	id_proposta_ano	Texto	Id da proposta + ano da proposta
2	id_proposta	Texto	Id da proposta
3	id_convenio	Texto	Id do instrumento
4	descricao_proposta	Texto	Descrição da proposta
5	nome_programa	Texto	Nome do programa pelo qual a proposta tenta conseguir recursos
6	acao_orcamentaria	Texto	Ação orçamentária dentro do programa
7	classificacao_nivel1	Texto	Classificação feita pela Enap, classificando nos setores: 'Saúde', 'Segurança e cidadania', 'Infraestrutura Urbana', 'Educação', 'Turismo, cultura e esportes', 'Agricultura'
8	classificacao_nivel2	Texto	Classificação feita pela Enap, classificando nos tipos: 'Obras', 'Equipamentos', 'Capacitação', 'Eventos e projetos'
9	classificacao_nivel3	Texto	Classificação feita pela Enap, classificando as obras de 'Infraestrutura Urbana' ou 'Turismo, teatros e cinemas' em subtipos
10	ano_proposta	Texto	Ano da proposta
11	ano_instrumento	Texto	Ano do instrumento
12	dummy_instrumento	Númerico	Uma binária (1: virou instrumento, 0: ficou no estágio de proposta) para simplificar contas do ponto de vista de proponentes
13	data_proposta	Data	Data de envio da proposta
14	ano_inicio_vigencia_convenio	Númerico	Ano de início de vigência do instrumento
15	ano_final_vigencia_convenio	Númerico	Ano final da vigência do instrumento
16	intervalo_vigencia_dias	Númerico	dia_fim_vigenc_conv' - 'dia_inic_vigenc_conv'
17	dataassinatura_convenio	Data	Data de assinatura do convênio
18	datapublicacao_convenio	Data	Data de publicação do convênio
19	datalimiteprestacao_contas	Data	Data limite de prestação de contas

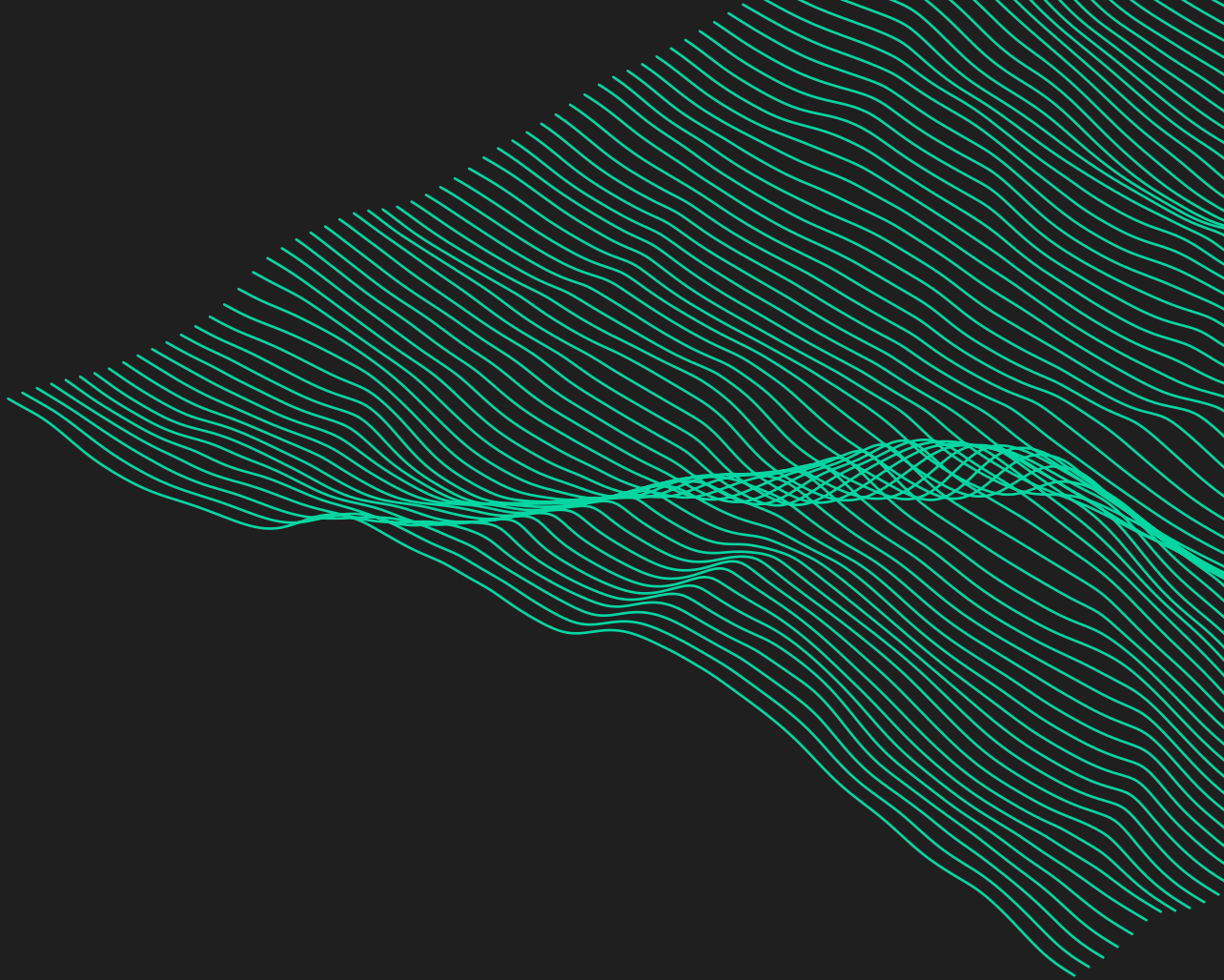
Index	Nome Variável	Tipo	Comentário
20	situacao_proposta_grupo	Texto	Agrupar-se as opções em 'Aprovada', 'Rejeitada', 'Em processo'
21	situacao_proposta	Texto	Situação da proposta
22	modalidade	Texto	Modalidade da proposta
23	situacao_contratacao	Texto	Situação da contratação
24	dummy_instrumento_ativo	Númerico	Troca-se de 'SIM' ou 'NÃO' para binária (1/0)
25	situacao_instrumento	Texto	Situação do instrumento
26	subsituacao_instrumento	Texto	Subsituação do instrumento
27	indicador_obtv	Texto	Indicador OBTV
28	data_inicio_execucao	Texto	Primeira data em que historico_sit vale 'EM_EXECUCAO' ou 'ASSINADA'
29	ano_inicio_execucao	Texto	Mesma coisa que a anterior, mas só o ano
30	data_saida_execucao	Timestamp	Primeira data em que historico_sit vale 'CONVENIO_ANULADO', 'CONVENIO_RESCINDIDO', 'PRESTACAO_CONTAS_APROVADA', 'PRESTACAO_CONTAS_APROVADA_COM_RESSALVAS'
31	ano_saida_execucao	Numérico	Mesma coisa que a anterior, mas só o ano
32	data_decisao_proposta	Timestamp	Primeira data em que historico_sit vale 'PROPOSTA_REPROVADA', 'PROPOSTA_REJEITADA_IMPEDIMENTO_TECNICO', 'PROPOSTA_ELIMINADA_EM_CHAMAMENTO_PUBLICO', 'ASSINADA' ou 'EM_EXECUCAO'
33	ano_decisao_proposta	Numérico	Mesma coisa que a anterior, mas só o ano
34	valor_total_proposta	Numérico	Valor total da proposta
35	valor_repasse_uniao_proposta	Numérico	Valor solicitado da união na proposta
36	valor_contrapartida_proposta	Numérico	Valor de contrapartida na proposta
37	valor_total_instrumento	Numérico	Valor total do instrumento
38	valor_repasse_uniao_instrumento	Numérico	Valor repassado pela união no instrumento
39	valor_contrapartida_instrumento	Numérico	Valor de contrapartida no instrumento
40	valor_empenhado_instrumento	Numérico	Valor empenhado no instrumento
41	valor_desembolsado_instrumento	Numérico	Valor desembolsado no instrumento
42	saldo_remanescente_tesouro_instrumento	Numérico	Saldo remanescente no tesouro
43	saldo_remanescente_conveniente_instrumento	Numérico	Saldo remanescente no conveniente

Index	Nome Variável	Tipo	Comentário
44	valor_rendimento_aplicacao_instrumento	Numérico	Valor do rendimento dos recursos do instrumento
45	valor_ingresso_contrapartida_instrumento	Numérico	Ingresso de contrapartida do instrumento
46	valor_saldo_instrumento	Numérico	Saldo do instrumento
47	valor_executado	Numérico	Somatória de todos os pagamentos de um mesmo convênio
48	quantidade_termos_aditivos	Texto	Quantidade de termos aditivos
49	quantidade_prorrogacoes	Texto	Quantidade de prorrogações do instrumento
50	valor_total_original_instrumento	Numérico	Prefere-se o 'vl_global_original_conv', mas em sua ausência vale o 'vl_global_original_conv'
51	data_final_original	Data	Data final do instrumento antes de prorrogações
52	intervalo_original	Numérico	Intervalo do instrumento antes de prorrogações
53	emenda_impositiva	Texto	Quando 'valor_emenda_impositiva' > 0 vale 'Emenda Impositiva'; quando 'valor_emenda_impositiva' igual a 0 vale 'Emenda Não-Impositiva' e quando 'valor_emenda_impositiva' sem valor (nulo) vale 'Sem emenda'
54	valor_emenda	Numérico	Valor de recursos via emenda
55	valor_emenda_impositiva	Numérico	Valor de recursos via emenda impositiva
56	emenda_individual	Texto	Pelo menos 1 emenda do tipo 'INDIVIDUAL'
57	emenda_bancada	Texto	Pelo menos 1 emenda do tipo 'BANCADA'
58	emenda_comissao	Texto	Pelo menos 1 emenda do tipo 'COMISSAO'
59	emenda_relator_geral	Texto	Pelo menos 1 emenda do tipo 'RELATOR GERAL'
60	emenda	Texto	Pelo menos 1 emenda de um dos 4 tipos citados acima
61	tipo_emenda	Texto	O tipo de emenda dominante. Relator geral > comissao > bancada > individual
62	regiao_proponente	Texto	Juntando as ufs nas 5 regiões
63	uf_proponente	Texto	Uf do proponente
64	municipio_proponente	Texto	Município do proponente
65	nm_proponente	Texto	Nome do proponente
66	natureza_juridica	Texto	natureza jurídica do proponente
67	orgao_sup	Texto	Órgão superior do órgão que recebe a proposta
68	orgao	Texto	Órgão que recebe a proposta

Proponentes

A tabela de proponentes é gerada com base na tabela de propostas, de maneira a juntar as propostas de um mesmo proponente consolidando informações sobre os solicitantes de recursos.

Index	Nome Variável	Tipo	Comentário
1	nm_proponente	Texto	No caso de natureza_juridica 'Administração Pública Municipal' ou 'Administração Pública Estadual ou do Distrito Federal', junta-se todos os proponentes da cidade ou estado relevante como um só. Preferimos isso a considerar a 'Secretaria de Educação do Município de x' como um proponente distinto da 'Secretaria de Saúde do município de x', como é feito nos dados originais.
2	natureza_juridica	Texto	Natureza jurídica
3	regiao_proponente	Texto	Juntando nas 5 regiões
4	uf_proponente	Texto	UF do proponente
5	municipio_proponente	Texto	Município do proponente
6	num_propostas	Numérico	Conta-se a quantidade de propostas do mesmo proponente ao longo do histórico da base de dados
7	quantidade_propostas_via_emenda	Numérico	Contagem de propostas com emendas do legislativo
8	quantidade_instrumentosassinados	Numérico	Soma de 'dummy_instrumento', a variável que vale 1 caso a proposta tenha virado um instrumento e 0 se permaneceu proposta
9	quantidade_de_subunidades	Numérico	Contagem de subunidades que foram unidas como o mesmo proponente
10	valor_repasse_uniao_instrumento	Numérico	Soma para o mesmo proponente
11	valor_repasse_solicitacao_proposta	Numérico	Soma para o mesmo proponente
12	ano_ingresso_na_base	Numérico	Da data da proposta mais antiga, extrai-se o ano
13	mes_ingresso_na_base	Numérico	Da data da proposta mais antiga, extrai-se o mês
14	dia_ingresso_na_base	Numérico	Da data da proposta mais antiga, extrai-se o dia
15	data_minima	Data	Data da proposta mais antiga



RegBR

O RegBR compila e apresenta dados de atos normativos brasileiros desde 1950. As atualizações serão semestrais.

Index	Nome Variável	Tipo	Descrição
1	id	Número	id único do ato normativo, segundo uma regra de criação interna
2	texto_lei	Texto	Texto do normativo, após limpeza inicial
3	situacao	Texto	Situação do ato normativo
4	tipo_lei	Texto	Tipo de normativo, vide tabela abaixo
5	setor	Texto	Setor da economia que o normativo rege vide tabela abaixo
6	agencia_reguladora	Texto	Se é normativo de Agência reguladora (e qual) ou Não. Agências reguladoras estudadas
7	ano	Número	Ano de publicação do normativo
8	situacao_infogov	Texto	Situação do ato normativo, com agrupamento de algumas classes

Setor Econômico

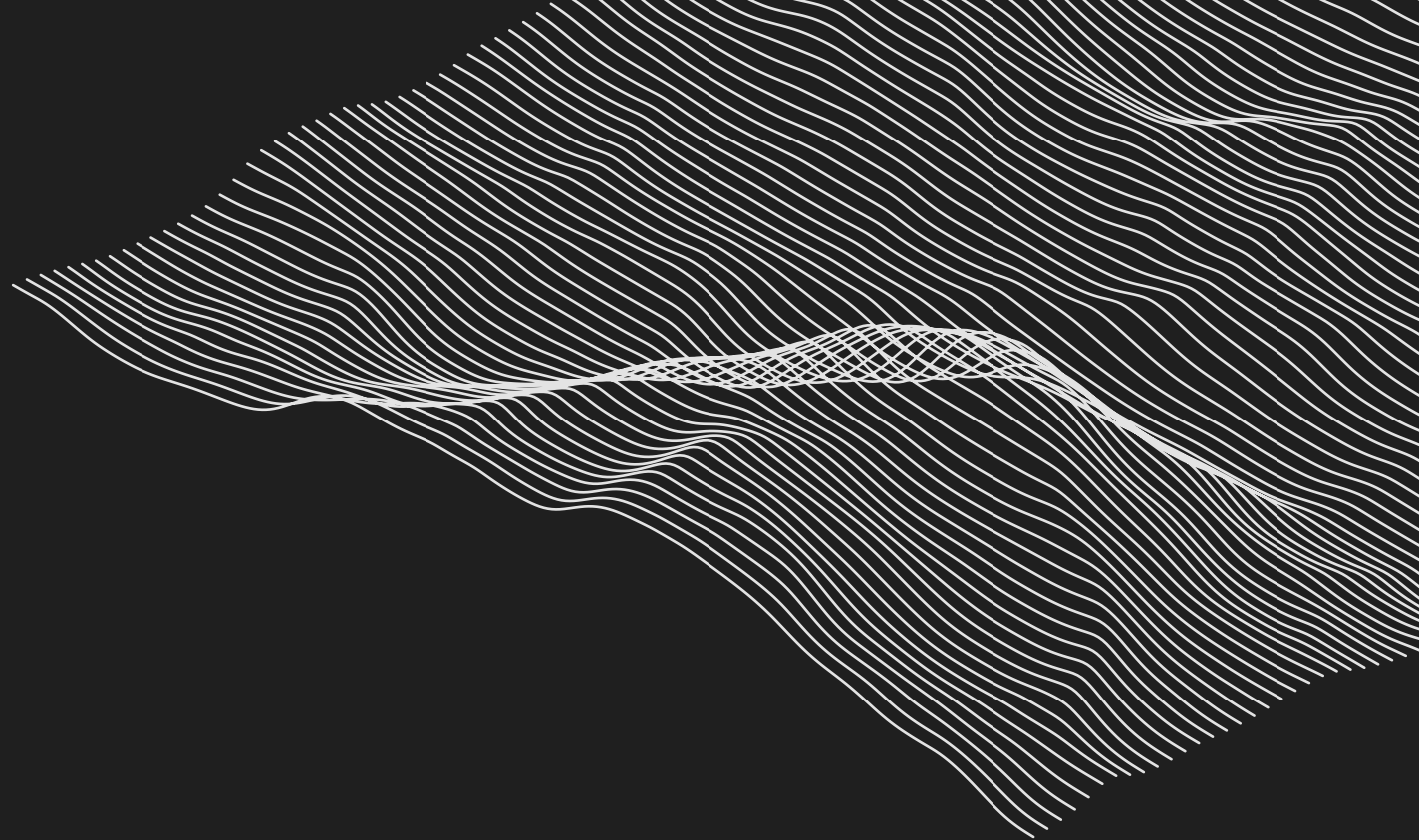
Index	Setor Econômico
1	Agricultura, pecuária, produção florestal, pesca e aquicultura
2	Indústria extrativas
3	Indústrias de transformação
4	Eletricidade e Gás
5	Água, esgoto, atividades de gestão de resíduos e descontaminação
6	Construção
7	Comércio, acomodação e serviços imobiliários
8	Transporte, armazenagem e correios
9	Informacao e comunicacao
10	Finanças, seguros e serviços relacionados
11	Atividades profissionais, científicas e técnicas
12	Atividades administrativas e serviços complementares
13	Administração pública, defesa e seguridade social
14	Educação
15	Saúde e serviços sociais
16	Artes, cultura, esportes e recreação
17	Outros serviços
18	Não-regulatório

Tipo de ato normativo

Index	Tipo de ato normativo
1	Emenda constitucional
2	Lei Ordinária
3	Decreto-lei
4	Medidas provisória
5	Lei complementar
6	Decreto
7	Resolução
8	Portaria
9	Instruções normativas
10	Súmulas

Agências reguladoras estudadas

Index	Tipo de ato normativo
1	ANA - Agência Nacional de Águas e Saneamento Básico
2	ANCINE - Agência Nacional do Cinema
3	ANEEL - Agência Nacional de Energia Elétrica
4	ANP - Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
5	ANM - Agência Nacional de Mineração
6	ANATEL - Agência Nacional de Telecomunicações
7	ANS - Agência Nacional de Saúde Suplementar
8	ANVISA - Agência Nacional de Vigilância Sanitária
9	ANAC - Agência Nacional de Aviação Civil
10	ANTAQ - Agência Nacional de Transportes Aquaviários
11	ANTT - Agência Nacional de Transportes Terrestres



info GOV

ENAP

<https://infogov.enap.gov.br/>