

## I - MOTIVAÇÃO

Análise de algumas bases de dados públicas do Ibama, discriminadas mais a frente, com o objetivo de verificar pontos que possam ser melhorados para facilitar o acesso aos dados pela população em geral.

## II - FONTES

Para fazer isso, iniciamos um processo de limpeza, organização e estruturação das informações disponibilizadas pelo Ibama:

Embargos:

- [Dados abertos - Termos de Embargo - CSV](#) (para simplificar, usaremos "te\_csv" para referir a essa tabela)
- [Siscom/ Geoserver/ Dados Abertos - Termos de Embargo - SHP](#) (vamos usar "te\_shp" para nos referirmos a colunas dessa tabela)

Autos de Infração:

- [Dados Abertos - Autos de Infração](#)
- [Siscom](#)

Autorizações

- **Dados abertos**
  - [UAS - Uso alternativo do Solo](#)
  - [ASV - Autorização de Supressão da Vegetação](#)
- **Siscom**
  - [ASV - Área de Autorização de Supressão de Vegetação](#)
  - [UAS - Área Solicitada Para Uso Alternativo do Solo](#)

Em paralelo, para confirmar algumas informações, também utilizamos informações dos sistemas abaixo:

- [Consulta Autuações Áreas Embargadas](#)
- [SEI - Ibama](#)

## III - OBSERVAÇÕES GERAIS

Após análise das bases acima, identificamos alguns pontos que dificultam a importação das informações para um banco de dados e, por essa razão, também exigem um trabalho maior e conhecimento técnico específico para utilizar esses dados. Nesse sentido, seguem algumas observações quanto a isso:

- máscara no CPF. O padrão adotado por diversos órgãos federais (por exemplo, [CGU](#), [RFB](#)) é o de ocultar os primeiros três e dois últimos dígitos do número, com base no [artigo 129, §2º da Lei nº 13.473/2017](#), que apesar de não determinar a obrigatoriedade da adoção desse padrão, trata da divulgação dados pessoais de empregados no caso de contratação de empresas para prestação de serviços. A adoção deste padrão é importante porque permite, junto com outros atributos, individualizar os registros, sem perder o anonimato.
- CNPJ não pode ser mascarado, já que a [Lei nº 13.709/2018 - LGPD](#), em seu artigo 1º, restringe suas normas especificamente às pessoas naturais:

Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.

Nesse sentido: [PARECER n. 00295/2020/CONJUR-CGU/CGU/AGU](#).

- Nem todas as tabelas possuem dicionário de dados;
- Não conseguimos identificar nas fontes as datas de atualização dos dados;
- Data fora do formato: AAAA-MM-DD e colunas que possuem data e hora, sendo que todos os dígitos do horário correspondem a zero (isso apenas deixa o arquivo mais pesado);
- Ausência de padrão da unidade métrica (m² ou ha);

### III - INCONSISTÊNCIAS ESPECÍFICAS

#### A. EMBARGOS

Em relação aos **embargos**, comparando as duas tabelas citadas acima, chegamos a algumas conclusões, que podem ser visualizadas na tabela ao fim da página:

- Com a última versão da tabela que temos, verificamos que elas não têm as mesmas quantidades de colunas: a **te\_shp** possui 70 colunas e a **te\_csv**, 38;
  - Obs.: em setembro de 2021, **te\_shp** possuía 67 colunas e a **te\_csv**, 41.
- As tabelas, extraídas na mesma data, também não possuem a mesma quantidade de registros:
  - até o início de março, a **te\_shp** possuía 96.102 únicos e a **te\_csv** 82.277 únicos
    - Por cerca de um mês, a **te\_shp** possuía cerca de 70000 registros e não conseguimos identificar a razão da diminuição de mais de 25 mil registros na tabela que tem origem do siscom. Recentemente, a quantidade de registros da **te\_shp** voltou para 96877.

- A partir dos nomes das colunas suas respectivas quantidades de valores não nulos, tentamos encontrar quais delas seriam as melhores para realizar o cruzamento entre as tabelas, para unificar as informações. Para isso, seguimos os seguintes passos:
  - Identificamos, a partir das nomes das colunas, de seus conteúdos e da quantidade registros, quais seriam as melhores para realizar esse cruzamento;
  - Com essa análise, concluímos que as melhores chaves de cruzamento seriam as colunas **te\_csv.num\_tad** e **te\_csv.seq\_tad**, **te\_shp.numero\_tad** e **te\_shp.seq\_tad**.
  - Entretanto, dentre todos os cenários que realizamos, chegamos apenas cerca de 70.000 registros coincidentes entre as tabelas, utilizando as colunas **te\_shp.numero\_tad** e **te\_csv.num\_tad**. Para isso, utilizamos a linguagem SQL e realizamos apenas um *join* entre as tabelas, sem analisar efetivamente se há correspondência entre os registros.
  - Sobre as colunas **te\_csv.num\_tad** e **te\_shp.numero\_tad**, verificamos uma diferença considerável de quantidade de caracteres entre os registros (tabela abaixo)
    - Em relação à **te\_csv.num\_tad**, os registros que possuem 6 ou menos caracteres parecem possuir apenas números, enquanto os com mais de 6 aparentemente letras e, de acordo com o número do processo, são mais de recentes (2020/2021);
    - Já sobre os registros da **te\_shp.numero\_tad**, filtramos por uma quantidade maior que 10 e percebemos que:
      - os números maiores têm várias composições diferentes:
        - alguns são dois números separados por "e",
        - outros são números aparentemente fora do padrão,
        - outros parecem iniciar com quatro dígitos que parecem corresponder a um ano, mas são incompatíveis com o ano do processo.

shp_qtd_ caracteres	shp_qtd_ registros	te_qtd_ caracteres	te_qtd_ registros
6	62442	6	70082
8	12799	5	6713
9	10333	8	3390
5	8844	4	2094
4	1869	3	163
10	254	7	45
3	148	2	5
7	81	10	2
11	73	1	2
15	10	9	1

2	4		
12	3		
19	2		
1	2		
18	2		
21	2		
43	1		
20	1		
35	1		
17	1		
<i>null</i>	0		

## B. AUTOS DE INFRAÇÃO

Em relação à tabela de **autuações**, identificamos também algumas informações inconsistentes, dentre elas, o número do processo administrativo. Como existem diversos formatos do número do PA no sistema, utilizamos as diretrizes da Portaria Interministerial nº 11/2019 para desenvolver uma função que converte os números para o formato do NUP - Número Único de Protocolo.

Dentre as diversas verificações que realizamos para confirmar o resultado dessa função, encontramos uma inconsistência em relação a alguns números de processo, conforme descrito abaixo (a coluna em vermelho é a única que foi criada por nós, todo o restante foi extraído diretamente do csv de autos de infração disponível no dados abertos):

nome_infrator	municipio	uf	num_processo	Num_processo_completo	Dat_hora_auto_infracao
GENESIO LANGA	COLATINA	ES	0200900175420	02009.001754/ 2020-84	17/09/2004 00:00
VIX STONES COMÉRCIO IMPORTAÇÃO E EXPORTAÇÃO LTDA	VILA VELHA	ES	0200900175420	02009.001754/ 2020-84	28/09/2005 00:00
RODRIGO LEAL CORREIA	ITAPEMIRIM	ES	02009001754202084	02009.001754/ 2020-84	24/08/2020 17:00

Após análise desses dados, verificamos que os registros só tinham em comum a Unidade Federativa e que os anos das infrações eram muito distantes do ano do PA, razão pela qual pesquisamos pelo documento dos autuados no Sistema de Consulta de Autuações e Embargos.

Como se pode visualizar pelas consultas e respectivos resultados abaixo, os processos que apareceram são exatamente iguais aos que chegamos por meio de nossa função (tabela acima):

**Consulta de Autuações Ambientais e Embargos**

Consulta Pública

Embargos  Autuações Ambientais

**Autuações Ambientais**

Dados do Autuado

CPF/CNPJ: 06.132.947/0001-32

Nome ou Razão Social:

Dados da Infração

Tipo de Infração: --selezione--

Bioma: --selezione--

Estado: --selezione-- Município: --VILTO--

Período de \* 01/01/2005 até 31/12/2005

MINISTÉRIO DO MEIO AMBIENTE - MMA  
 INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS - IBAMA  
 DIRETORIA DE PROTEÇÃO AMBIENTAL

**Relatório de Autuações Ambientais**

Estado	Município	CNPJ/CPF	Nome Autuado	Nº A.I.	Valor Multa	Nº Processo
ESPIRITO SANTO	VILA VELHA	06.132.947/0001-32	VIX STONES COMÉRCIO IMPORTAÇÃO E EXPORTAÇÃO LTDA	305969	1.500,00	02009.001754/2020-84

**Consulta de Autuações Ambientais e Embargos**

Consulta Pública

Embargos  Autuações Ambientais

**Autuações Ambientais**

Dados do Autuado

CPF/CNPJ: 133.214.387-39

Nome ou Razão Social:

Dados da Infração

Tipo de Infração: --selezione--

Bioma: --selezione--

Estado: --selezione-- Município: --VILTO--

Período de \* 01/01/2020 até 31/12/2020

MINISTÉRIO DO MEIO AMBIENTE - MMA  
 INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS - IBAMA  
 DIRETORIA DE PROTEÇÃO AMBIENTAL

**Relatório de Autuações Ambientais**

Estado	Município	CNPJ/CPF	Nome Autuado	Nº A.I.	Valor Multa	Nº Processo
ESPIRITO SANTO	ITAPEMIRIM	133.214.387-39	RODRIGO LEAL CORREIA	MEGTISA	66.000,00	02009.001754/2020-84

**Consulta de Autuações Ambientais e Embargos**

Consulta Pública

Embargos  Autuações Ambientais

**Autuações Ambientais**

Dados do Autuado

CPF/CNPJ: 157.144.547-15

Nome ou Razão Social:

Dados da Infração

Tipo de Infração: --selezione--

Bioma: --selezione--

Estado: --selezione-- Município: --VILTO--

Período de \* 01/01/2004 até 31/12/2004

MINISTÉRIO DO MEIO AMBIENTE - MMA  
 INSTITUTO BRASILEIRO DO MEIO AMBIENTE E DOS RECURSOS NATURAIS RENOVÁVEIS - IBAMA  
 DIRETORIA DE PROTEÇÃO AMBIENTAL

**Relatório de Autuações Ambientais**

Estado	Município	CNPJ/CPF	Nome Autuado	Nº A.I.	Valor Multa	Nº Processo
ESPIRITO SANTO	COLATINA	157.144.547-15	GENESIO LANGA	107391	500,00	02009.001754/2020-84
ESPIRITO SANTO	COLATINA	157.144.547-15	GENESIO LANGA	107390	500,00	02009.001755/2020-29
ESPIRITO SANTO	COLATINA	157.144.547-15	GENESIO LANGA	107389	50,00	02009.001756/2020-73

Entretanto, quando realizamos a busca pelo número "200900175420" no [SEI](#), chegamos ao seguinte resultado:

6 resultados

Fiscalização Ambiental: Apuração de Infrações Ambientais N° 02009.001754/2020-84 Unidade Geradora: SUPES-ES Data: 02/09/2020	02009.001754/2020-84
Fiscalização Ambiental: Apuração de Infrações Ambientais N° 02009.001754/2019-41 Unidade Geradora: NEA-ES Data: 16/08/2019	02009.001754/2019-41
Arrecadação: Cobrança N° 02009.001754/2018-60 Unidade Geradora: NUARRE-ES Data: 20/06/2018	02009.001754/2018-60
Fiscalização Ambiental: Apuração de Infrações Ambientais N° 02009.001754/2008-98 Unidade Geradora: UT-CACHOEIRO DE ITAPEMIRIM-ES Data: 10/12/2008	02009.001754/2008-98
Fiscalização Ambiental: Apuração de Infrações Ambientais N° 02009.001754/2006-26 Unidade Geradora: SUPES-RO Data: 25/09/2006	02009.001754/2006-26
Fiscalização Ambiental: Apuração de Infrações Ambientais N° 02009.001754/2005-45 Unidade Geradora: NUIP-ES Data: 28/09/2005	02009.001754/2005-45

Ao clicarmos no primeiro número, que deveria corresponder a um dos três autuados, como visto acima, o resultado do SEI não traz qualquer informação além daquelas demonstradas na página de resultado de busca, não sendo possível verificar sequer a qual autuado o processo se refere:

## Pesquisa Processual

Autuação	
Processo:	02009.001754/2020-84
Tipo:	Fiscalização Ambiental: Apuração de Infrações Ambientais
Data de Registro:	02/09/2020
Interessados:	

Informações quanto à solicitação de vistas, cópia ou acompanhamento integral do processo: <https://www.gov.br/ibama/pt-br/ace>

Por outro lado, ao acessarmos o resultado do processo de número 02009.001754/**2005-45**", temos acesso ao processo da VIX, que no outro sistema aparece como "02009.001754/**2020-84**":

## Pesquisa Processual

Autuação	
Processo:	02009.001754/2005-45
Tipo:	Fiscalização Ambiental: Apuração de Infrações Ambientais
Data de Registro:	28/09/2005
Interessados:	VIX STONES COMERCIO IMPORTAÇÃO E EXPORTAÇÃO LTDA (06.132.947/0001-32)

Lista de Andamentos (33 registros):

Data/Hora	Unidade	
06/08/2019 11:42	GT-DBA-ES	Processo recebido na unidade
29/07/2019 11:14	GT-DBA-ES	Processo remetido pela unidade SUPES-ES
29/07/2019 11:14	SUPES-ES	Remoção de sobrestamento

## C. AUTORIZAÇÕES

### 1. ASV

Analisamos duas bases, uma proveniente da sinaflor ("**autorizacao\_asv\_shp**") e outra dos dados abertos ("**autorizacao\_asv**").

dados	autorizacao_asv_shp	autorizacao_asv
Total de linhas	6283	76504
Quantidade de colunas	10	77
Total de valores (sem null) coluna car	2179	12449
Total de valores únicos colunas car	710	920

Identificamos um aumento considerável (quase o triplo) da quantidade de autorizações na tabela **autorizacao\_asv\_shp** no período de aproximadamente um mês, sem uma razão aparente para isso (o total de registros disponíveis nessa tabela para 2022 é de aproximadamente 300).

### 2. UAS

Analisamos duas bases, uma proveniente da sinaflor (“**autorizacao\_uas\_shp**”) e outra dos dados abertos (“**autorizacao\_uas**”).

<b>dados</b>	<b>autorizacao_uas_shp</b>	<b>autorizacao_uas</b>
Total de linhas	8403	2035583
Quantidade de colunas	10	75
Total de valores (sem null) coluna car	5375	2035583
Total de valores únicos colunas car	2262	2145

Há uma quantidade imensa de linhas para a mesma autorização, na qual todas as informações se repetem, e aparentemente as únicas informações que se alteram são a espécie das árvores, o que deixa a tabela com um tamanho gigantesco e dificulta o acesso à informação.

Sugestões para resolver essa situação seriam agrupar as espécies na mesma coluna ou fazer uma tabela derivada, apenas com essa informação, vinculando à tabela principal por meio de uma chave primária.

#### **IV - CONCLUSÕES**

Analisando as bases, não é possível:

1. Compreender a razão da diferença de registros entre bases de embargos extraídas na mesma data, mas de sistemas diferentes;
2. Identificar a chave primária que podemos utilizar para realizar, com segurança, o cruzamento de registros entre as tabelas;
3. Verificar, pelo número do processo, com segurança, qual é o autuado (informações divergentes entre sites de consulta oficiais);
4. Ainda sobre os Embargos, definir um padrão que a coluna **num\_tad** / **numero\_tad** deveria seguir;
5. Compreender a razão da alteração na quantidade de registros na tabela de embargos proveniente do sinaflor (diminuição de cerca de 25.000 e posterior reinclusão desses registros) e na estrutura da tabela;
6. Realizar o cruzamento com segurança entre as tabelas de autorizações provenientes do portal de dados abertos e do sinaflor, pela diferença na estrutura das tabelas;
- 7.

