

# Notebook meramente preparatório, mantendo os dados brutos. Serve, apenas, para concatenar as 27 tabelas de devedores da PFN, uma para cada UF. Resultando no arquivo "concatenadoordenado.csv", de 4GB.

Fonte: [https://dadosabertos.pgfn.gov.br/2021\\_trimestre\\_02/Dados\\_abertos\\_Nao\\_Previdenciario.zip](https://dadosabertos.pgfn.gov.br/2021_trimestre_02/Dados_abertos_Nao_Previdenciario.zip)  
([https://dadosabertos.pgfn.gov.br/2021\\_trimestre\\_02/Dados\\_abertos\\_Nao\\_Previdenciario.zip](https://dadosabertos.pgfn.gov.br/2021_trimestre_02/Dados_abertos_Nao_Previdenciario.zip)), acessado em 25/07/2021.

In [1]:

```
1 import pandas as pd
```

In [2]:

```
1 acre = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_lai
```

In [3]:

```
1 acre.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50998 entries, 0 to 50997
Data columns (total 13 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   CPF_CNPJ                             50998 non-null  object
1   TIPO_PESSOA                           50998 non-null  object
2   TIPO_DEVEDOR                           50998 non-null  object
3   NOME_DEVEDOR                           50998 non-null  object
4   UF_UNIDADE_RESPONSAVEL                 50998 non-null  object
5   UNIDADE_RESPONSAVEL                     50998 non-null  object
6   NUMERO_INSCRICAO                       50998 non-null  int64
7   TIPO_SITUACAO_INSCRICAO                 50998 non-null  object
8   SITUACAO_INSCRICAO                     50998 non-null  object
9   RECEITA_PRINCIPAL                       50998 non-null  object
10  DATA_INSCRICAO                         50998 non-null  object
11  INDICADOR_AJUIZADO                      50998 non-null  object
12  VALOR_CONSOLIDADO                       50998 non-null  float64
dtypes: float64(1), int64(1), object(11)
memory usage: 5.1+ MB
```

In [ ]:

```
1
```

In [4]:

```
1 alagoas = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_
```

In [ ]:

1

In [5]:

```
1 amazonas = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
```

In [ ]:

1

In [6]:

```
1 amapa = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
```

In [ ]:

1

In [7]:

```
1 bahia = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
```

In [ ]:

1

In [8]:

```
1 ceara = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
```

In [ ]:

1

In [9]:

```
1 distritofederal = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\
```

In [ ]:

1

In [10]:

```
1 espiritosanto = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\ar
```

In [ ]:

1

In [11]:

```
1 goias = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_1a
```

In [ ]:

```
1
```

In [12]:

```
1 maranhao = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo
```

```
2
```

In [ ]:

```
1
```

In [13]:

```
1 minasgerais = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arqu
```

```
2
```

In [ ]:

```
1
```

In [14]:

```
1 matogrossodosul = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\
```

```
2
```

In [ ]:

```
1
```

In [15]:

```
1 matogrosso = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arqui
```

```
2
```

In [ ]:

```
1
```

In [16]:

```
1 matogrossodosul = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\
```

```
2
```

In [ ]:

```
1
```

In [17]:

```
1 para = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
2
```

In [ ]:

```
1
```

In [18]:

```
1 paraiba = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_
2
```

In [ ]:

```
1
```

In [19]:

```
1 pernambuco = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arqui
2
```

In [ ]:

```
1
```

In [20]:

```
1 piaui = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_la
2
```

In [ ]:

```
1
```

In [21]:

```
1 parana = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_I
2
```

In [ ]:

```
1
```

In [22]:


```
1 riodejaneiro = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arc
2
```

In [ ]:

```
1
```

In [23]:

```
1 riograndedonorte = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_riogrande_donorte.csv')
2
```




In [ ]:

```
1
```

In [24]:

```
1 rondonia = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_rondonia.csv')
2
```



In [ ]:

```
1
```

In [25]:

```
1 roraima = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_roraima.csv')
2
```




In [ ]:

```
1
```

In [26]:

```
1 riograndedosul = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_riogrande_sul.csv')
2
```




In [ ]:

```
1
```

In [27]:

```
1 santacatarina = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_santa_catarina.csv')
2
```



In [ ]:

```
1
```

In [28]:

```
1 sergipe = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquivo_sergipe.csv')
2
```



In [ ]:

1

In [29]:

```
1 saopaulo = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquiv
```

In [ ]:

1

In [30]:

```
1 tocantins = pd.read_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\PFN\arquiv
```

**Concatenamos todos os 27, ordenando pela disposição alfabética das UF, resultando em um DataFrame de 20.090.612 registros.**

**Exportamos para o concatenadoordenado.csv.**

In [31]:

```
1 df_pfn = pd.concat([acre,alagoas,amazonas,amapa,bahia,ceara,distritofederal,espiritosar
```

In [34]:

```
1 df_pfn.shape
```

Out[34]:

```
(20090612, 13)
```

In [32]:

```
1 df_pfn.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 20090612 entries, 0 to 127506
Data columns (total 13 columns):
 #   Column                                Dtype
---  -
 0   CPF_CNPJ                             object
 1   TIPO_PESSOA                           object
 2   TIPO_DEVEDOR                           object
 3   NOME_DEVEDOR                           object
 4   UF_UNIDADE_RESPONSAVEL                 object
 5   UNIDADE_RESPONSAVEL                     object
 6   NUMERO_INSCRICAO                       int64
 7   TIPO_SITUACAO_INSCRICAO                 object
 8   SITUACAO_INSCRICAO                     object
 9   RECEITA_PRINCIPAL                       object
10   DATA_INSCRICAO                         object
11   INDICADOR_AJUIZADO                     object
12   VALOR_CONSOLIDADO                       float64
dtypes: float64(1), int64(1), object(11)
memory usage: 2.1+ GB
```

In [ ]:

```
1
```

In [ ]:

```
1 df_pfn.to_csv(r'C:\Users\73594253368\Desktop\Curso\Datasets\Procon\concatenadoordenado.
```

In [ ]:

```
1
```