

MANUAL DE INSTALAÇÃO

Instalar a versão PYTHON 3.6 ou superior

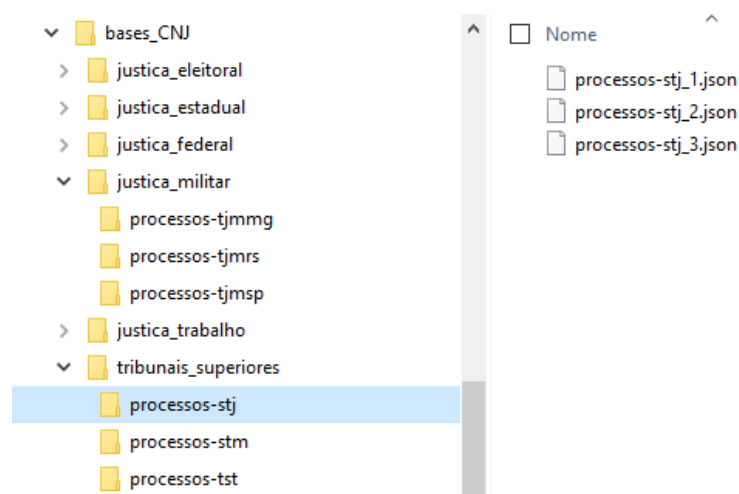
Instalar a biblioteca PANDAS:

➤ `Pip install pandas`

É necessário ter 2 sub-diretórios, um para gravação do LOG e outro para saída dos arquivos processados:



Os arquivos originais JSON (dados do DATAJUD) podem estar dentro de uma estrutura similar:



Em todos os códigos fonte, no início do corpo principal, tem a configuração dos diretórios de entrada, de LOG e de saída. Eles precisam ser, necessariamente alterados, conforme o diretório de instalação:

```
# -----  
# MAIN - Corpo Principal  
# -----  
  
# Constantes Iniciais - Parametros - CONFIGURAR NA INSTALACAO !!!  
# -----  
• dir_input = 'C:\\bases_CNJ'      # diretorio RAIZ de entrada (pra pegar todos os arquivos JSON)  
• dir_output = 'D:\\CNJ\\output\\'  # diretorio de saida  
• dir_log = 'D:\\CNJ\\LOG\\'        # diretorio de LOG  
# -----
```

```
# -----  
# MAIN - Corpo Principal  
# -----  
  
# Constantes Iniciais - Parametros - CONFIGURAR NA INSTALACAO !!!  
# -----  
• dir_input = 'D:\\CNJ\\output\\'  # diretorio de entrada  
• dir_output = 'D:\\CNJ\\output\\'  # diretorio de saida  
• dir_log = 'D:\\CNJ\\LOG\\'        # diretorio de LOG  
# -----
```

Os programas precisam ser executados na sequência, conforme documento de apresentação.

No nome dos programas tem a ordem de execução.

`cleanjud_1_parse_json.py`

`cleanjud_2_pivot_tbus.py`

`cleanjud_3_inconsistencias.py`

`cleanjud_4_machine_learning.py`

`cleanjud_5_correcao.py`

Para executar:

➤ `python <nome_do_programa>`

```
D:\CNJ>python cleanjud_2_pivot_tbus.py
Inicio execucao : 2020-10-21 12:34:54
DONE!
```

```
D:\CNJ>python cleanjud_3_inconsistencias.py
Inicio execucao : 2020-10-21 12:36:08
Arq de Entrada: D:\CNJ\output\cnj_datajud.txt
Arq de Saida: D:\CNJ\output\cnj_rel_inconsistencia.txt
... carregando parametros...
... leitura arquivo de entrada...
... inicio processamento relatorio inconsistencias...
#
... gravando arquivo de relatorio ...
DONE!
```

```
D:\CNJ>python cleanjud_4_machine_learning.py
Inicio execucao : 2020-10-21 12:37:09
Arq de Entrada: D:\CNJ\output\cnj_datajud_OK.txt
Arq Machine Learning: D:\CNJ\output\cleanjud_machine_learning.txt
Arq Saida Outliers: D:\CNJ\output\cleanjud_anomalias.txt
... carregando parametros...
... leitura do arquivo de entrada...
... inicio Modelo Aprendizado ...
... Agrupa Variaveis de Comparacao ...
#####
... Calculo dos Percentuais Distribuicao ...
... Ordenacao dos Percentis ...
... Calculo do Percentual Acumulado ...
... gravando a base de conhecimento do Machine Learning ...
... Carrega a base de conhecimento pra memoria ...
... Grava Arquivo de OUTLIERS (provaveis anomalias) ...
DONE!
```

```
D:\CNJ>python cleanjud_5_correcao.py
Inicio execucao : 2020-10-21 12:42:56
Arq de Entrada: D:\CNJ\output\cnj_datajud.txt
Arq de Saida: D:\CNJ\output\cnj_datajud_corrigido.txt
... carregando parametros...
... leitura arquivo de entrada...
... inicio processamento das CORRECOES - SANEAMENTO...
##
... gravando arquivo SANEADO ...
DONE!
```

São necessário também MANTER no diretório da aplicação os arquivos de PARAMETROS abaixo:

- cleanjud_de_para_cnj7_orgao.csv
- cleanjud_de_para_tribunal_sistema.csv
- cleanjud_hierarquia_assunto.csv
- cleanjud_hierarquia_serventia.csv
- cleanjud_movimento.csv
- cleanjud_par_e_par.csv
- cleanjud_serventias.csv
- cleanjud_sgt_assunto.csv
- cleanjud_sgt_classe.csv

Os LOGS de execucao serão gravados no subdiretorio LOG