



Analizando datos sobre

# ENERGIA



# Insighters

Consultoria

# Integrantes



**Adriano Kim**



**José Aurelio**



**Marcos Paulo**



**Paulo Vitorino**



**Renato G. Costa**



**Wesley Luiz**

# TRANSIÇÃO ENERGÉTICA NO BRASIL



## Panorama Nacional

Cenário de energia até o momento e onde estamos



## Taxa de crescimento

Qual fonte de energia está com maior taxa de crescimento?



## Sustentabilidade

Como estão distribuídas as fontes de energia sustentável com maior crescimento?



## Case de estudo - Eletromobilidade

O uso da fonte solar para a expansão da oferta de veículos híbridos e elétricos.

# Informações dos dados

**Capacidade instalada por região, UF e por fonte geradora**

**Capacidade de geração por fonte**

**Capacidade instalada de usinas solares e eólicas**

**Geração de energia de usinas solares e eólicas, por fonte, região e UF  
(31 Datasets, somando um total de mais de 8.600.000 registros)**

**Geração e capacidade por geração distribuída e distribuição geográfica  
(1 dataset com mais de 2 milhões de registros)**

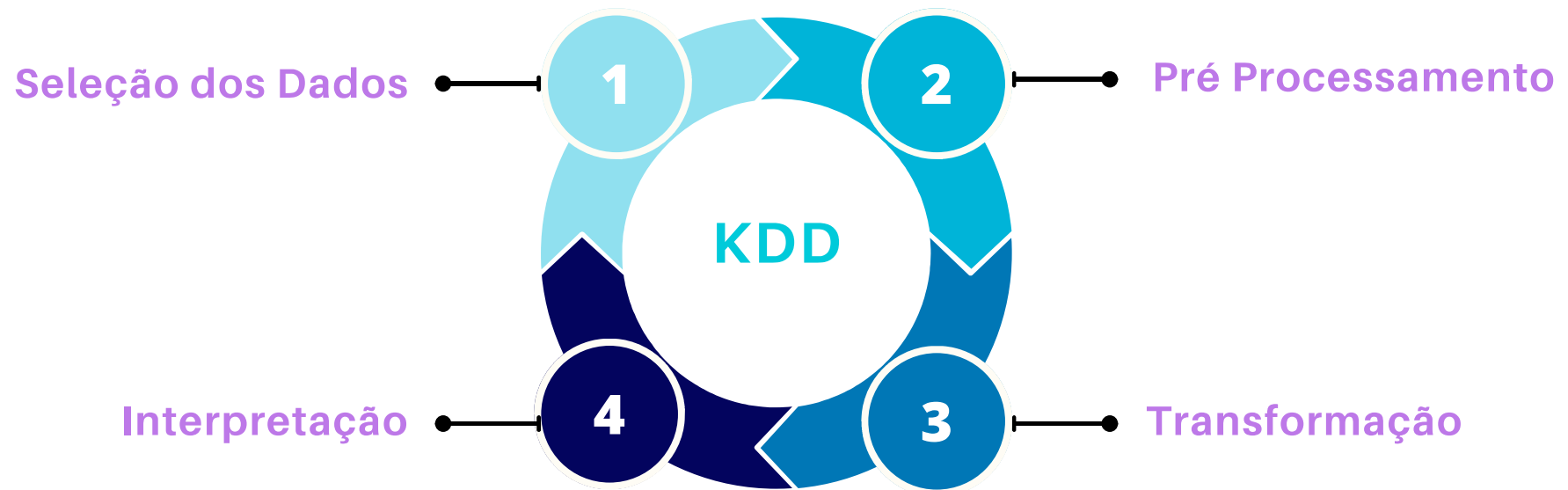
**Quantidade de pontos de recarga de veículos elétricos, por ano**

**Quantidade de vendas de veículos elétricos, por ano**

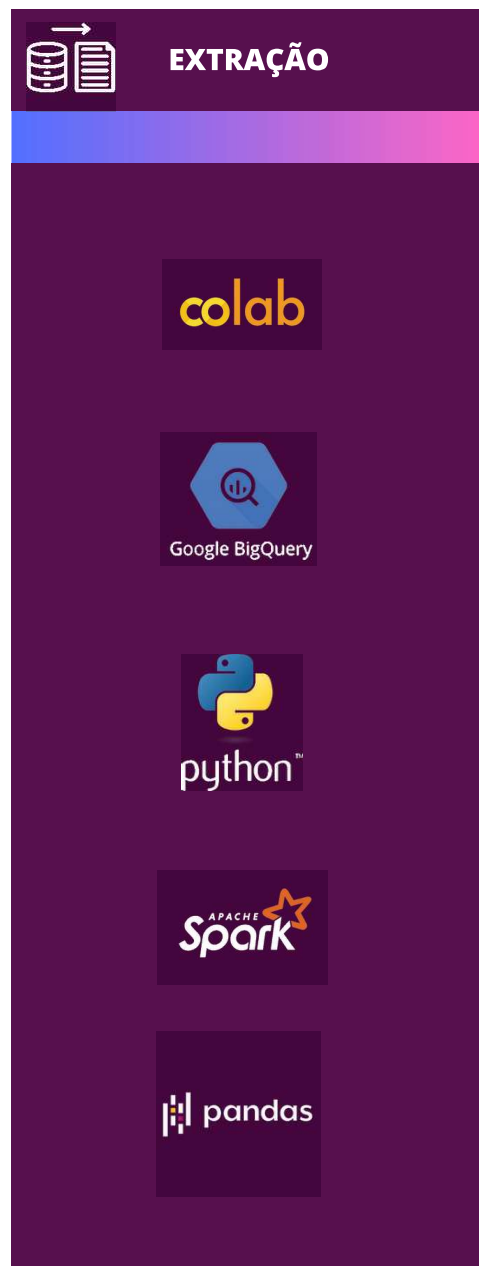
**Distribuição de incidência solar no Brasil**

# METODOLOGIA

## Descoberta de Conhecimento em Banco de Dados (KDD)

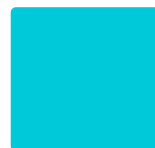


Observação: A técnica de "Mineração de Dados" foi ajustada para integração no dashboard





# **DASHBOARDS**



**ANUÁRIO ESTATÍSTICO DE  
ENERGIA ELÉTRICA**



**FATOR DE CAPACIDADE DE  
GERAÇÃO EÓLICA E SOLAR**



**RELAÇÃO DE EMPREENDIMENTOS  
DE GERAÇÃO DISTRIBUÍDA**



**ATLAS SOLAR IRRADIAÇÃO  
GLOBAL HORIZONTAL**



**PONTOS DE RECARGA PARA  
VEÍCULOS ELÉTRICOS NO MUNDO**



**HISTÓRICO DO LICENCIAMENTO DE  
VEÍCULOS ELÉTRICOS NO BRASIL**



# SWOT

## Strenghts, Weaknesses, Opportunities e Threats

### Forças

**Recursos Hidrelétricos Abundantes**  
**Energia Renovável Diversificada**  
**Incentivos Governamentais**

### Fraquezas

**Dependência de Hidrelétricas**  
**Vulnerabilidade Climática**  
**Desafios Socioambientais**

### Oportunidades

**Diversificação de Fontes**  
**Tecnologias Inovadoras**  
**Eficiência Energética**  
**Microgeração e Autossuficiência**  
**Projetos Socioambientais Sustentáveis**

### Ameaças

**Instabilidade Climática**  
**Riscos Socioambientais**  
**Dependência de Importações**  
**Falta de Investimento**

# SUGESTÃO

Gas Efeito estufa



Eletropontos



Geração distribuída

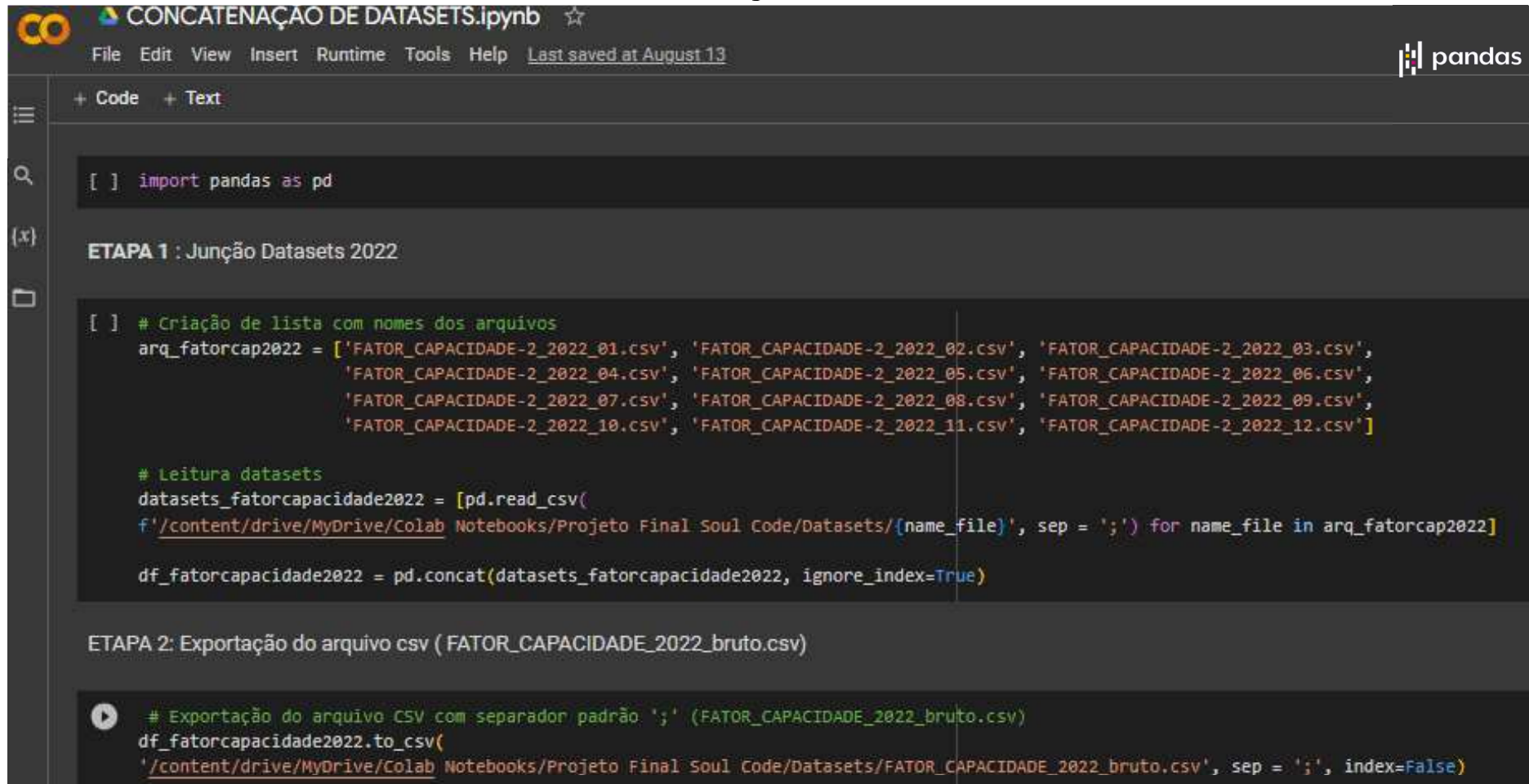


Transição energética



# CODIFICAÇÃO

## Concatenação Datasets



```
CONCATENAÇÃO DE DATASETS.ipynb ☆
File Edit View Insert Runtime Tools Help Last saved at August 13 pandas

+ Code + Text

[ ] import pandas as pd

ETAPA 1 : Junção Datasets 2022

[ ] # Criação de lista com nomes dos arquivos
    arq_fatorcap2022 = ['FATOR_CAPACIDADE-2_2022_01.csv', 'FATOR_CAPACIDADE-2_2022_02.csv', 'FATOR_CAPACIDADE-2_2022_03.csv',
                        'FATOR_CAPACIDADE-2_2022_04.csv', 'FATOR_CAPACIDADE-2_2022_05.csv', 'FATOR_CAPACIDADE-2_2022_06.csv',
                        'FATOR_CAPACIDADE-2_2022_07.csv', 'FATOR_CAPACIDADE-2_2022_08.csv', 'FATOR_CAPACIDADE-2_2022_09.csv',
                        'FATOR_CAPACIDADE-2_2022_10.csv', 'FATOR_CAPACIDADE-2_2022_11.csv', 'FATOR_CAPACIDADE-2_2022_12.csv']

    # Leitura datasets
    datasets_fatorcapacidade2022 = [pd.read_csv(
        f'/content/drive/MyDrive/Colab Notebooks/Projeto Final Soul Code/Datasets/{name_file}', sep = ';') for name_file in arq_fatorcap2022]


    df_fatorcapacidade2022 = pd.concat(datasets_fatorcapacidade2022, ignore_index=True)

ETAPA 2: Exportação do arquivo csv ( FATOR_CAPACIDADE_2022_bruto.csv)

# Exportação do arquivo CSV com separador padrão ';' (FATOR_CAPACIDADE_2022_bruto.csv)
df_fatorcapacidade2022.to_csv(
    '/content/drive/MyDrive/Colab Notebooks/Projeto Final Soul Code/Datasets/FATOR_CAPACIDADE_2022_bruto.csv', sep = ';', index=False)
```

# CODIFICAÇÃO


## Concatenação Datasets

 CONCATENAÇÃO\_DE\_DATASETS.ipynb ☆  
File Edit View Insert Runtime Tools Help Last saved at 1:11 PM

Comment Share

+ Code + Text Connect

ETAPA 5: Junção Datasets 2015 a 2023



```
[ ] # Criação de lista com nomes dos arquivos
arq_fatorcapacidade = ['FATOR_CAPACIDADE_2015_bruto.csv', 'FATOR_CAPACIDADE_2016_bruto.csv', 'FATOR_CAPACIDADE_2017_bruto.csv',
                        'FATOR_CAPACIDADE_2018_bruto.csv', 'FATOR_CAPACIDADE_2019_bruto.csv', 'FATOR_CAPACIDADE_2020_bruto.csv',
                        'FATOR_CAPACIDADE_2021_bruto.csv', 'FATOR_CAPACIDADE_2022_bruto.csv', 'FATOR_CAPACIDADE_2023_bruto.csv']

# Leitura datasets
datasets_fatorcapacidade = [pd.read_csv(
    f'/content/drive/MyDrive/Colab Notebooks/Projeto Final Soul Code/Datasets/{name_file}',
    sep = ';') for name_file in arq_fatorcapacidade]

# União de todos os datasets, ignora os 'indexes' originais para evitar criação de duplicatas
df_fatorcapacidade = pd.concat(datasets_fatorcapacidade, ignore_index=True)
```

ETAPA 6: UPLOAD arquivo FATOR\_CAPACIDADE\_2021a2923.csv para Google Storage/Bucket

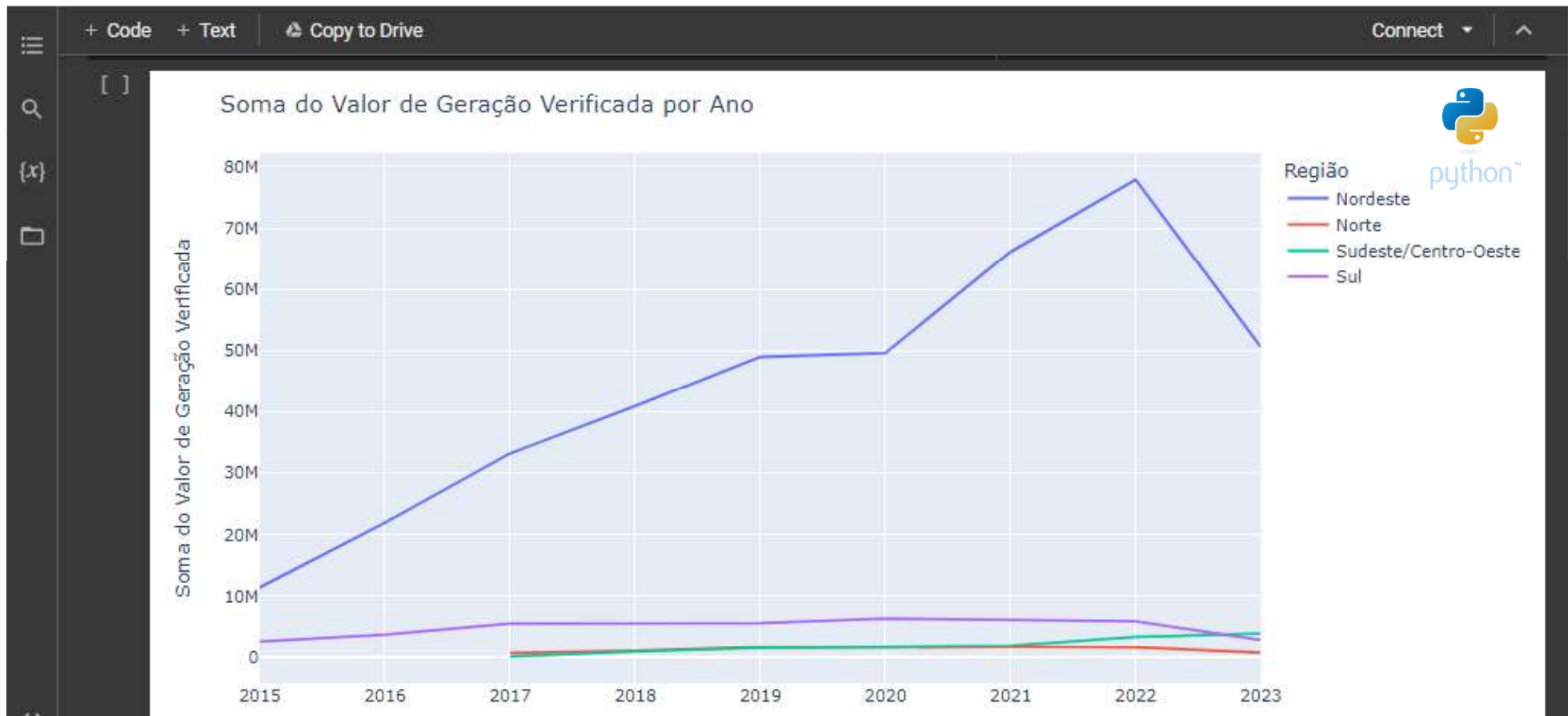
Filtrar apenas pelo prefixo do nome Filtro Filtrar objetos e pastas

Mostrar dados excluídos

<input type="checkbox"/>	Nome	Tamanho	Tipo	Criado ?	Classe de armazenamento	Última modificação	
<input type="checkbox"/>	Bruto_Total_Usinas.csv	574 KB	text/csv	1 de ago. de 2023 12:28:12	Standard	1 de ago. de 2023 12:28:12	↓ ⋮
<input type="checkbox"/>	FATOR_CAPACIDADE_2015a2023_brut.	1,7 GB	text/csv	4 de ago. de 2023 16:19:18	Standard	4 de ago. de 2023 16:19:18	↓ ⋮
<input type="checkbox"/>	Fator capacidade Eólica e Solar/	—	Pasta	—	—	—	⋮
<input type="checkbox"/>	Veiculos eletricos brasil anfaeva/	—	Pasta	—	—	—	⋮
<input type="checkbox"/>	geracao_distribuida_bruto.csv	571,5 MB	text/csv	1 de ago. de 2023 12:27:02	Standard	1 de ago. de 2023 12:27:02	↓ ⋮

# CODIFICAÇÃO

## Gráfico Python



# CODIFICAÇÃO

 pandas

ETL

```
1 # Criação de coluna com valor Ano
2 df['anoCadastral'] = df.DthAtualizaCadastralEmpreend.dt.year
```

**Extração do valor Ano em uma nova coluna do dataframe**

```
1 # filtro NAN para UF
2 filtroUF = df.loc[df['SigUF'] == 'NAN']

1 # Laço de repetição para substituição dos valores NAN em SigUF
2 for n in filtroUF['CodEmpreendimento']:
3     for m in listaUF:
4         if str(n).startswith(f'GD.{m}'):
5             df['SigUF'] = df['SigUF'].replace('NAN', f'{m}')
6         else:
7             pass
```

**Laço de repetição para extração da UF presente em CodEmpreendimento e preenchimento da coluna SigUF com valores NAN**

```
1 # Verificando valores únicos para SigUF, retirando NAN
2 df.SigUF.unique()
```

```
array(['SC', 'SP', 'RJ', 'ES', 'MG', 'BA', 'SE', 'PE', 'AL', 'PB', 'RN',
       'CE', 'PI', 'MA', 'PA', 'AP', 'AM', 'RR', 'AC', 'DF', 'GO', 'RO',
       'TO', 'MT', 'MS', 'PR', 'RS'], dtype=object)
```



# CODIFICAÇÃO

## MONGODB



### Verificação

```
[ ] 1 # Verificando valores no MongoDB
    2 doc_count_csv = collection_csv.count_documents({})
    3 doc_count_excel = collection_excel.count_documents({})
    4
    5 print( f'\n-Csv: {doc_count_csv} \n-EXCEL: {doc_count_excel}')
```

```
-CSV: 93014
-EXCEL: 660
```

## Colab para conexão MongoDB

### backup

LOGICAL DATA SIZE: 30.86MB STORAGE SIZE: 10.45MB INDEX SIZE: 2.62MB TOTAL COLLECTIONS: 2

CREATE

Collection Name	Documents	Logical Data Size	Avg Document Size	Storage Size	Indexes	Index Size	Avg Index Size
csv	93014	30.78MB	347B	10.41MB	1	2.57MB	2.57MB
excel	660	87.52KB	136B	40KB	1	48KB	48KB

## Console MongoDB

# CODIFICAÇÃO

## Projeto Final

Escola: SoulCode Academy

Curso: Bootcamp Analista de Dados - Martech - AD2

### Equipe 08

Alunos: Adriano Kim, José Aurelio, Marcos Paulo, Paulo Vitorino, Renato e Wesley

Professores: Douglas Ribeiro, Franciane Rodrigues e Jonathas Carneiro

## ▾ Preparação de Ambiente

Instalações e importações das bibliotecas necessárias para o processo de ETL.

```
1 # Instalação Bibliotecas
2 !pip install gcsfs -q

[ ] 1 # Importando Bibliotecas
2 import os
3 import pandas as pd
4 import numpy as np
5
6 from google.colab import data_table
7 from google.cloud import storage
8 from google.colab import drive
9 from google.colab import auth
10 from pydrive.auth import GoogleAuth
11 from pydrive.drive import GoogleDrive
```

## Análise de Dados



# CODIFICAÇÃO



**BIGQUERY**



# Agradecimentos



meio&mensagem

