



**Pontifícia Universidade Católica do Rio de Janeiro – PUC Rio**

**Especialização em Ciência de Dados e Analytics**

**Sprint III: Engenharia de Dados**

**Minimun Viable Product – MVP**

**Dayana de Lima e Silva**

\*\*\*\*\*

**ANTES DE TUDO:** Este MVP tem por base o MVP entregue na Sprint I e disponibilizado em: [ciencia\\_dados/MVP\\_SprintI\\_DayanaLima.ipynb at main · dayanallima/ciencia\\_dados \(github.com\)](https://github.com/dayanallima/ciencia_dados/blob/main/ciencia_dados/MVP_SprintI/DayanaLima.ipynb)

## **1. Definição do Problema**

### **Objetivo**

A ANP divulgou os dados de produção de óleo dos campos de terra, mar e pré-sal dos meses de janeiro e fevereiro de 2023. O objetivo desse MVP é traçar um panorama do cenário atual da produção no Brasil.

### **Modelo de Aprendizado**

Esse MVP é um modelo de aprendizado supervisionado, pois é construído a partir dos dados de entrada que serão apresentados para um algoritmo na forma de pares ordenados (entrada – saída desejada). Esses dados são rotulados, pois já é de conhecimento, a saída esperada para cada entrada de dados. Nesse caso, o aprendizado (ou treinamento) consiste em apresentar para o algoritmo um número suficiente de exemplos de entradas e saídas desejadas (já rotuladas previamente). Assim, o objetivo do algoritmo é aprender uma regra geral que mapeie as entradas nas saídas corretamente, o que consiste no modelo final.

### **Premissas**

Apenas a produção de óleo é de interesse para apresentação. Os arquivos bases serão usados no formato .xlsx conforme bases originais para evitar perdas para outras extensões.

### **Condições e Restrições**

Dados apenas de 2023 retirados do domínio público: <https://www.gov.br/anp/pt-br/centrais-de-conteudo/dados-estatisticos>

Caminho do material:

1. Produção de Petróleo e Gás Natural em Dados Estatísticos.
2. Produção por poços: Panorama anual da produção nacional de petróleo e gás natural por poço, divididos em mar e terra.
3. Produção por poço em 2023.

Para facilitar a análise, foi criada em cada arquivo uma coluna "Exploração" identificando o ambiente: terra, mar ou pré-sal. Isso foi feito ainda no Excel tendo em vista a dificuldade de fazer esse agrupamento via código pois o meio que poderia ser usado, seria o nome do poço o que implicaria numa atribuição extensa de cada poço a seu meio de exploração.

### Atributos

Para a análise, o foco será na produção de óleo e como esta se distribui nos estados, campos, bacias etc.

- Estado: Unidade federativa do país. O Brasil tem 26 estados mais o distrito federal.
- Exploração: Meio onde o petróleo foi explorado, podendo ser em terra, mar (não pré sal) e no pré sal. Estruturalmente, a exploração do mar com pré sal se diferencia da exploração no mar "comum" por apresentar uma camada de sal antes da área de produção.
- Bacia: Bacia ou bacia sedimentar é o espaço onde há a formação de petróleo e alguns tipos de fósseis.
- Nome\_ANP: Código de descrição do poço segundo nomenclatura da Agência Nacional de Petróleo - ANP.
- Campo: Um é uma área, na bacia, onde estão instalados equipamentos para extração de combustível fóssil.
- Operador: Operador ou empresa operadora é a companhia responsável por operacionalizar as atividades de um campo.
- Período: Intervalo de tempo que foi feito o registro da produção. Nesse *dataset*, janeiro e fevereiro de 2023.
- Oleo\_bpd: Produção de óleo em barris por dia (bpd). 1m<sup>3</sup>/d equivale a 6,2898 bpd.
- Grau\_API: O Grau API é uma escala que mede a densidade dos líquidos derivados do petróleo. A escala API, medida em graus, varia inversamente à densidade relativa, isto é, quanto maior a densidade relativa, menor o grau API. O grau API é maior quando o petróleo é mais leve. Se o grau API for superior a 10, o petróleo flutua na água. Se for menor que 10, o petróleo é mais denso que a água e afunda. Petróleos com grau API maior que 31 são considerados leves; entre 22 e 31 graus API, são médios; abaixo de 22 graus API, são pesados; com grau API igual ou inferior a 10, são petróleos extrapesados. Quanto maior o

grau API, maior o valor do petróleo no mercado, pois, sendo mais leve, ele fornece derivados de maior valor agregado e tem um processo de refino mais barato.

## 2. Plataforma Databricks

Para a análise dos dados e processamento em nuvem foi utilizado o ambiente da Plataforma Databricks por meio da versão Databricks Community Edition.

## 3. Etapas do Trabalho

### Primeiros Passos

Primeiramente o notebook, usado para a Sprint I, foi importado do *github* (Figura 1) e feito os devidos ajustes: retirada das etapas que não eram de interesse da Sprint III e configurações do código Python para que o mesmo fosse executado na plataforma.

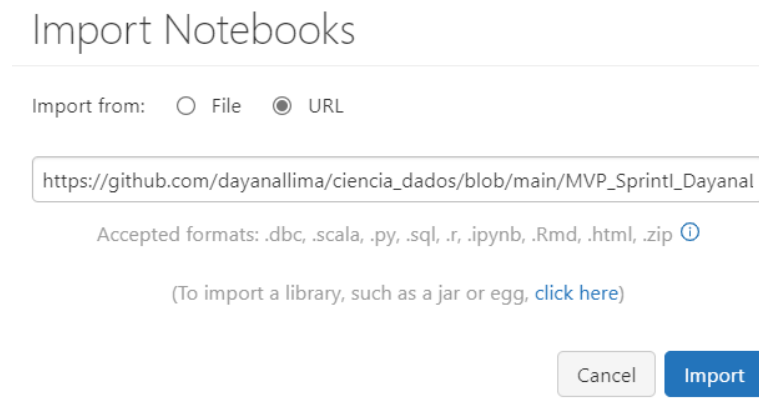


Figura 1- Importação do Notebook base a partir do Github

### Criação do Cluster

Para executar as tarefas dentro da Plataforma, um Cluster que é um conjunto de recursos computacionais foi criado, conforme Figura 2.

## Attach to a compute resource



Your notebook must be attached to a compute resource to execute commands. Create new compute to run your command.

Name

Cluster\_MVPIII\_DayanaLima

Runtime

Runtime: 12.2 LTS (Scala 2.12, Spark 3.3.2)



Advanced Configuration

Cancel

Create, Attach, & Run

Figura 2- Criação do Cluster na Plataforma Databricks

## ETL (Extract, Transform, Load) - Etapa de Extração

A base de dados, inicialmente descrita nesse documento, foi extraída via notebook após os arquivos serem inseridos no *Github*, como mostra a Figura 3.

```
2 minutes ago (20s) 12 Python
```

```
## Carga da Base de Dados
# O dataset está em domínio público no GitHub, uma plataforma de hospedagem de código-fonte.

## Caminho das Bases de Dados

# Base 1 - Dados de produção de Terra em Jan/23
base_01 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_01_producao_Terra.xlsx?raw=true"

# Base 2 - Dados de produção de Terra em Fev/23
base_02 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_02_producao_Terra.xlsx?raw=true"

# Base 3 - Dados de produção do Mar em Jan/23
base_03 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_01_producao_Mar.xlsx?raw=true"

# Base 4 - Dados de produção do Mar em Fev/23
base_04 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_02_producao_Mar.xlsx?raw=true"

# Base 5 - Dados de produção do Pré sal em Jan/23
base_05 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_01_producao_Presal.xlsx?raw=true"

# Base 6 - Dados de produção do Pré sal em Fev/23
base_06 = "https://github.com/dayanallima/ciencia_dados/blob/main/2023_01_producao_Presal.xlsx?raw=true"

## Leitura dos Arquivos

dataset_01 = pd.read_excel(base_01)
dataset_02 = pd.read_excel(base_02)
dataset_03 = pd.read_excel(base_03)
dataset_04 = pd.read_excel(base_04)
dataset_05 = pd.read_excel(base_05)
dataset_06 = pd.read_excel(base_06)
```

Figura 3- EXTRAÇÃO dos Dados

A Figura 4 mostra a visualização dos dados após concatenação das bases:

2 minutos ago (<1s)14

# Cria um dataset unido as 6 bases importadas  
dataset = pd.concat([dataset\_01, dataset\_02, dataset\_03, dataset\_04, dataset\_05, dataset\_06])

1 minute ago (2s)15Python

# Exibe visualmente o conjunto de dados  
display(dataset)

	Estado	Exploração	Bacia	Nome_ANP	Nome_Operador	Campo	Ope
1	Espírito Santo	Terra	Espírito Santo	1-VITA-2A-ES	1PG01AES	TUCANO	Vipetro
2	Espírito Santo	Terra	Espírito Santo	7-CNC-21-ES	7CNC21ES	CANCÃ	Petrobras
3	Rio Grande do Nor...	Terra	Potiguar	7-FP-211-RN	7FP 0211 RN	FAZENDA POCINHO	Petrobras
4	Bahia	Terra	Recôncavo	1-ALV-11-BA	1-ALV-198-A1-BA	CABURÉ LESTE	Alvopetro
5	Bahia	Terra	Recôncavo	8-DJ-754-BA	8DJ 0754 BA	DOM JOÃO	3R Canc
6	Rio Grande do Nor...	Terra	Potiguar	3-STAR-15-RN	3PBC2RN	SABIÁ BICO-DE-OSSO	Potiguar
7	Rio Grande do Nor...	Terra	Potiguar	7-SVM-8-RN	7SVM 0008 RN	SERRA VERMELHA	Petrobras
8	Rio Grande do Nor...	Terra	Potiguar	7-ET-1598D-RN	7ET1598DRN	ESTREITO	Petrobras
9	Espírito Santo	Terra	Espírito Santo	1-VITA-4-ES	1CQ1ES	TUCANO	Vipetro
10	Espírito Santo	Terra	Espírito Santo	7-CNC-19-ES	7CNC19ES	CANCÃ	Petrobras
11	Rio Grande do Nor...	Terra	Potiguar	7-FP-207-RN	7FP 0207 RN	FAZENDA POCINHO	Petrobras
12	Espírito Santo	Terra	Espírito Santo	7-LPN-31-ES	7LPN 0031 ES	LAGOA PARDA NORTE	Capixab
13	Bahia	Terra	Recôncavo	7-DJ-815H-BA	7DJ815HBA	DOM JOÃO	3R Canc
14	Rio Grande do Nor...	Terra	Potiguar	3-STAR-28-RN	3PBC3RN	SABIÁ BICO-DE-OSSO	Potiguar

4,516+ rows | Truncated data due to byte limit | 2.10 seconds runtimeRefreshed 1 minute ago

Figura 4 - Visualização, no Databricks, da base de dados

Como a base de dados da Figura 4 tinha informações além das necessárias para o objetivo deste MVP, uma nova base foi definida, conforme a Figura 5.

4 minutos ago (<1s)16Python

## Cria uma base para análise apenas com as colunas desejadas  
# Arquivo para Análise  
dataset\_analise = dataset[['Estado', 'Exploração', 'Bacia', 'Nome\_ANP',  
| | | | | 'Campo', 'Operador', 'Período', 'Óleo\_bpd', 'Grau\_API']]

4 minutos ago (1s)17Python

# Exibe visualmente o conjunto de dados  
display(dataset\_analise)

	Estado	Exploração	Bacia	Nome_ANP	Campo	Operador	Período	1.2 Óleo_bpd	1.2 Grau_API
1	Espírito Santo	Terra	Espírito Santo	1-VITA-2A-ES	TUCANO	Vipetro	2023/01	5.9461	16
2	Espírito Santo	Terra	Espírito Santo	7-CNC-21-ES	CANCÃ	Petrobras	2023/01	69.2869	19.6
3	Rio Grande do Nor...	Terra	Potiguar	7-FP-211-RN	FAZENDA POCINHO	Petrobras	2023/01	0.6247	23.2
4	Bahia	Terra	Recôncavo	1-ALV-11-BA	CABURÉ LESTE	Alvopetro	2023/01	0	null
5	Bahia	Terra	Recôncavo	8-DJ-754-BA	DOM JOÃO	3R Candeias S.A.	2023/01	4.9405	36.5
6	Rio Grande do Nor...	Terra	Potiguar	3-STAR-15-RN	SABIÁ BICO-DE-OSSO	Potiguar E&P S.A.	2023/01	22.4175	25.43
7	Rio Grande do Nor...	Terra	Potiguar	7-SVM-8-RN	SERRA VERMELHA	Petrobras	2023/01	1.2726	23.2
8	Rio Grande do Nor...	Terra	Potiguar	7-ET-1598D-RN	ESTREITO	Petrobras	2023/01	0.7836	23.2
9	Espírito Santo	Terra	Espírito Santo	1-VITA-4-ES	TUCANO	Vipetro	2023/01	3.5172	16
10	Espírito Santo	Terra	Espírito Santo	7-CNC-19-ES	CANCÃ	Petrobras	2023/01	41.3713	19.6
11	Rio Grande do Nor...	Terra	Potiguar	7-FP-207-RN	FAZENDA POCINHO	Petrobras	2023/01	4.7419	23.2
12	Espírito Santo	Terra	Espírito Santo	7-LPN-31-ES	LAGOA PARDA NORTE	Capixaba Energia	2023/01	0	null
13	Bahia	Terra	Recôncavo	7-DJ-815H-BA	DOM JOÃO	3R Candeias S.A.	2023/01	42.1805	36.5
14	Rio Grande do Nor...	Terra	Potiguar	3-STAR-28-RN	SABIÁ BICO-DE-OSSO	Potiguar E&P S.A.	2023/01	4.3521	25.43
15	Rio Grande do Nor...	Terra	Potiguar	3-BRSA-873-RN	SIBITE	Potiguar E&P S.A.	2023/01	14.9142	null

10,000+ rows | Truncated data due to row limit | 1.22 seconds runtimeRefreshed 4 minutos ago

Figura 5 - Base de dados que será usada para a análise

Foi feito download desta base (Figura 6) para posterior criação de tabela (Figura 7).

Just now (7s)

# Exibe visualmente o conjunto de dados

display(dataset\_analise)

(1) Spark Jobs

Table

Download first 10000 rows

Download all rows (up to 5GB compressed)

Add to notebook dashboard

Rio Grande do Nor...

Bahia

Bahia

Rio Grande do Nor...

Rio Grande do Nor...

Rio Grande do Nor...

Espirito Santo

Espirito Santo

Rio Grande do Nor...

Espirito Santo

Bahia

Rio Grande do Nor...

Rio Grande do Nor...

Rio Grande do Nor...

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Terra

Recôncavo

Recôncavo

Recôncavo

Potiguar

Potiguar

Potiguar

Potiguar

Espirito Santo

Espirito Santo

Potiguar

Recôncavo

Potiguar

Potiguar

Potiguar

1-VITA-2A-ES

7-CNC-21-ES

7-FP-211-RN

1-ALV-11-BA

8-DJ-754-BA

3-STAR-15-RN

7-SVM-8-RN

7-ET-1598D-RN

1-VITA-4-ES

7-CNC-19-ES

7-FP-207-RN

7-LPN-31-ES

7-DJ-815H-BA

3-STAR-28-RN

3-BRSA-873-RN

TUCANO

CANCÃ

FAZENDA POCINHO

CABURÉ LESTE

DOM JOÃO

SABIÁ BICO-DE-OSSO

SERRA VERMELHA

ESTREITO

TUCANO

CANCÃ

FAZENDA POCINHO

LAGOA PARDA NORTE

DOM JOÃO

SABIÁ BICO-DE-OSSO

SIBITE

Vipetro

Petrobras

Petrobras

Alvopetro

3R Candeias S.A.

Potiguar E&P S.A.

Petrobras

Petrobras

Vipetro

Petrobras

Petrobras

Capixaba Energia

3R Candeias S.A.

Potiguar E&P S.A.

Potiguar E&P S.A.

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

2023/01

1.2 Óleo\_bpd

1.2 Grau\_API

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

1.2 Óleo\_bpd

5.9461

69.2869

0.6247

0

4.9405

22.4175

1.2726

0.7836

3.5172

41.3713

4.7419

0

42.1805

4.3521

14.9142

16

19.6

23.2

null

36.5

25.43

23.2

23.2

16

19.6

23.2

null

36.5

25.43

null

10,000+ rows

Truncated data due to row limit

6.55 seconds runtime

Refreshed now

Figura 6 - Download da Base de Dados

## Create New Table

Data source ?

Upload File S3 Other Data Sources

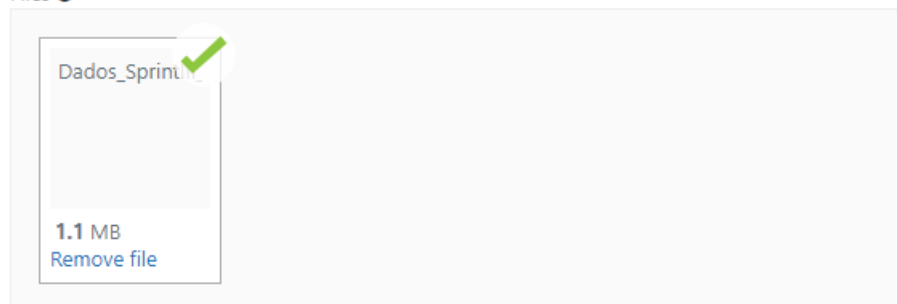
DBFS Target Directory ?

/FileStore/tables/ SprintIII

Select

Files uploaded to DBFS are accessible by everyone who has access to this workspace. [Learn more](#)

Files ?



✓ File uploaded to /FileStore/tables/SprintIII/Dados\_SprintIII\_DayanaLima-2.csv

Create Table with UI

Create Table in Notebook ?

Figura 7 – Criando uma tabela a partir dos dados do notebook

## ETL (Extract, Transform, Load) - Etapas de Transformação e Carregamento

### Ambiente Apache Spark

Conforme documentação da Plataforma, o Apache Spark está no centro do Databricks e é a tecnologia que alimenta clusters compute e SQL warehouse.

No notebook foi inserido um código, para que o arquivo CSV, contendo a base de dados, fosse lido e no DataFrame para que pudesse ser manipulado e processado posteriormente usando as funcionalidades do Spark. Para o contexto proposto, tais funcionalidades estão relacionadas à interface da Linguagem SQL, de acordo com a Figura 8.

```

# Leitura do arquivo CSV no ambiente do Apache Spark, utilizando a linguagem de programação Python.
# ".format("CSV")": especifica o formato do arquivo que será lido, que é CSV neste caso.
# ".option("header", "true")": define a opção de que a primeira linha do arquivo CSV contém o cabeçalho, especificando que a opção "header" é verdadeira.
# ".option("inferSchema", "true")": define a opção de inferir automaticamente o esquema dos dados, ou seja, o Spark tentará deduzir os tipos de dados para cada coluna com base nos valores presentes no arquivo CSV.
# ".load("dbfs:/FileStore/tables/SprintIII/Dados_SprintIII_DayanaLima-1.csv")": carrega o arquivo CSV especificado com o caminho "dbfs:/FileStore/tables/SprintIII/export.csv" e preenche o DataFrame "df" com os dados.

df = spark.read \
    .format("CSV") \
    .option("header", "true") \
    .option("inferSchema", "true") \
    .load("dbfs:/FileStore/tables/SprintIII/Dados_SprintIII_DayanaLima-2.csv")
  
```

(2) Spark Jobs

df: pyspark.sql.dataframe.DataFrame = [Estado: string, Exploração: string ... 7 more fields]

Figura 8 - Ambiente Apache para uso da Linguagem SQL

A partir do DataFrame foi criada uma tabela temporária na memória do Spark para ser consultada usando a linguagem SQL. E por causa da necessidade de usar SQL, as colunas com títulos grifados com “ç” e acentos ortográficos foram renomeadas. Esse processo é mostrado na Figura 9.

```

df = df.withColumnRenamed("Exploração", "Exploracao") \
    .withColumnRenamed("Óleo_bpd", "oleo_bpd")

# Criando uma tabela temporária para ser consultada a partir da linguagem SQL
df.createOrReplaceTempView("petro_tbl")
  
```

```

%sql
SELECT * FROM petro_tbl -- Visualizando os dados da tabela
  
```

(1) Spark Jobs

\_sqldf: pyspark.sql.dataframe.DataFrame = [Estado: string, Exploracao: string ... 7 more fields]

	Estado	Exploracao	Bacia	Nome_ANP	Campo	Operador	Periodo	oleo_bpd	Grau_API
1	Espirito Santo	Terra	Espirito Santo	1-VITA-2A-ES	TUCANO	Vipetro	2023-01-01	5.9461	16.0
2	Espirito Santo	Terra	Espirito Santo	7-CNC-21-ES	CANCÃ	Petrobras	2023-01-01	69.2869	19.6
3	Rio Grande do Nor...	Terra	Potiguar	7-FP-211-RN	FAZENDA POCINHO	Petrobras	2023-01-01	0.6247	23.2
4	Bahia	Terra	Recôncavo	1-ALV-11-BA	CABURÉ LESTE	Alvopetro	2023-01-01	0	null
5	Bahia	Terra	Recôncavo	8-DJ-754-BA	DOM JOÃO	3R Candelas S.A.	2023-01-01	4.9405	36.5
6	Rio Grande do Nor...	Terra	Potiguar	3-STAR-15-RN	SABIA BICO-DE-OSSO	Potiguar E&P S.A.	2023-01-01	22.4175	25.43
7	Rio Grande do Nor...	Terra	Potiguar	7-SVM-8-RN	SERRA VERMELHA	Petrobras	2023-01-01	1.2726	23.2
8	Rio Grande do Nor...	Terra	Potiguar	7-ET-1598D-RN	ESTREITO	Petrobras	2023-01-01	0.7836	23.2
9	Espirito Santo	Terra	Espirito Santo	1-VITA-4-ES	TUCANO	Vipetro	2023-01-01	3.5172	16.0
10	Espirito Santo	Terra	Espirito Santo	7-CNC-19-ES	CANCÃ	Petrobras	2023-01-01	41.3713	19.6
11	Rio Grande do Nor...	Terra	Potiguar	7-FP-207-RN	FAZENDA POCINHO	Petrobras	2023-01-01	4.7419	23.2
12	Espirito Santo	Terra	Espirito Santo	7-LPN-31-ES	LAGOA PARDA NORTE	Capixaba Energia	2023-01-01	0	null
13	Bahia	Terra	Recôncavo	7-DJ-815H-BA	DOM JOÃO	3R Candelas S.A.	2023-01-01	42.1805	36.5
14	Rio Grande do Nor...	Terra	Potiguar	3-STAR-28-RN	SABIA BICO-DE-OSSO	Potiguar E&P S.A.	2023-01-01	4.3521	25.43
15	Rio Grande do Nor...	Terra	Potiguar	3-BRSA-873-RN	SIBITE	Potiguar E&P S.A.	2023-01-01	14.9142	null

10,000+ rows | Truncated data due to row limit | 1.32 seconds runtime

Refreshed 8 minutes ago

Figura 9 - Criação e Visualização da Tabela Temporária para consulta em SQL

#### 4. Análises e Conclusões

##### Exploração: Qual o ambiente de maior produção de óleo?

As Figuras 10 e 11 mostram as análises feitas considerando os três ambiente de exploração: Terra, Mar (convencional) e Mar (pré-sal).

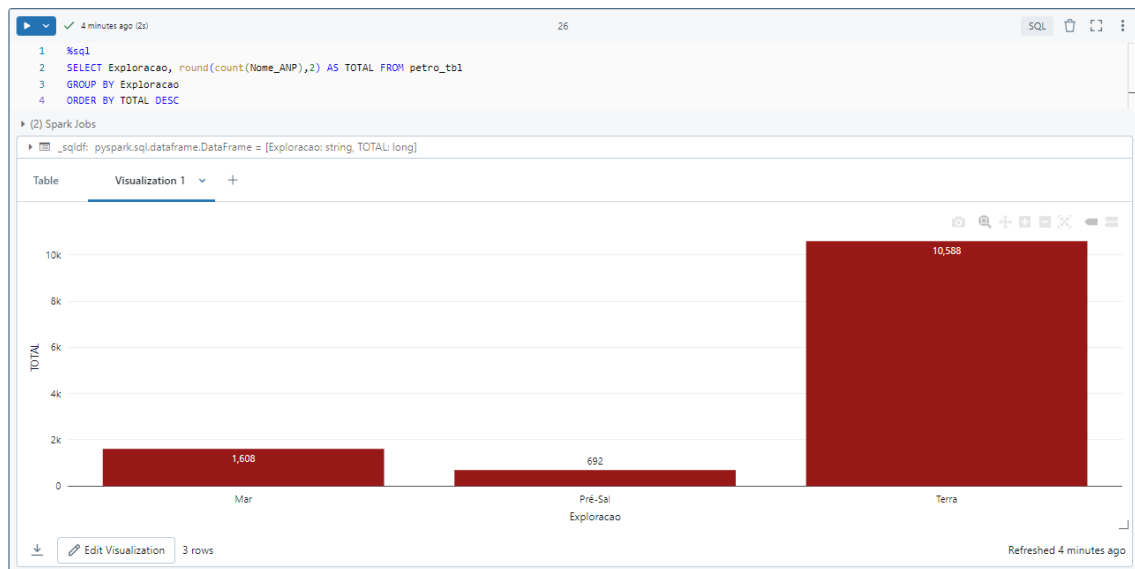


Figura 10 - Ambiente de Exploração X Número de Poços

O gráfico apresenta um quantitativo muito superior de poços *onshore* (terra) comparados com os números offshore (e mar e pré-sal). Dentre os fatores que contribuem para essa diferença estão: geologia, custo e logística. Muitas das principais reservas de petróleo estão localizadas em depósitos de rochas sedimentares em terra. Em relação aos custos, perfurar um poço em terra é quase sempre mais barato do que em mar. As plataformas offshore são mais caras de construir e manter, além de exigir equipamentos e técnicas de perfuração mais avançados e complexos. Com respeito à logística, a extração de petróleo em terra é geralmente mais fácil do que em mar, já que não há necessidade de transportar equipamentos e suprimentos para o alto mar. Além disso, os poços em terra estão localizados perto de infraestruturas existentes, como rodovias e redes de transporte, tornando mais fácil e econômico o transporte de equipamentos, suprimentos e pessoal.



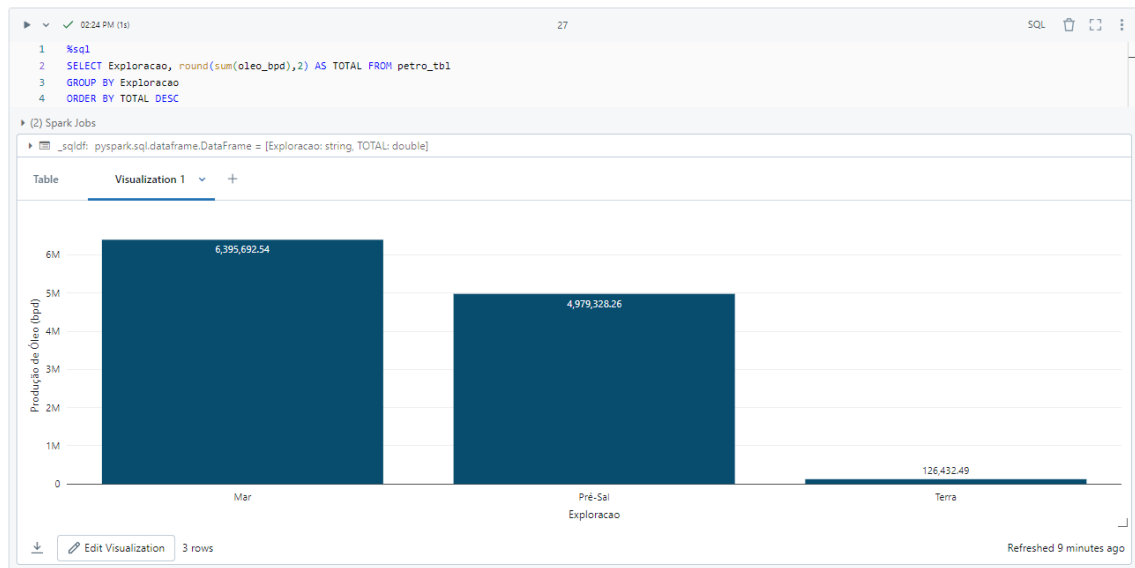


Figura 11 - Ambiente de Exploração x Produção de Óleo (bpd)

Esse gráfico vai de encontro ao anterior. O senso comum espera que mais poços entreguem mais produção. No entanto, os poços offshore, em mar (1608) e no pré-sal (692) produzem mais do que os poços de terra (10588). Para justificar, dois pontos podem ser considerados: reservas e pressão. Os depósitos offshore de petróleo são geralmente maiores do que aqueles em terra. Isso ocorre porque os depósitos em mar são formados por processos geológicos complexos que podem resultar em grandes acumulações de hidrocarbonetos. Além disso, a pressão no reservatório em mar pode ser maior do que em terra. A alta pressão ajuda a forçar o petróleo para fora do reservatório. A pressão também pode ajudar a manter a produção de petróleo em um nível mais alto e constante ao longo do tempo.

### Bacia Sedimentar: Qual a bacia de maior produção de óleo?

As Figuras 12 e 13 mostram as análises feita considerando as bacias sedimentares.

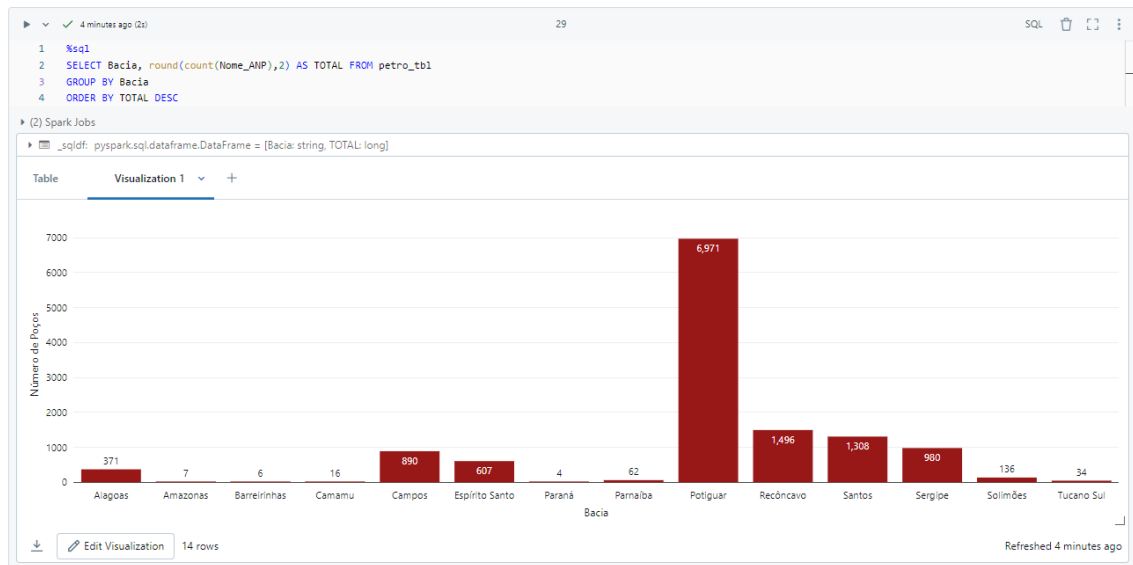


Figura 12 - Bacia Sedimentar X Número de Poços

O gráfico, apesar de não separar por ambiente de exploração, confirma a maior distribuição em terra dos poços, visto que, dentre as 5 maiores, as bacias Potiguar, Recôncavo e Sergipe são majoritariamente terrestres.

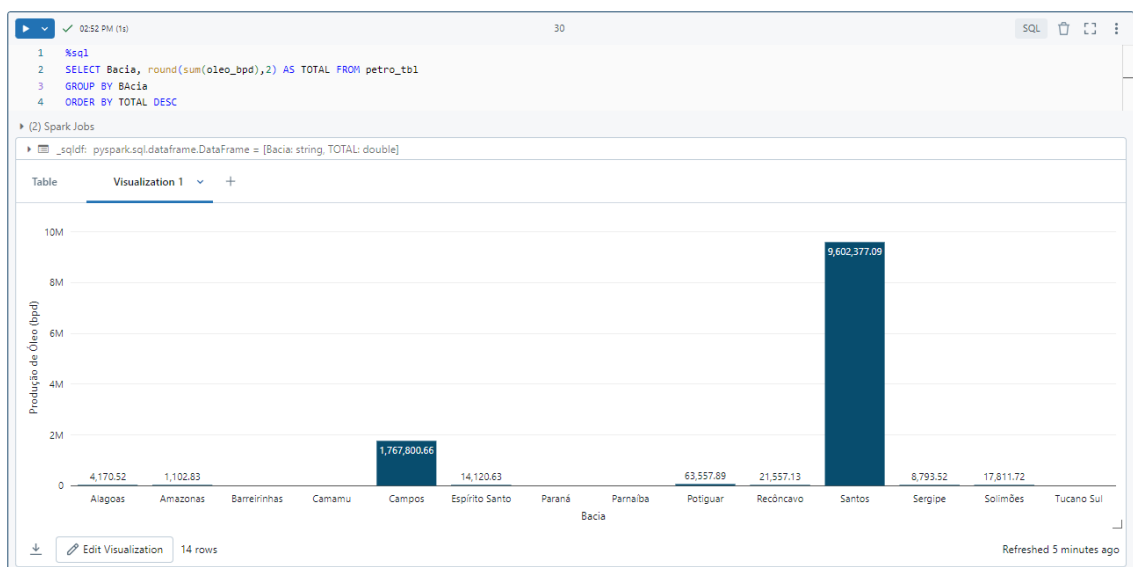


Figura 13 - Bacia Sedimentar X Produção de óleo (bpd)

O gráfico evidencia que as bacias de Campos e Santos que estão entre as 5 com mais poços perfurados são as que entregam maior volume de óleo produzido.

### Estado: Qual o estado brasileiro tem a maior produção de óleo?

As Figuras 14 e 15 mostram as análises feitas considerando os estados brasileiros.

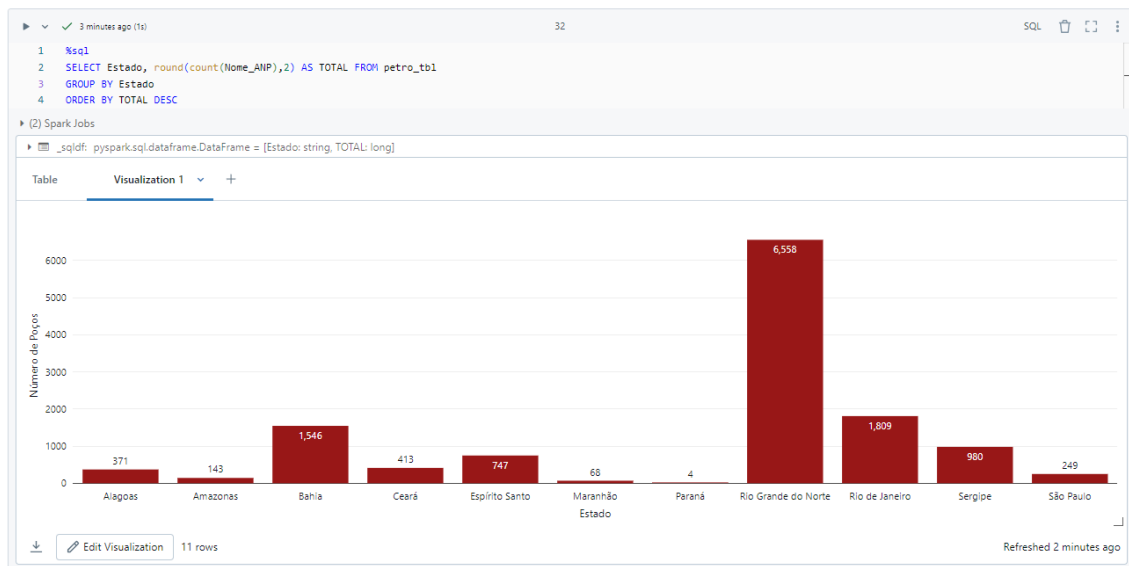


Figura 14 - Estado X Número de Poços

Esse gráfico mostra que o Rio Grande do Norte é o estado com maior número de poços. A bacia sedimentar terrestre da região justifica esse quantitativo e o tempo de exploração na região, também. Do outro lado, com menor número de poços o estado do Paraná.

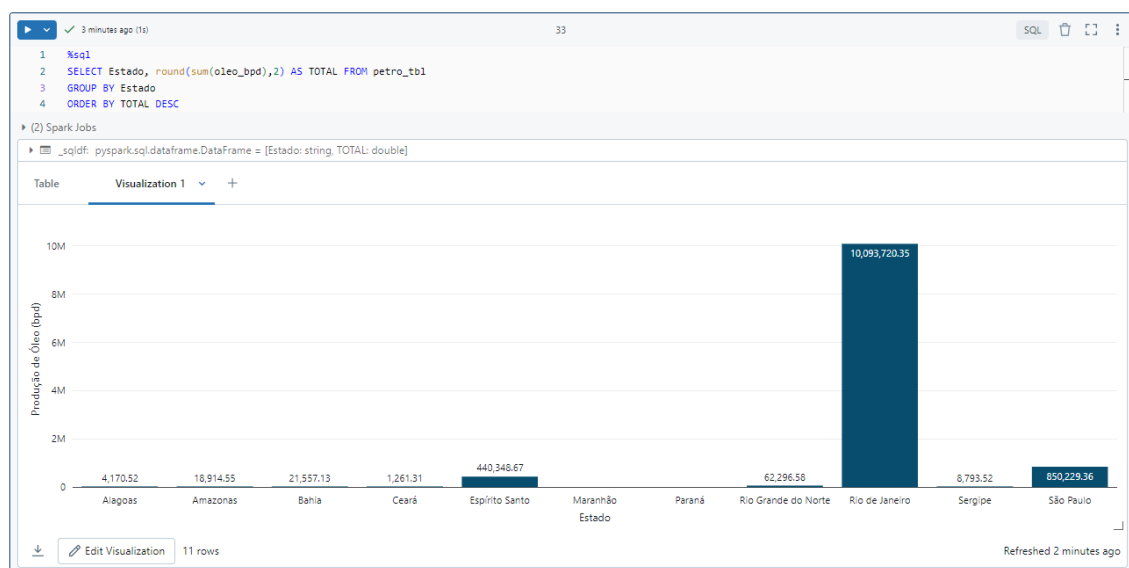


Figura 15 - Estado X Produção de Óleo (bpd)

O gráfico apresenta os estados do Rio de Janeiro e São Paulo em concordância com as bacias de Campos e de Santos, apresentando o maior volume de óleo.

### **Campo de Petróleo: Qual o campo de petróleo brasileiro é o maior produtor de óleo?**

As Figuras 16 e 17 mostram as análises feitas considerando os campos de petróleo brasileiros.

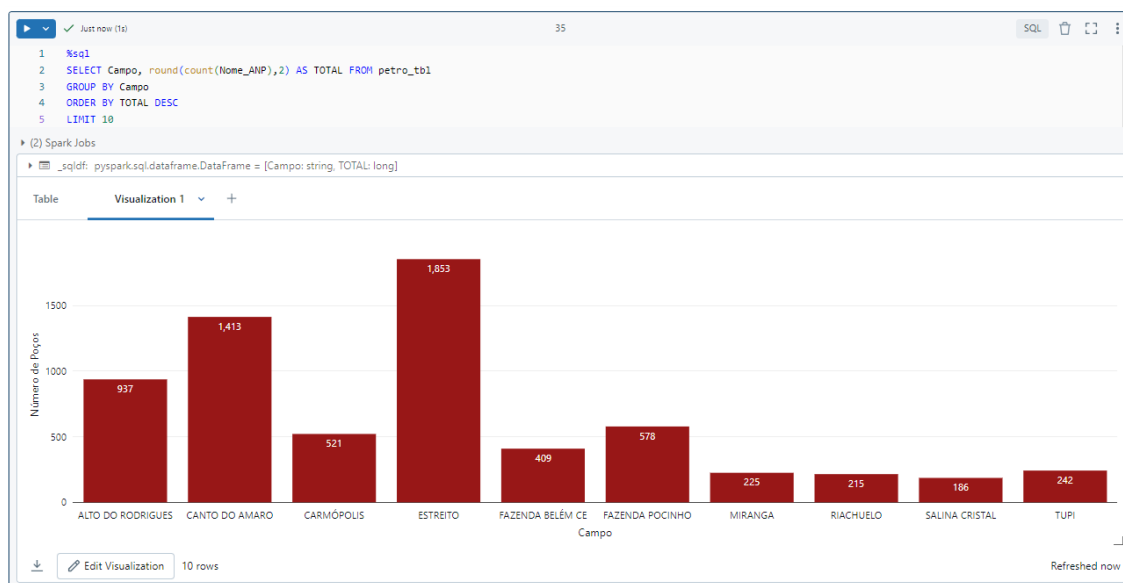


Figura 16 - Campo X Número de Poços

Nesse gráfico por campo, observa-se que dentre os 10 campos com maior número de poços, 6 estão no RN e CE, logo, na Bacia Potiguar.

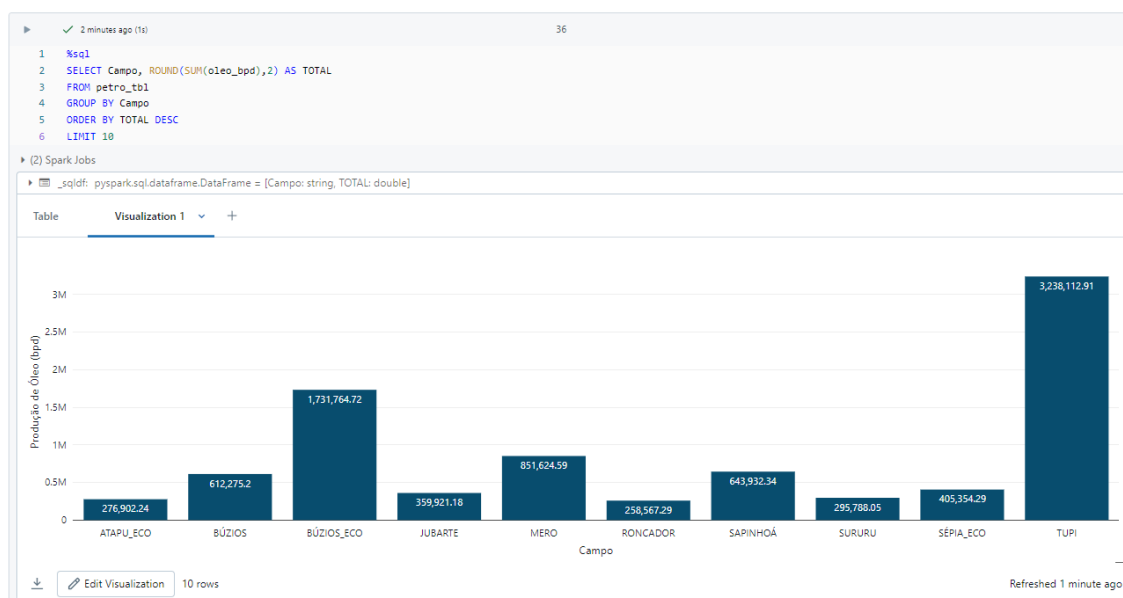


Figura 17 - Campo X Produção de Óleo (bpd)

O gráfico das 10 maiores produções de óleo concentra os campos do sudeste, com 7 registros no Rio de Janeiro.

### Grau API: Qual a avaliação do óleo produzido no Brasil quanto à sua densidade?

A Figura 18 mostra os códigos para o cálculo da Variância, do Desvio Padrão e da Média do Grau API em relação aos 3 Ambientes de Exploração. A Tabela 1 resume esses resultados em melhor visualização.

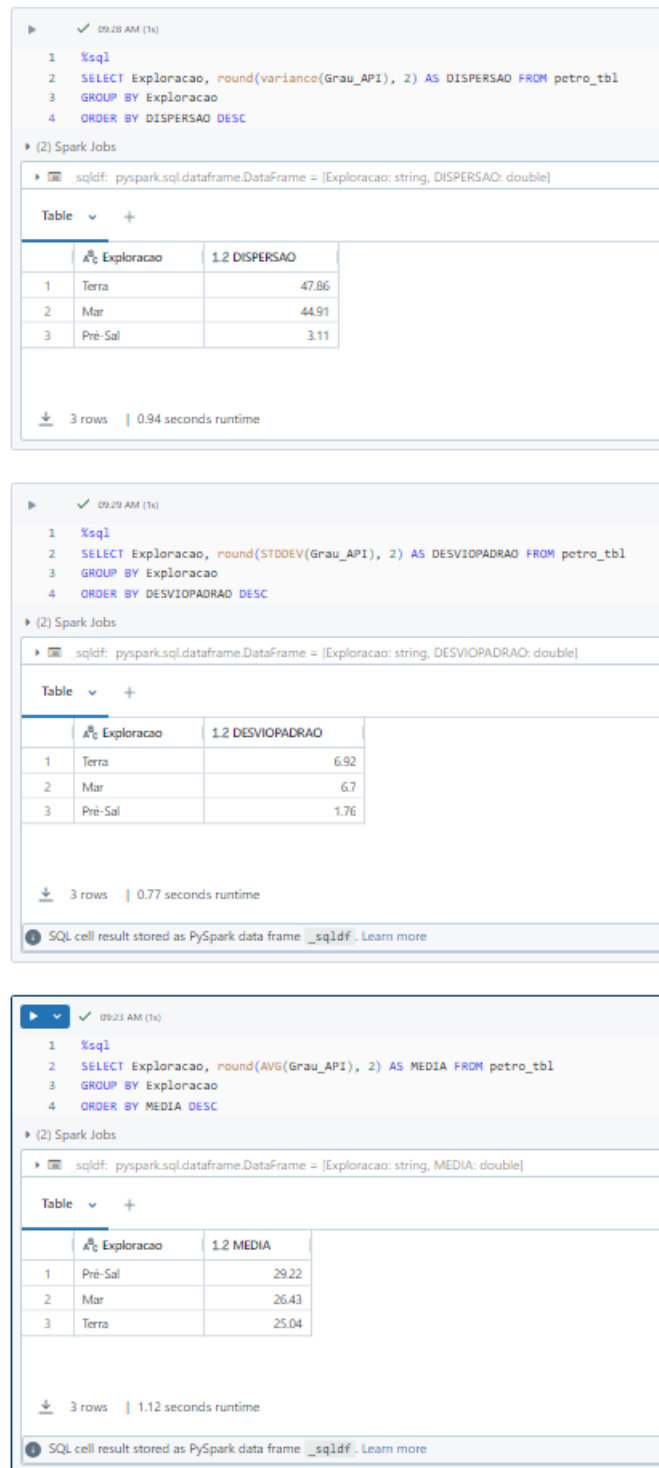


Figura 18 - Análise do Grau API por Ambiente de Exploração

Ambiente de Exploração	Variância	Desvio Padrão	Média
Terra	47.86	6.92	25.04
Mar	44.91	6.7	26.43
Pré-Sal	3.11	1.76	29.22

Tabela 1 - Análise do Grau API por Ambiente de Exploração

Os dados mostram que a maior parte do óleo produzido é de médio a leve pois, com grau API maior que 31 são considerados leves; entre 22 e 31 graus API, são médios. Por definição, quanto maior a densidade

relativa, menor o grau API. Quanto maior o grau API, maior o valor do petróleo no mercado, pois, sendo mais leve, ele fornece derivados de maior valor agregado e tem um processo de refino mais barato. Vale destacar o Grau API do Pré-Sal, nota-se que pouco varia em relação ao médio, bem como, é pequeno também o desvio em relação a esta.

### **Conclusões**

São pontos relevantes para o cenário de produção de óleo, em bpd, no Brasil no primeiro bimestre de 2023:

- A exploração offshore concentra a maior produção de óleo com pouco mais de 17% dos poços do total perfurado no país;
- O Sudeste, é a região brasileira com maior produção de óleo e o Nordeste a região com maior número de poços perfurados;
- Em termos de Grau API, No Brasil, a maioria do petróleo produzido tem um grau API entre 20 e 30, o que significa que é relativamente pesado;
  - Os poços considerados leves estão distribuídos em diferentes estados e regiões, do país. Estes, tem alto valor agregado;
  - O Brasil por sua história e resultados, firma-se no cenário mundial de petróleo com forte potencial.

### **Autoavaliação**

Este trabalho está bastante simplificado, requer bem mais robustez.

Para tanto é necessário dedicação e constância para o aprendizado do manuseio das plataformas e das linguagens necessárias.

Reconheço minhas limitações na área e para o entendimento dos conceitos e aplicações. Mas, ao mesmo tempo, vejo com satisfação a finalização desse MVP pois, ainda assim, em sua simplicidade foi um desafio vencido.