

Relatório Técnico – Simulador Final de Análise de Dados

Programa “Vem pra Neuro, Mulher!”

Télós, Neurotech e Porto Digital

Análise de dados do Programa de Monitoramento de Praias - Petrobrás e IBAMA

Tainah Guerra Santos

31/07/2025

1. Contexto e Objetivo

Os Projetos de Monitoramento de Praias (PMP) são desenvolvidos para o atendimento de condicionantes do licenciamento ambiental federal, conduzido pelo IBAMA, referentes às atividades de exploração e produção de petróleo e gás natural offshore da PETROBRAS.

As regiões monitoradas pelo PMP são a Bacia de Santos, as Bacias de Campos e Espírito Santo, a Bacia Potiguar e a Bacia de Sergipe-Alagoas. Os monitoramentos são realizados por equipes técnicas de diferentes instituições parceiras e focam no registro da fauna alvo, que incluem aves marinhas e limícolas, répteis e mamíferos marinhos.

As informações coletadas incluem classificação taxonômica do indivíduo, sua condição física e características ambientais no local e no momento do registro.

Esta análise visa gerar insights para apoiar a tomada de decisões estratégicas por parte dos órgãos envolvidos, a partir da interpretação dos dados históricos coletados entre 2020 e 2024 pelo Projeto de Monitoramento de Praias. Os objetivos incluem:

- Avaliar a efetividade das ações realizadas pelo PMP, por meio de indicadores-chave de desempenho;
- Identificar padrões geográficos e temporais, contribuindo para um planejamento mais eficiente e preventivo;
- Analisar a frequência e distribuição de diferentes grupos da fauna-alvo, destacando espécies mais vulneráveis ou recorrentes nos registros;
- Investigar indícios de impactos ambientais, com foco especial na presença de óleo e em interações antrópicas prejudiciais.

2. Fontes de Dados e LGPD

A base utilizada foi pública, acessada por meio do site do Sistema de Informação de Monitoramento da Biota Aquática ([SIMBA](#)).

O uso da base atende à LGPD, pois:

- Não contém dados pessoais identificáveis de indivíduos;

- Apresenta informações de interesse público com finalidade ambiental e científica;
- Está disponível para consulta pública e uso educacional, promovendo transparência e responsabilidade socioambiental.

3. Requisitos

- Python 3.12+: Linguagem utilizada para manipulação e análise dos dados;
- Pandas: Biblioteca para análise e manipulação de dados no Python;
- Power BI 2.140.1351.0+: Ferramenta de visualização de dados.

4. Tratamento e Análise Exploratória de Dados

O processo de análise foi conduzido inicialmente em Python no Jupyter notebook, utilizando a biblioteca pandas, e com módulo criado para atender às necessidades do projeto, definido como `utils.py`. Foi criado um [repositório no github](#) para garantir boas práticas de versionamento e modularização no desenvolvimento do projeto.

Uma vez gerado o dataset final tratado e consolidado, ele foi exportado e utilizado no PowerBI para criar métricas e gráficos que atendessem aos objetivos da análise para apresentação e dashboard finais.

I. Estratégia de limpeza e transformação (ETL)

No site do SIMBA, utilizou-se filtro de datas para obter registros entre 01/01/2020 e 31/12/2024. Uma vez baixados os arquivos em csv referentes às diferentes regiões monitoradas pelo PMP, os dados foram armazenados em Google Sheets e importados para o Jupyter notebook através de métodos do pandas.

Colunas interessantes para análise foram previamente definidas, a fim de otimizar o processamento de dados e evitar bases desnecessariamente grandes. As datas foram convertidas para o tipo nativo de séries temporais do pandas durante a importação.

A inspeção inicial dos dados serviu para conferir as importações, checando tipos de dados, filtro das datas, tamanhos dos datasets,

registros duplicados e nulos. Uma vez conferidos, seguiu-se para etapa de transformação, cujo objetivo principal foi de padronizar registros, criar, remover e renomear colunas de acordo com necessidades da análise, para por fim concatenar datasets em um único arquivo consolidado.

As transformações realizadas incluíram:

- Remoção de colunas irrelevantes para os objetivos da análise;
- Criação de coluna "Região", definindo local do registro de acordo com as regiões atendidas pelo programa;
- Padronização da coluna "Interações antrópicas", mantendo somente a informação do tipo de interação e preenchendo nulos com "Ausente";
- Tratamento de registros nulos em "Quantidade de óleo", preenchendo como "Ausente", e remoção de demais nulos;
- Ajuste na coluna "Data/Hora", mantendo apenas a data (sem a informação de horário);
- Padronização da coluna "Cidade", registrando apenas a primeira cidade mencionada quando múltiplas foram inseridas;
- Renomeação de colunas para melhor organização e compreensão dos dados;
- Concatenação dos datasets em um único arquivo consolidado

Ao fim deste relatório encontra-se o dicionário de dados referente ao dataset consolidado final, resultado das etapas de tratamento realizadas neste projeto.

II. Análise Exploratória dos Dados (EDA)

Uma etapa inicial de Análise Exploratória dos Dados foi realizada ainda no Jupyter notebook, que serviu tanto para garantir que os registros estivessem de acordo com o esperado e para entender, de forma geral, a distribuição de categorias na base consolidada.

Para isso, foram calculadas as contagens de valores únicos para cada coluna. Os principais achados dessa etapa foram:

1 - Constatou-se que os esforços de monitoramento são desiguais entre as diferentes regiões, com as Bacias de Sergipe-Alagoas e Potiguar sendo pouco representadas, tanto em volume de registros quanto de cidades monitoradas, quando comparadas às Bacias de Santos e de Campos e Espírito Santo.

Região	Registros	Cidades monitoradas
Bacia de Santos	84225	54
Bacias de Campos e Espírito Santo	24106	28
Bacia de Sergipe-Alagoas	8491	7
Bacia Potiguar	3246	14

2 - As aves foram a classe mais representada nos registros (52%), no entanto, quando analisamos ao nível de ordem, as tartarugas lideram o ranking (43%), ainda representando 3 das 5 espécies mais registradas.

3 - A maioria dos animais registrados foram de espécies não ameaçadas (72%), não tinham vestígios visíveis de óleo (99,9%) e nem de interações antrópicas (84%).

4 - Dos poucos registros com presença de óleo, a maioria foi registrada com quantidade até 25% (74%), já quanto àqueles com interação antrópica, a maioria apresentou interação com pesca (81%).

5. Visualização dos dados com dashboard em PowerBI

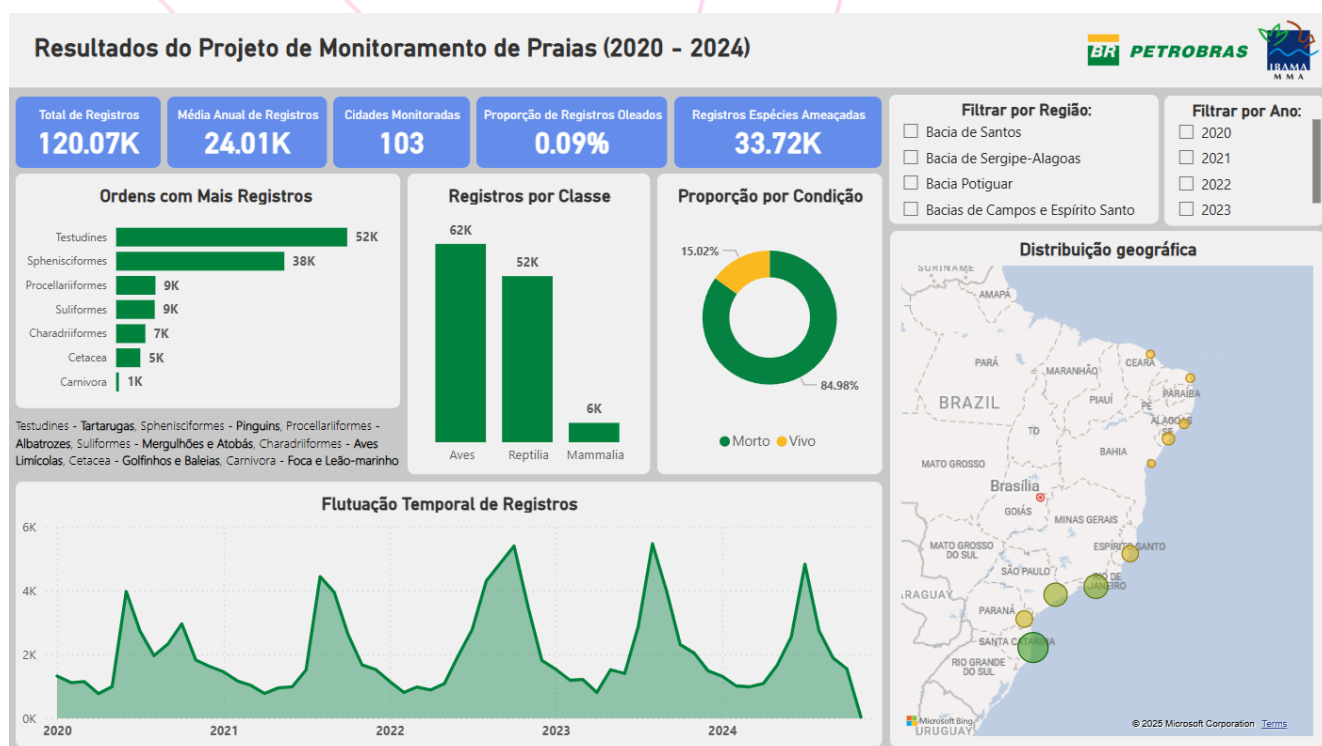
O tratamento e análise exploratória realizados anteriormente permitiram desenvolver o dashboard de forma mais otimizada, diminuindo a necessidade de operações em power query ou DAX que aumentassem a lentidão da visualização.

Ainda assim, algumas operações foram feitas, foram elas:

- Checagem de tipos de dados;
- Criação da coluna “Ano” a partir da coluna “Data”;
- Criação de medida de média anual de registros;
- Criação de medida de percentual de registros oleados.

Essas medidas, juntas ao total de registros, total de cidades monitoradas e total de registros de espécies ameaçadas, foram consideradas KPIs, pois trazem informações relevantes à análise de desempenho do programa.

Abaixo, vemos uma imagem estática do dashboard criado.



Quanto às visualizações, foram criadas as que seguem:

- Gráfico de barras com top 7 ordens com mais registros;
- Gráfico de colunas com distribuição de registros entre as diferentes classes;
- Gráfico de rosca com proporção de registros por condição, morto ou vivo;
- Gráfico de linhas com flutuação temporal de registros;
- Mapa com bolhas representando volumetria de registros nos diferentes estados monitorados.

Além disso, foram adicionados filtros de região e ano, permitindo a exploração dinâmica dos dados em diferentes cenários.

Os números e visualizações trazidos no dashboard focam nos principais objetivos da análise de forma clara e sucinta. O seu uso seria adequado tanto em contexto de análise histórica, como a realizada neste projeto, como para monitoramento contínuo de métricas e resultados úteis à tomada de decisões.

6. Insights e Recomendações

A análise dos dados históricos do Programa de Monitoramento de Praias (PMP) revelou insights cruciais para o aprimoramento das estratégias de monitoramento e conservação da fauna marinha. Foram eles:

1 - Desigualdade no Monitoramento: Bacias de Sergipe-Alagoas e Potiguar são sub-representadas em registros e cidades monitoradas, em contraste com as Bacias de Santos e Campos e Espírito Santo.

2 - Padrões Sazonais Relevantes: A análise ao longo dos cinco anos revelou picos consistentes de registros entre agosto e outubro, especialmente em setembro. Esse padrão indica uma sazonalidade importante, possivelmente associada ao comportamento de alguns grupos de animais, como os de aves migratórias.

2 - Fauna Alvo: Aves são a classe mais registrada (52%), mas tartarugas lideram em nível de ordem (43%), sendo 3 das 5 espécies mais comuns.

3 - Baixa Incidência de Impactos: A maioria dos animais não apresenta vestígios de óleo (99,9%) ou interações antrópicas (84%), sendo a pesca a interação mais comum (81%).

Quanto às recomendações, é pertinente listar:

1 - Expandir o Monitoramento em Regiões Sub-representadas: Sugere-se um estudo aprofundado para entender as causas da baixa representatividade nas Bacias de Sergipe-Alagoas e Potiguar. Com base nos resultados, considerar a expansão dos esforços de monitoramento nessas áreas, incluindo o aumento do número de cidades monitoradas e a frequência das incursões.

2 - Otimizar Esforços de Monitoramento Sazonalmente:

Recomenda-se intensificar os esforços de monitoramento e resposta nos períodos de maior ocorrência, especialmente entre agosto e outubro, conforme identificado nos dados. Além de realizar estudos sobre a variação sazonal dos diferentes grupos taxonômicos, permitindo o planejamento e a execução de ações específicas e direcionadas ao longo do ano, otimizando recursos e aumentando a eficácia da atuação do programa.

3 - Aprimorar o Monitoramento de Presença de Óleo e Interações Antrópicas: Apesar da baixa incidência registrada, muitos animais não foram submetidos a avaliações detalhadas, como necrópsias, o que pode limitar a identificação precisa das causas. É importante investigar os motivos dessa lacuna (como falta de estrutura, equipamentos ou capacitação) e buscar formas de fortalecer os procedimentos de diagnóstico, especialmente em casos suspeitos de contaminação ou interação humana.

4 - Manutenção e Expansão do Dashboard: O dashboard em Power BI provou ser uma ferramenta valiosa. Recomenda-se a manutenção regular do mesmo com dados atualizados e a exploração da adição de novas métricas ou visualizações que possam enriquecer ainda mais a análise e o acompanhamento contínuo dos resultados do PMP.

ANEXO – Dicionário de dados

Campo	Descrição
Código	Código único do registro
Estado	Estado em que foi feito o registro
Estratégia do trecho	Estratégia do trecho monitorado
Data	Data em que foi feito o registro
Classe	Classe taxonômica do animal registrado
Ordem	Ordem taxonômica do animal registrado
Espécie	Espécie do animal registrado
Espécie ameaçada	Se a espécie é ameaçada, podendo ser “Sim” ou “Não”
Caracterização do ambiente	Caracterização do tipo de ambiente em que foi feito o registro
Condição	Condição do animal registrado, podendo ser “Vivo” ou “Morto”
Presença de óleo	Se foi observada presença de óleo no animal, podendo ser “Sim” ou “Não”
Quantidade de óleo	Quantidade de óleo observada, podendo ser “Ausente”, “até 25%”, “até 50%”, “até 75%” ou “mais do que 75%”
Estágio de desenvolvimento	Estágio de desenvolvimento do indivíduo registrado
Interações antrópicas	Se há indícios de interação antrópica no animal, podendo ser “Ausente” ou algum dos tipos listados
Região	Região atendida pelo PMP em que foi feito o registro
Cidade	Cidade em que foi feito o registro