

# Projeto Fauna Marinha (PFM)

**Case Técnico de Gestão de Dados:  
estruturação, qualificação e  
visualização**

Frente: Monitoramento de Tartarugas  
Parceria CEMAM + Comunidade DS

# CASE TÉCNICO

## CENÁRIO ATUAL

- +10 anos de monitoramento ambiental
- Alto volume de registros de campo
- Diversidade de fichas (aves, cetáceos, peixes, quelônios, sirênios, ninhos)
- Crescimento da base que leva ao aumento da complexidade

### Desafio central:

Transformar registros operacionais em informação estratégica.

## DIAGNÓSTICO

- Campos preenchidos como texto livre
- Possibilidade de dados duplicados
- Falta de identificador único
- Planilhas não relacionais
- Dificuldade de análise temporal

### Impacto:

Indicadores instáveis e retrabalho constante

## OBJETIVO

- Garantir padronização
- Aumentar confiabilidade
- Organizar histórico
- Automatizar processos
- Apoiar prestação de contas, captação de recursos, relatórios técnicos e dados para produção científica

### Resultado:

Plano de Ação para estruturar e qualificar a base do Monitoramento de Tartarugas

### Ferramentas:

Google Sheets, Python (Pandas) no Google Colab, Looker

1 Organização e padronização da documentação técnica e dos processos

8

2 Tratamento e padronização dos dados históricos (ETL)



2.1 Melhoria da forma de registro em campo (prevenção de erro)



3 Desenvolvimento de dashboards e indicadores no Looker Studio



**Governança → Estruturação → Prevenção → Visualização**

## DADOS NO CEMAM

Tipo	Exemplo
Estruturado	Planilha de monitoramento
Semi-estruturado	Observação em texto nas Fichas Biométricas
Não estruturado	Fotos e PDFs

### Atividades a serem feitas:

1. Levantamento de todas as bases existentes
2. Identificação de inconsistências
3. Atualização do dicionário de dados
4. Definição de responsáveis (governança)
5. Estruturação do fluxo de dados

## DICIONÁRIO DE DADOS

Conteúdo:

- Nome da coluna
- Tipo de dado (texto, número, data)
- Valores permitidos
- Unidade de medida
- Obrigatoriedade ou Não

Exemplo:

Classe\_carcaca → inteiro (1-5)

Sexo → M / F / Ind

Estagio → Juvenil / Adulto / Ind

## GOVERNANÇA DE DADOS

Papel	Responsabilidade
Coletor(a)	Preencher ficha
Digitador(a)	Inserir na base
Revisor(a)	Conferir
Gestor(a) de Dados	Manter padrão

## FLUXO DOS DADOS

1. Registro em campo
2. Digitalização padronizada
3. Consolidação na base oficial
4. Revisão
5. Atualização do dashboard

## ARQUITETURA DE DADOS

Nível 1 — Armazenamento bruto (Drive): Pastas organizadas por ano, Subpastas por tipo (fotos, fichas, planilhas)

Nível 2 — Processamento (Python no Google Colab): Ler Fichas Biométricas para padronizar nomes de colunas, mapear categorias, ID único, tratamento de nulos, remoção de duplicados, converter datas, validar coordenadas, exportar base limpa

Nível 3 — Base Normalizada (Sheets): Planilhas com campos obrigatórios, Validação de dados ativada, Lista suspensa

Nível 4 — BI (Base normalizada no Sheets alimenta o Looker)

\* Importante ter uma cópia offline das pastas em caso de perda de conteúdo em produção

## RISCOS SE NÃO HOUVER GESTÃO DE DADOS

- Perda histórica de informação
- Inconsistência em relatórios
- Ausência de Monitoramento e Avaliação do impacto na conservação da vida marinha
- Redução de credibilidade para captação de recursos
- Dificuldade em publicar notícias, artigos e livros



### ETL NO CEMAM

Extract -> Carregar as fichas

Transform -> Limpar, padronizar, validar

Load -> Gerar base oficial estruturada

### PASSOS QUE SERÃO FEITOS NO GOOGLE COLAB COM PYTHON:

1. Carregar as fichas
2. Padronização de categorias e conversão dos tipos (o que é texto, número inteiro, número decimal, data)
3. Limpeza de inconsistências e duplicidades
4. Tratamento de nulos
5. Criação de identificadores únicos
6. Separação em modelo relacional
7. Validação de qualidade
8. Base de dados a ser conectada com o Looker



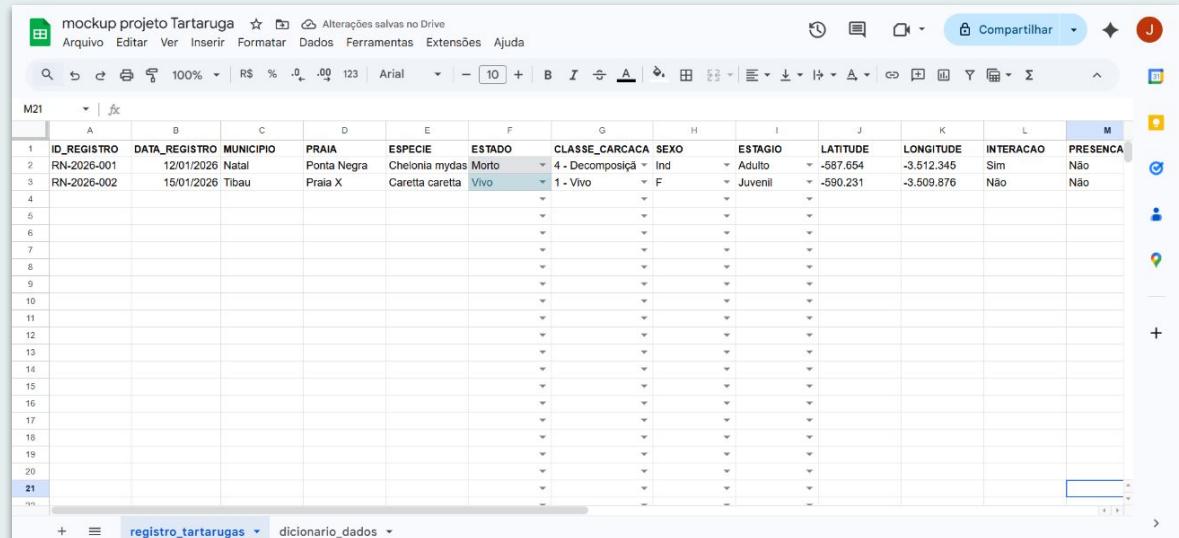
## 2.1 Melhoria da forma de registro em campo (prevenção de erro)

### CONTROLE DE QUALIDADE NAS PLANILHAS DE MONITORAMENTO:

Regras:

- Campos obrigatórios
- Lista Suspensa
- Proteção de células/colunas
- Não permitir texto livre
- Validar latitude e longitude
- Impedir datas inválidas
  
- Logs de alteração
- Backup automático no Sheets

#### MODELO SIMPLIFICADO



The screenshot shows a Google Sheets document titled "mockup projeto Tartaruga". The spreadsheet has a header row with column labels: ID\_REGISTRO, DATA\_REGISTRO, MUNICIPIO, PRAIA, ESPECIE, ESTADO, CLASSE\_CARCACA, SEXO, ESTAGIO, LATITUDE, LONGITUDE, INTERACAO, and PRESENCA. Below the header, there are two rows of data. Row 2 contains: RN-2026-001, 12/01/2026, Natal, Praia Negra, Chelonia mydas, Morto, 4 - Decomposição, Ind, Adulto, -587.654, -3.512.345, Sim, Não. Row 3 contains: RN-2026-002, 15/01/2026, Tibau, Praia X, Caretta caretta, Vivo, 1 - Vivo, F, Juvenil, -590.231, -3.509.876, Não, Não. The rest of the table is empty with rows numbered 4 through 20.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID_REGISTRO	DATA_REGISTRO	MUNICIPIO	PRAIA	ESPECIE	ESTADO	CLASSE_CARCACA	SEXO	ESTAGIO	LATITUDE	LONGITUDE	INTERACAO	PRESENCA
2	RN-2026-001	12/01/2026	Natal	Praia Negra	Chelonia mydas	Morto	4 - Decomposição	Ind	Adulto	-587.654	-3.512.345	Sim	Não
3	RN-2026-002	15/01/2026	Tibau	Praia X	Caretta caretta	Vivo	1 - Vivo	F	Juvenil	-590.231	-3.509.876	Não	Não
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													

## Atividades a serem feitas:

- Definição de KPIs
- Criação de dashboard com filtros estratégicos, visualização espacial, automatização de relatórios

Ano

Município

Espécie

Classe de carcaça

## Indicadores-Chave de Performance (KPIs) com exemplos na CEMAM:

### Monitoramento Geral

- Total de ocorrências por mês
- % mortos vs vivos
- Distribuição por classe de carcaça

### Impacto Antrópico

- % com interação
- Tipo mais frequente
- Região corporal mais afetada

### Tartarugas

- N° de ninhos por temporada
- N° de Ovos
- Taxa média de sucesso



## MODELO SIMPLIFICADO

**CEMAM**  
Programa Fauna Marinha

Panorama Geral de Tartarugas   Mortalidade e Reprodução



# PLANO DE AÇÃO

Atividade/Mês	1	2	3
Padronização da base (variáveis das Fichas Biométricas devem refletir no Colab e no Google Sheets)			
Atualização do dicionário “Documentação Monitora Mar Norte”			
Treinamento e definição da governança			
Migração histórica e Implementação do fluxo			
Testes na Planilha de Monitoramento (entrada de dados -> Python -> Looker)			
Criação de dashboard			
Manual interno			
Deploy			

# MELHORIAS FUTURAS

## AUTOMAÇÃO COM GOOGLE APPS SCRIPT

Possibilidade de reduzir etapas manuais e aumentar controle de qualidade diretamente no Google Sheets.

### O que pode ser adotado:

- Validação automática ao salvar registros
- Geração automática de ID único
- Notificação por e-mail quando houver inconsistências
- Bloqueio de envio se campos obrigatórios estiverem vazios

# OBRIGADA!



Joice Oliveira  
22/02/2026

