



# Monitoramento dos Estoques de Camarões: Variações Mensais e Distribuição Geográfica

Pedro Petroski Ruiz<sup>1</sup>

10390289

<sup>1</sup>Faculdade de Computação e Informática (FCI)  
Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

<sup>2</sup>Programa de pós-graduação em Computação Aplicada – Faculdade de Computação e Informática (FCI) – Universidade Presbiteriana Mackenzie – São Paulo, SP – Brasil

10390289@mackenzista.com.br

**Resumo.** Este projeto apresenta uma análise exploratória e preditiva dos estoques de camarões (Rosa, Sete Barbas, Branco e Santana/Vermelho), com o objetivo de identificar padrões temporais e geográficos, e prever o volume de estoque conforme características como tipo de pessoa, estado, município e local de armazenamento. A base de dados utilizada inclui registros de data de envio, localização e quantidade em quilogramas para cada espécie de camarão.

No pré-processamento, os dados foram otimizados, tratados valores faltantes e categóricos, e normalizados quando necessário. Foi feita análise descritiva para estudar distribuições, tendências sazonais e correlações entre as variáveis. Em seguida, aplicaram-se três algoritmos supervisionados de classificação/predição: Naive Bayes, K-Nearest Neighbors (KNN) e Decision Tree (Árvore de Decisão), avaliando seu desempenho por métricas como acurácia, precisão, recall e matriz de confusão, através de divisão entre conjuntos de treino e teste.

**Palavras-chave:** Estoque de camarões, Análise de Dados, Sazonalidade, Modelos de Classificação

**Abstract.** This project presents an exploratory and predictive analysis of shrimp stocks (Pink, Seven-Beards, White, and Santana/Red), with the goal of identifying temporal and geographic patterns, and forecasting stock volumes based on attributes such as person type, state, municipality, and storage location. The dataset used includes records of shipping date, location, and weight in kilograms for each shrimp species.

In preprocessing, the data were optimized; missing and categorical values were handled; and normalization was applied where needed. A descriptive analysis was conducted to examine distributions, seasonal trends, and correlations among variables. Then, three supervised classification/prediction algorithms were applied: Naive Bayes, K-Nearest Neighbors (KNN), and Decision Tree, evaluating their performance using metrics such as accuracy, precision, recall, and confusion matrix, through splitting into training and test sets.

**Keywords:** *Shrimp Stocks, Data Analysis, Seasonality, Classification Models*

## 1. Introdução

Segundo o *blog Opinião CE* carcinicultura (produção de camarão em cativeiro) tem se destacado como um setor de grande relevância econômica no Brasil, especialmente na Região Nordeste. Segundo dados recentes do IBGE, a produção nacional de camarão alcançou recordes — em 2023, foram produzidas aproximadamente **127,5 mil toneladas**, com um valor de produção de cerca de **R\$ 2,63 bilhões**, crescimento de 13% em comparação ao ano anterior.

Neste contexto, gerir eficientemente os estoques de diferentes espécies de camarão torna-se essencial não só para produtores, mas também para distribuidores e autoridades, com vistas à sustentabilidade econômica, ao planejamento logístico, ao controle de perdas e à garantia de disponibilidade para o consumo e exportação.

Embora existam muitos dados sobre volume de produção, há menos estudos focados em analisar de forma integrada os estoques de camarões diferenciados por espécie, por local de armazenamento, por município/estado, e por tipo de pessoa (ou seja, distinção entre diferentes agentes ou entidades envolvidas).

Entender padrões temporais (por exemplo sazonalidade) e geográficos desses estoques pode auxiliar na previsão de oferta, para tomada de decisão em cadeia de suprimentos, determinar políticas fiscais ou ambientais, otimizar logística e reduzir desperdícios.

Além disso, o uso de técnicas de aprendizado de máquina para prever volumes de estoque pode agregar valor prático, ajudando produtores ou gestores de estoques a se prepararem para variações futuras, melhorarem eficiência operacional e fortalecerem a competitividade do setor.

O objetivo deste trabalho é realizar uma análise exploratória e preditiva dos estoques de camarões (espécies Rosa, Sete Barbas, Branco e Santana/Vermelho), identificando padrões temporais e geográficos, e construir modelos que permitam prever o volume de estoque a partir de variáveis como tipo de pessoa, estado, município e local de armazenamento. Especificamente, busca-se:

1. Descrever as distribuições e tendências sazonais dos estoques de cada espécie ao longo do tempo e entre localidades;
2. Avaliar a correlação entre variáveis relacionadas à localização e tipo de armazenamento com os volumes estocados;
3. Aplicar métodos de aprendizado de máquina (Naive Bayes, K-Nearest Neighbors e Decision Tree) para gerar previsões de estoque e comparar seu desempenho por métricas como acurácia, precisão, recall;

4. Apresentar recomendações práticas com base nos achados, considerando limitações do estudo.

## 2. Referencial Teórico

O referencial teórico apresenta a base conceitual relacionada ao problema de pesquisa e aos objetivos do trabalho. Apresenta, portanto, conteúdos teóricos essenciais que ajudam o leitor a entender conceitos importantes e necessários para o correto entendimento do projeto.

Nessa seção, devem ser apresentados tópicos que possibilitem compreender o estado da arte da área de pesquisa. No entanto, não deve ser entendido como um conjunto de sinopses e resumos de artigos ou livros. Portanto, deve ser construído pelo autor de forma fluída e comprehensível, sempre indicando como as obras referenciadas contribuíram para o desenvolvimento do projeto.

Recomenda-se que essa seção seja organizada em formato de漏il, no sentido de que o tema mais genérico preceda os temas específicos.

## 3. Descrição do Problema

No setor da produção e comercialização de camarões, gerir o estoque de diferentes espécies (como Rosa, Sete Barbas, Branco e Santana/Vermelho) apresenta desafios complexos. Entre eles:

1. **Variação temporal:** As quantidades estocadas de camarão não são uniformes ao longo do tempo — há períodos de maior e menor entrada, envio ou demanda. Essa variabilidade torna difícil prever com precisão os estoques em determinados momentos, o que pode levar a excesso ou falta.
2. **Diferenças geográficas:** Os estoques variam bastante entre estados e municípios, bem como entre diferentes locais de armazenamento. Isso significa que as condições logísticas, climáticas, de infraestrutura e de mercado locais interferem fortemente no volume estocado e sua movimentação.
3. **Heterogeneidade das espécies:** Cada espécie de camarão pode ter diferentes padrões de produção, demanda, percepção ou custos de armazenamento. Isso dificulta tratar todas as espécies da mesma forma, exigindo modelos ou previsões diferenciadas.
4. **Dados incompletos ou inconsistentes:** Muitas vezes, há dados faltantes, erros nos registros ou variáveis categóricas que não estão padronizadas (por exemplo, “tipo de pessoa”, “local de armazenamento”). Isso reduz a confiabilidade das análises descritivas e preditivas.

5. **Decisões de gestão com base em incerteza:** Sem previsões confiáveis, os gestores de estoque ficam tomando decisões com bastante incerteza — se armazenar muito, se arriscam perdas; se armazenar pouco, pode faltar produto. Isso pode gerar desperdício, perdas econômicas ou desabastecimento.
6. **Capacidade e eficiência de armazenamento:** A infraestrutura de armazenamento pode variar em qualidade e capacidade. Lugares com espaço limitado, condições inadequadas ou logística lenta agravam os problemas de estoque — por exemplo, demora no transporte ou na movimentação do camarão entre localidades.

## 4. Metodologia

Para alcançar esses objetivos, o projeto seguirá o seguinte escopo metodológico:

- Coleta ou uso da base de dados contendo registros com variáveis como: data de envio, tipo de pessoa (entidades envolvidas), estado, município, local de armazenamento, e quantidades em quilogramas para cada espécie de camarão (Rosa, Sete Barbas, Branco, Santana/Vermelho).
- Tratamento dos dados no pré-processamento: limpeza, tratamento de valores faltantes, codificação de variáveis categóricas, normalização quando apropriado.
- Análise exploratória: estatísticas descritivas, visualizações (por exemplo séries temporais, mapas geográficos, boxplots) para entender tendências, sazonalidade, outliers, correlações.
- Modelagem preditiva: uso de três algoritmos supervisionados — **Naive Bayes**, **K-Nearest Neighbors (KNN)** e **Decision Tree** — para previsão de estoque. Os modelos serão treinados e testados usando divisão de dados em conjuntos de treino e teste.
- Avaliação de desempenho dos modelos através de métricas como acurácia, precisão, recall e análise da matriz de confusão.
- Interpretação dos resultados: quais variáveis se mostram mais relevantes, qual modelo performa melhor, implicações para gestão de estoque; além de apontar limitações (tamanho da amostra, viés, variáveis não consideradas) e possíveis extensões para trabalhos futuros.

## 5. Referências

[Produção aquícola aumenta 16% no Brasil; Ceará lidera na criação de camarão – Opinião CE](#)

<https://www.ibama.gov.br/sophia/cnia/livros/camaroesmarinhosdobrasildigital.pdf>