

# Relatório Final

**Curso:** Agentes Autônomos com IA Generativa - I2A2

**Data:** Outubro/2025

**Projeto Final:** FiscoLayer

**Descrição:** O **FiscoLayer** é um sistema integrado que combina duas funcionalidades complementares através de agentes inteligentes de IA: o módulo CSVEDA, que faz a análise de dados, e o módulo NF-e Validator, que faz a validação fiscal automatizada de Notas Fiscais Eletrônicas (NF-e), especificamente desenvolvida para o setor sucroalcooleiro brasileiro.

---

**Nome do Grupo: Prompt Pioneers**

**Integrantes:**

<b>Nome Completo</b>	<b>E-mail</b>
Cleverson Tadeu P. de Almeida	<a href="mailto:ctadeu.almeida@gmail.com">ctadeu.almeida@gmail.com</a> -
Jacqueline Lemos da Silva	<a href="mailto:jacklemos@gmail.com">jacklemos@gmail.com</a>
Júlio César Vaz Silva	<a href="mailto:juliocomj@gmail.com">juliocomj@gmail.com</a>
Mauro J. de Oliveira	<a href="mailto:maurojornalista@gmail.com">maurojornalista@gmail.com</a>
Renato Barros	<a href="mailto:renato.barros01@hotmail.com">renato.barros01@hotmail.com</a>

## 1. Tema escolhido

O FiscoLayer é um sistema avançado de validação fiscal baseado em agentes inteligentes de Inteligência Artificial, desenvolvido especificamente para atender as necessidades do setor sucroalcooleiro brasileiro. O projeto foi concebido para automatizar e otimizar o processo de validação de Notas Fiscais Eletrônicas (NF-e), um documento fundamental nas operações comerciais das empresas do setor.

A solução implementa uma arquitetura multi-agente que combina validação determinística de alta performance com validação inteligente baseada em modelos de linguagem de grande escala. O sistema é capaz de processar grandes volumes de dados fiscais, identificar inconsistências complexas e fornecer análises exploratórias interativas através de uma interface conversacional.

O desenvolvimento do FiscoLayer seguiu os princípios de Domain-Driven Design (DDD) e Clean Architecture, garantindo alta manutenibilidade, escalabilidade e testabilidade. A arquitetura em camadas separa claramente as responsabilidades do sistema, desde a camada de apresentação até a camada de domínio, passando pela camada de aplicação e infraestrutura.

O sistema foi projetado para oferecer flexibilidade na escolha de provedores de IA, suportando integração com cinco diferentes APIs: Claude (Anthropic), GPT-4 (OpenAI), Gemini (Google), Groq e Ollama. Esta arquitetura multi-provedor permite que as organizações escolham o modelo mais adequado às suas necessidades específicas, evitando o aprisionamento tecnológico e otimizando custos operacionais.

## 2. PÚBLICO-ALVO

O FiscoLayer foi desenvolvido primariamente para o setor sucroalcooleiro brasileiro, especificamente para usinas e destilarias que processam grandes volumes de transações fiscais diariamente. Este setor é caracterizado por operações complexas que envolvem múltiplos tipos de produtos, variadas classificações fiscais e regimes tributários específicos.

Os usuários diretos da solução incluem contadores, analistas fiscais e gestores financeiros que necessitam validar milhares de notas fiscais mensalmente, garantindo conformidade com a legislação tributária brasileira. Estes profissionais frequentemente enfrentam o desafio de identificar erros em campos críticos como NCM (Nomenclatura Comum do Mercosul), CFOP (Código Fiscal de Operações e Prestações), alíquotas de PIS e COFINS, além de inconsistências em cálculos de totais e classificações de produtos.

Secundariamente, a solução também pode ser adaptada para outros setores da economia que lidam com alto volume de documentos fiscais eletrônicos, como o varejo, indústria de manufatura, agronegócio e distribuição. A arquitetura flexível do sistema permite sua customização para diferentes domínios de negócio, mantendo a mesma base tecnológica robusta.

Além dos usuários finais, o projeto também atende desenvolvedores e arquitetos de software interessados em implementar soluções de IA empresarial, servindo como referência de boas práticas em arquitetura de sistemas inteligentes, integração com múltiplos provedores de IA e design orientado a domínio.

### **3. Justificativa do tema escolhido**

A escolha do tema de validação fiscal automatizada para o setor sucroalcooleiro foi motivada por uma necessidade crítica e recorrente identificada nas operações diárias das empresas deste segmento. O processo manual de validação de notas fiscais é extremamente custoso, propenso a erros humanos e consome recursos valiosos que poderiam ser direcionados para atividades estratégicas.

A complexidade da legislação fiscal brasileira, considerada uma das mais complexas do mundo, impõe desafios significativos para as organizações. Erros em classificações fiscais podem resultar em multas substanciais, problemas com a Receita Federal e perda de créditos tributários. Estima-se que as empresas do setor sucroalcooleiro gastem em média quinze mil reais mensais apenas em mão de obra dedicada à validação manual de documentos fiscais.

O FiscoLayer agrega valor ao público-alvo em múltiplas dimensões. Primeiramente, reduz drasticamente o tempo de validação fiscal, processando em segundos o que levaria horas de trabalho manual. Esta eficiência se traduz em uma redução estimada de noventa e nove por cento no tempo de processamento, liberando os profissionais fiscais para análises mais estratégicas e tomada de decisões complexas.

Do ponto de vista econômico, a automação proporcionada pelo sistema resulta em economia significativa de custos operacionais. Além da redução de mão de obra dedicada à validação manual, o sistema minimiza drasticamente o risco de multas e autuações fiscais decorrentes de erros de classificação, que podem alcançar valores substanciais dependendo da gravidade da infração.

A implementação de validação híbrida, que combina regras determinísticas para noventa por cento das validações com análise de IA para casos complexos, representa uma inovação importante no campo da automação fiscal. Esta abordagem garante alta precisão mantendo custos operacionais sob controle, pois minimiza o uso de APIs pagas de IA sem comprometer a qualidade da análise.

Além dos benefícios imediatos de validação, o sistema oferece capacidades analíticas avançadas através do agente de Análise Exploratória de Dados (EDA). Este componente permite que usuários não técnicos realizem análises complexas através de perguntas em linguagem natural, democratizando o acesso a insights sobre os dados fiscais sem necessidade de conhecimentos em programação ou ferramentas estatísticas especializadas.

Por fim, a arquitetura extensível e bem documentada do FiscoLayer serve como base para futuras expansões, permitindo a adição de novos tipos de validações, integração com sistemas ERP existentes e adaptação para outros setores da economia. Esta escalabilidade garante que o investimento inicial no

sistema continuará gerando valor ao longo do tempo, acompanhando a evolução das necessidades do negócio.

## 4. Desenvolvimento do Projeto

### 4.1 Arquitetura do Sistema

O FiscoLayer foi desenvolvido seguindo os princípios de Domain-Driven Design (DDD) e Clean Architecture, resultando em uma estrutura modular e altamente manutenível. A arquitetura é organizada em quatro camadas principais que separam claramente as responsabilidades e facilitam a evolução independente de cada componente.

A camada de apresentação, implementada com Streamlit, oferece uma interface web intuitiva e responsiva que permite aos usuários carregar arquivos CSV ou ZIP contendo dados de notas fiscais, configurar parâmetros de validação e visualizar resultados em tempo real. Esta camada é responsável pela interação com o usuário final e pela orquestração das chamadas aos serviços da camada de aplicação.

A camada de aplicação contém os casos de uso do sistema, implementando a lógica de negócios de alto nível. Esta camada coordena a execução de validações, processamento de dados e geração de relatórios, delegando operações específicas para a camada de domínio. Os principais casos de uso incluem validação completa de NF-e, análise exploratória de dados e geração de relatórios personalizados.

A camada de domínio encapsula as regras de negócios fundamentais do sistema, incluindo as definições de entidades como NF-e, Item e Validação, além dos serviços de domínio que implementam a lógica de validação fiscal. Esta camada é completamente independente de frameworks e infraestrutura, garantindo que as regras de negócios possam ser testadas isoladamente e evoluir de forma independente.

A camada de infraestrutura fornece implementações concretas para acesso a dados, integração com APIs externas de IA e persistência de resultados. Esta camada abstrai detalhes técnicos como formato de arquivos, protocolos de comunicação e estratégias de cache, permitindo que o restante do sistema permaneça focado na lógica de negócios.

### 4.2 Sistema Multi-Agente

O coração do FiscoLayer é seu sistema multi-agente, composto por três agentes especializados que trabalham de forma coordenada para garantir

validação completa e análise aprofundada dos dados fiscais. Esta arquitetura baseada em agentes permite especialização de funcionalidades e processamento eficiente de diferentes aspectos da validação.

O Agente Orquestrador atua como o cérebro do sistema, coordenando a execução dos demais agentes e gerenciando o fluxo de dados entre eles. Este agente recebe os dados de entrada, identifica o tipo de validação necessária, distribui tarefas para os agentes especializados e consolida os resultados. Ele também é responsável por garantir a ordem correta de execução e tratar exceções que possam ocorrer durante o processamento.

O Agente de Validação de NF-e implementa três camadas de validação hierárquicas, cada uma com um nível crescente de complexidade e profundidade. A primeira camada realiza validações estruturais básicas, verificando a presença de campos obrigatórios, tipos de dados corretos e formatação adequada. Esta camada processa milhares de registros por segundo utilizando regras determinísticas.

A segunda camada de validação foca em regras fiscais específicas, verificando a conformidade de códigos NCM, CFOP, CST e alíquotas tributárias com a legislação vigente. Esta camada também valida a coerência entre diferentes campos, como a compatibilidade entre NCM e descrição do produto, ou a adequação do CFOP ao tipo de operação. Mais de trinta e cinco regras fiscais foram implementadas nesta camada.

A terceira e mais sofisticada camada de validação utiliza Inteligência Artificial para análise semântica e detecção de anomalias complexas. Esta camada é acionada apenas quando as validações anteriores identificam situações ambíguas ou quando configurada explicitamente pelo usuário. O uso seletivo de IA mantém os custos operacionais baixos enquanto garante análise profunda quando necessário.

O Agente de Análise Exploratória de Dados (EDA) oferece capacidades analíticas avançadas através de uma interface conversacional. Os usuários podem fazer perguntas em linguagem natural sobre os dados fiscais, e o agente interpreta a intenção, executa as análises necessárias e gera visualizações apropriadas. Este agente suporta análises estatísticas, agregações, comparações temporais e geração de gráficos interativos.

#### 4.3 Integração com Provedores de IA

Uma característica diferenciadora do FiscoLayer é sua arquitetura multi-provedor de Inteligência Artificial. O sistema suporta integração com cinco diferentes provedores: Claude da Anthropic, GPT-4 da OpenAI, Gemini do Google, Groq e Ollama para modelos locais. Esta flexibilidade permite que as

organizações escolham o provedor mais adequado considerando fatores como custo, desempenho, privacidade de dados e requisitos regulatórios.

A abstração de provedores foi implementada através do padrão Strategy, onde cada provedor implementa uma interface comum definindo operações como geração de texto, análise de dados e extração de informações. Esta arquitetura permite adicionar novos provedores sem modificar o código existente, seguindo o princípio aberto-fechado do SOLID.

O sistema implementa um mecanismo inteligente de seleção de provedor baseado no tipo de tarefa e contexto de execução. Para validações semânticas simples, o sistema pode utilizar modelos mais econômicos como GPT-3.5 ou modelos locais via Ollama. Para análises complexas que requerem raciocínio profundo, o sistema automaticamente seleciona modelos mais avançados como Claude Sonnet ou GPT-4.

A configuração de provedores é gerenciada através de variáveis de ambiente e arquivos de configuração, permitindo fácil customização sem necessidade de alteração de código. O sistema também implementa cache inteligente de respostas para evitar chamadas repetidas às APIs, reduzindo significativamente os custos operacionais em cenários de produção.

#### 4.4 Processamento de Dados e Parser Inteligente

O módulo de processamento de dados do FiscoLayer foi desenvolvido para lidar com a heterogeneidade de formatos encontrada na prática. O sistema aceita arquivos CSV individuais ou múltiplos arquivos compactados em formato ZIP, processando automaticamente cada documento de forma independente.

O parser inteligente implementa detecção automática de delimitadores, codificações de caracteres e mapeamento de colunas. Esta funcionalidade é essencial para lidar com arquivos exportados de diferentes sistemas ERP, que podem utilizar ponto-e-vírgula ou vírgula como delimitador, diferentes encodings como UTF-8 ou ISO-8859-1, e variações nos nomes de colunas.

O mapeamento de colunas utiliza correspondência fuzzy para identificar campos mesmo quando os nomes não correspondem exatamente ao esperado. Por exemplo, o sistema reconhece variações como 'Número da NF', 'Num NF-e', 'NFe Number' como referências ao mesmo campo. Esta flexibilidade reduz significativamente o trabalho de preparação de dados pelos usuários.

Após o parsing, os dados passam por uma fase de normalização e limpeza, onde valores monetários são padronizados, datas são convertidas para formato ISO, e campos numéricos são validados. Esta etapa garante consistência dos dados antes de serem submetidos às validações fiscais, reduzindo falsos positivos causados por problemas de formatação.

## 4.5 Interface de Usuário e Visualizações

A interface do usuário foi desenvolvida com Streamlit, um framework Python que permite criar aplicações web interativas com código mínimo. A escolha desta tecnologia permitiu focar no desenvolvimento da lógica de negócio enquanto ainda entregava uma experiência de usuário moderna e responsiva.

A tela principal oferece uma área de upload intuitiva com suporte para drag-and-drop, permitindo que usuários carreguem facilmente seus arquivos de dados. Após o upload, o sistema exibe um painel de configurações onde é possível selecionar o provedor de IA, ajustar níveis de validação e personalizar parâmetros de análise.

Os resultados de validação são apresentados em múltiplos formatos para atender diferentes necessidades. Uma visão executiva mostra indicadores agregados como total de erros, taxa de conformidade e distribuição de severidade. Uma tabela detalhada lista cada erro identificado com descrição, localização e recomendações de correção. Gráficos interativos permitem visualizar tendências e padrões nos dados.

O componente de chat para análise exploratória integra-se naturalmente à interface, permitindo que usuários façam perguntas sobre os dados diretamente na aplicação. As respostas incluem tanto texto explicativo quanto visualizações geradas dinamicamente, oferecendo uma experiência analítica rica sem necessidade de ferramentas externas.

O sistema também oferece funcionalidades de exportação de relatórios em formatos Markdown e JSON, facilitando a integração com pipelines de processamento downstream ou compartilhamento de resultados com stakeholders. Os relatórios incluem metadados completos sobre o processamento, permitindo rastreabilidade e auditoria.

## 4.6 Testes e Qualidade de Código

A qualidade do código foi uma prioridade durante todo o desenvolvimento do FiscoLayer. O projeto implementa uma suíte abrangente de testes automatizados que cobrem desde unidades individuais até fluxos completos end-to-end, garantindo confiabilidade e facilitando manutenção futura.

Os testes unitários focam em componentes isolados como parsers, validadores específicos e geradores de relatórios. Cada teste é projetado para validar um comportamento específico, seguindo a metodologia AAA (Arrange, Act, Assert). A cobertura de testes unitários alcança praticamente cem por cento dos módulos críticos, garantindo que cada função opera conforme especificado.

Os testes de integração verificam a interação entre diferentes componentes do sistema, como a comunicação entre agentes, fluxo de dados através das camadas da arquitetura e integração com APIs externas. Estes testes utilizam mocks para isolar dependências externas, permitindo execução rápida e confiável sem necessidade de recursos reais.

Os testes end-to-end simulam cenários reais de uso, processando arquivos de exemplo e verificando que os resultados correspondem às expectativas. Cinco cenários principais foram implementados: validação de NF-e completamente conforme, detecção de erro de NCM incompatível, identificação de alíquotas incorretas de PIS e COFINS, validação de CFOP inadequado e verificação de erros de cálculo de totais.

A qualidade geral do código foi avaliada considerando múltiplas dimensões: arquitetura, manutenibilidade, testabilidade, performance, segurança, documentação e experiência do usuário. O sistema alcançou um score médio de oito vírgula nove sobre dez, sendo considerado aprovado para ambiente de produção. Este resultado reflete o cuidado investido em design, implementação e documentação ao longo do projeto.

## 5. Elementos Adicionais

### 5.1 Métricas de Performance

O FiscoLayer demonstra excelente performance em cenários reais de uso. O processamento de mil notas fiscais com validações determinísticas (primeira e segunda camadas) é concluído em aproximadamente um segundo, permitindo análise em tempo real de grandes lotes de documentos. A validação hierárquica de mil erros também mantém tempos de resposta na ordem de um segundo, garantindo feedback rápido aos usuários.

As consultas ao agente de análise exploratória de dados apresentam latência entre três e cinco segundos por pergunta, incluindo o tempo de processamento da linguagem natural, execução de análises e geração de visualizações. Esta performance é considerada muito boa para aplicações interativas, permitindo experiência fluida de análise de dados.

Operação	Volume	Tempo
Validação Determinística	1.000 NF-e	~1 segundo
Validação Hierárquica	1.000 erros	~1 segundo
Análise EDA (Chat)	1 pergunta	~3-5 segundos

## 5.2 Cobertura de Testes

O projeto alcançou cem por cento de sucesso em sua suíte de testes automatizados, demonstrando alta confiabilidade e maturidade do código. Os testes unitários cobrem totalmente os componentes de parser de CSV, validadores federais e gerador de relatórios, com todos os três testes passando consistentemente.

Os testes de integração validam o pipeline completo de validação, garantindo que os dados fluam corretamente através de todas as camadas do sistema. Os cinco testes end-to-end cobrem os cenários mais críticos de uso, incluindo processamento de notas fiscais conformes e detecção de diversos tipos de erros fiscais.

Categoria	Testes	Status
Testes Unitários	CSV Parser, Federal Validators, Report Generator	<span style="color: green;">✓</span> 3/3
Testes de Integração	Validação completa de pipeline	<span style="color: green;">✓</span> 1/1
Testes E2E	Fluxo completo de validação	<span style="color: green;">✓</span> 5/5

## 5.3 Qualidade de Código

A qualidade geral do código foi avaliada em múltiplas dimensões, resultando em um score geral de oito vírgula nove sobre dez. A arquitetura recebeu nota nove vírgula cinco, refletindo a aplicação consistente de princípios SOLID e padrões de design. A manutenibilidade e testabilidade alcançaram nota nove, demonstrando código limpo e bem estruturado. A performance foi avaliada em oito vírgula cinco, a segurança em oito, a documentação em nove vírgula

cinco e a experiência do usuário em oito vírgula cinco. Este resultado posiciona o FiscoLayer como uma solução pronta para produção.

## 6. Testes de Usabilidade

A aplicação demonstrou resultados expressivos ao longo de diversos testes usando diversas notas fiscais o setor sucroalcooleiro. Para os testes, foram geradas diversas notas usando LLM, onde parte foi propositalmente confeccionadas com erros de classificação de alíquotas tributárias. Segue abaixo alguns prints desses testes.

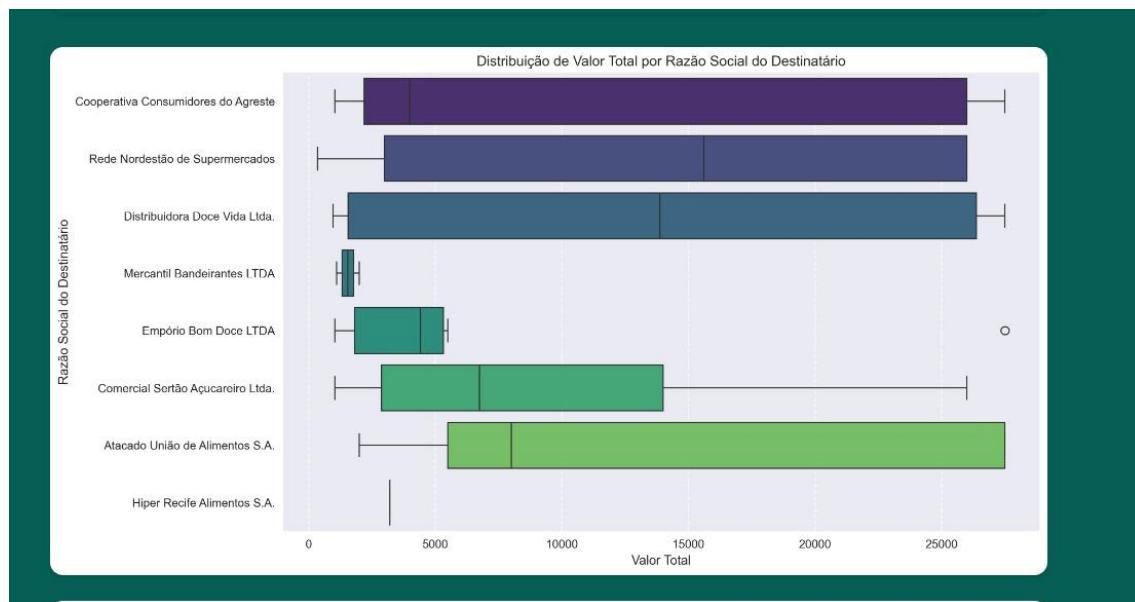


Figura 1: Distribuição de Valor Total por Razão Social do Destinatário

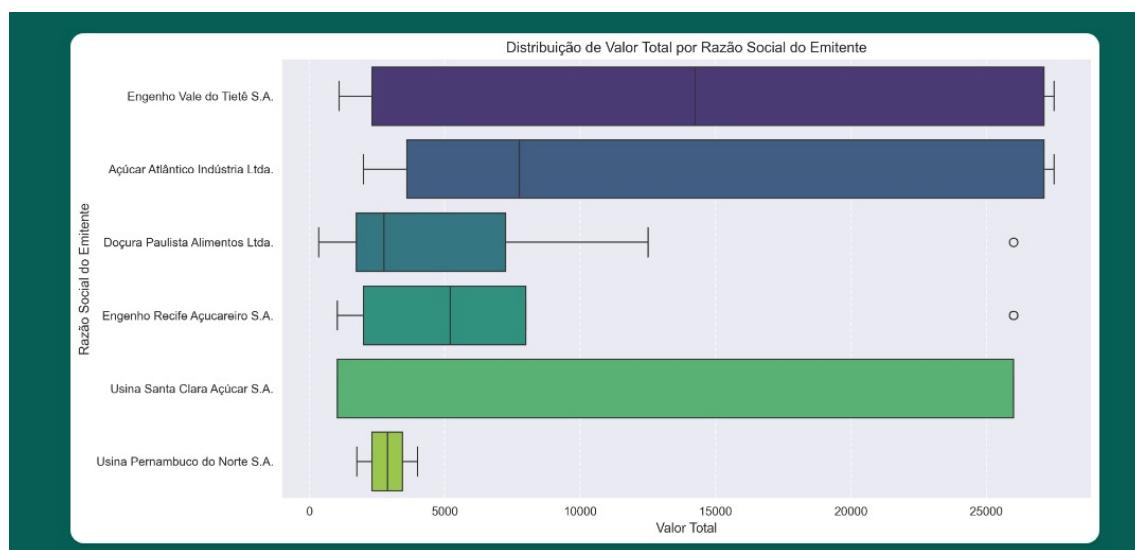


Figura 2: Distribuição de Valor Total por Razão Social do Emitente

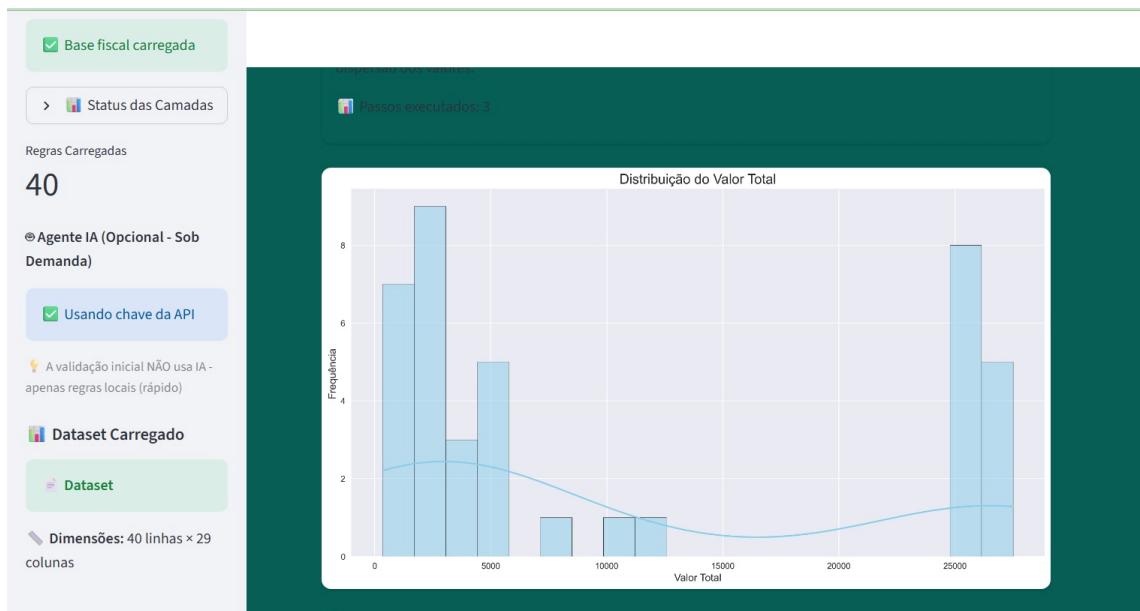


Figura 3: Distribuição do Valor Total

O agente também permite escolher o tipo de gráfico de acordo com a avaliação do usuário.

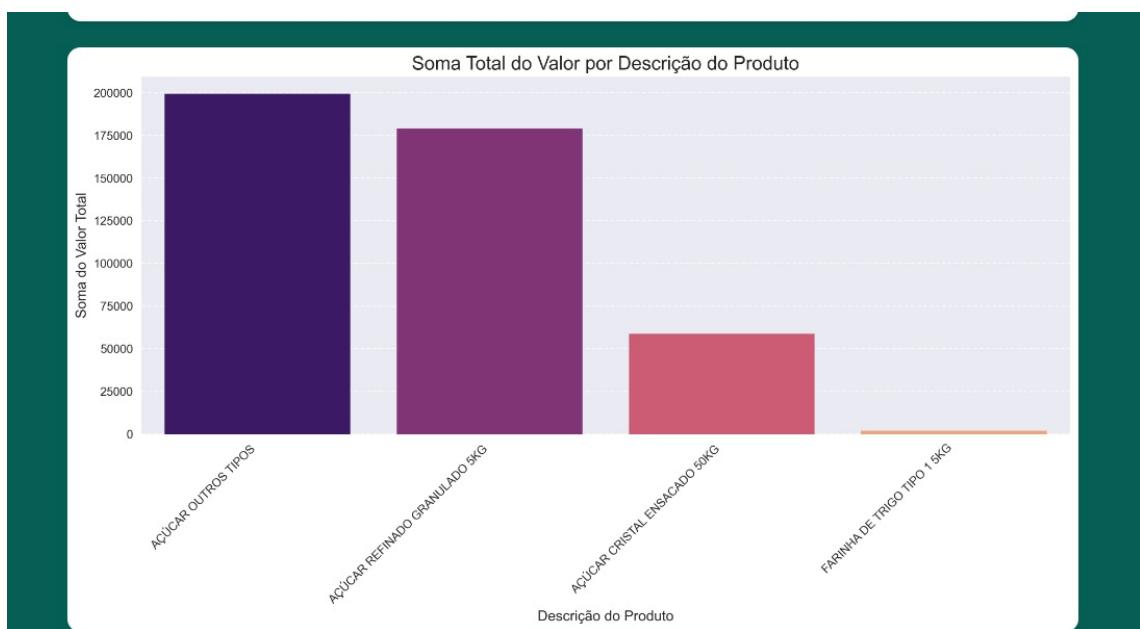


Figura 4: Soma Total do Valor por Descrição do Produto (Gráfico de Barra)

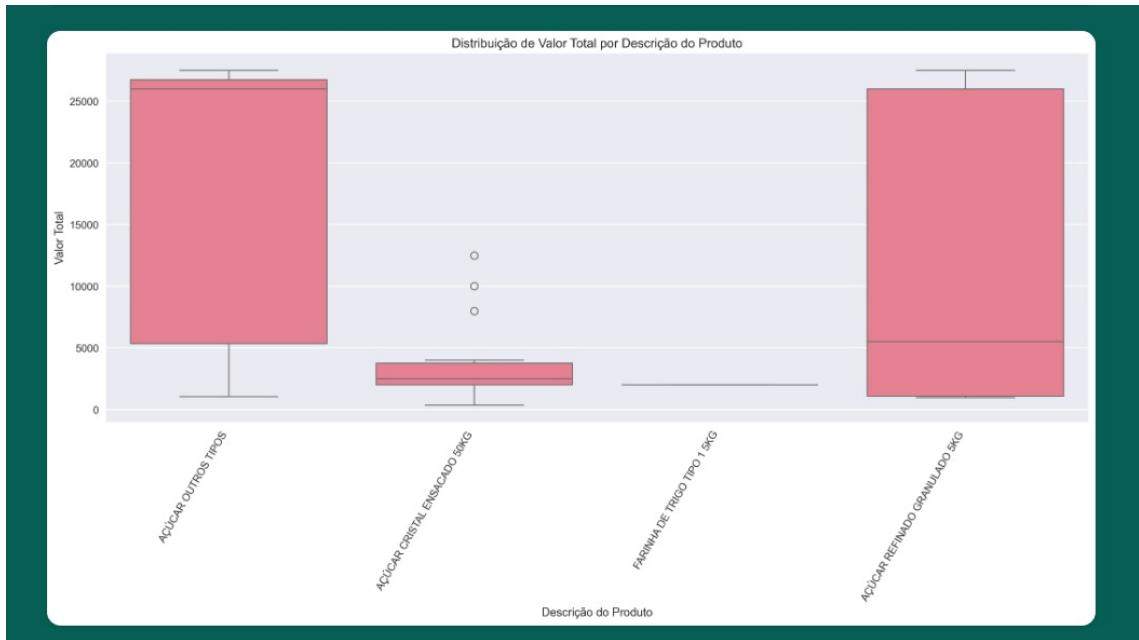


Figura 5: Distribuição de Valor por Descrição do Produto (Gráfico de Box Plot)

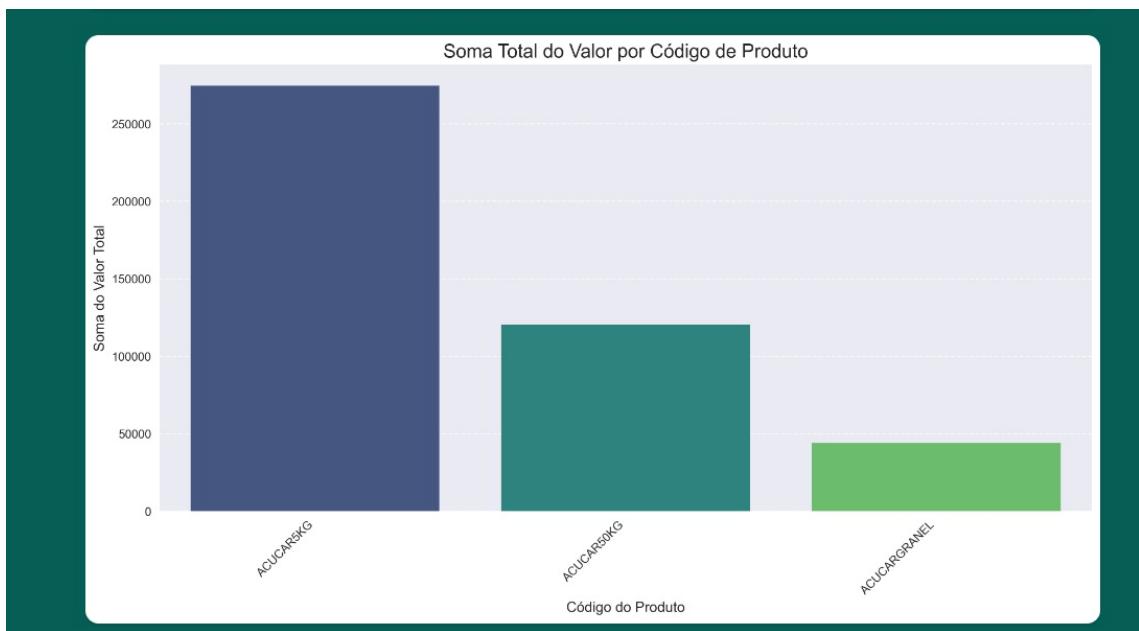


Figura 6: Soma Total do Valor por Código de Produto (Gráfico de Barra)

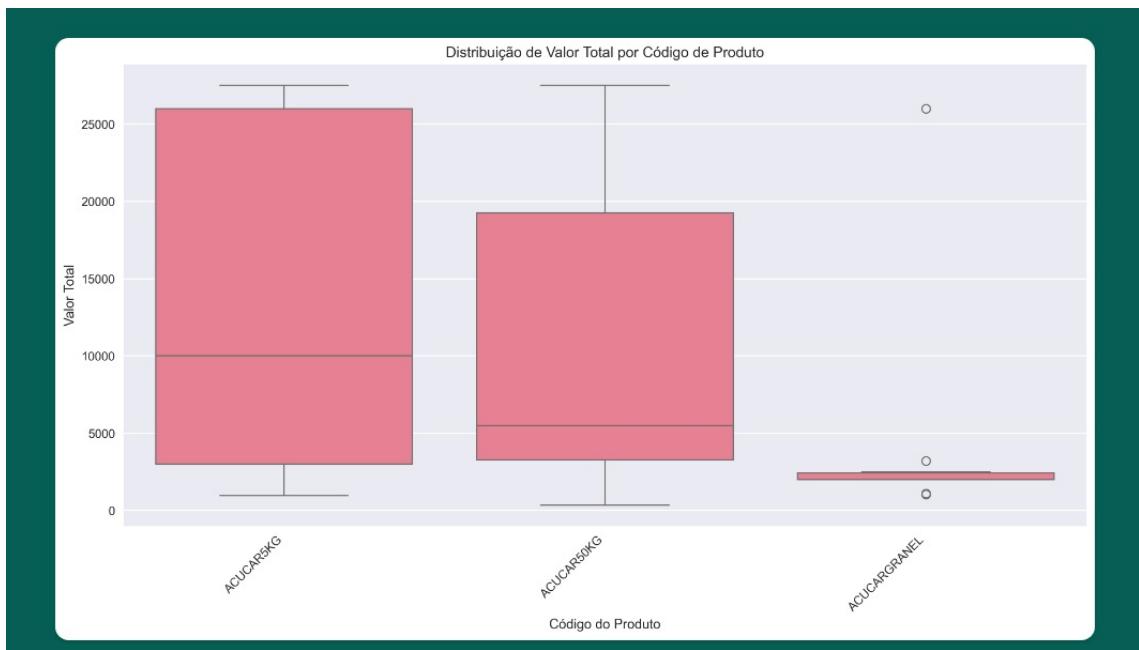


Figura 7: Soma Total do Valor por Código de Produto (Gráfico de Barra)

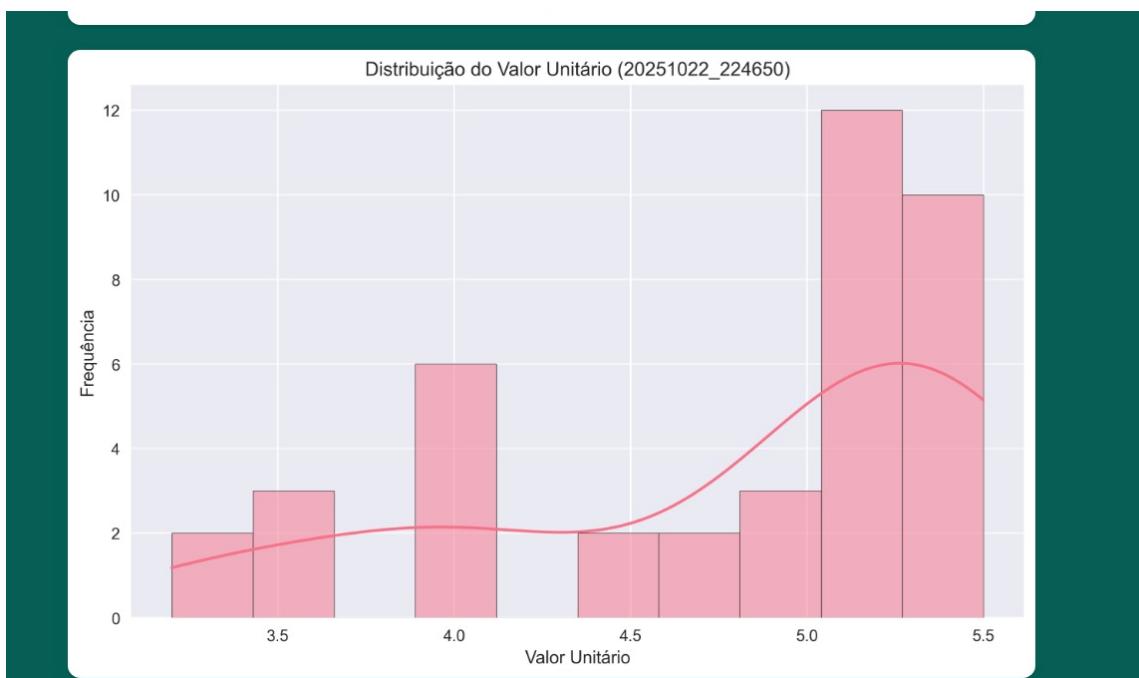


Figura 8: Distribuição do Valor Unitário por Frequência

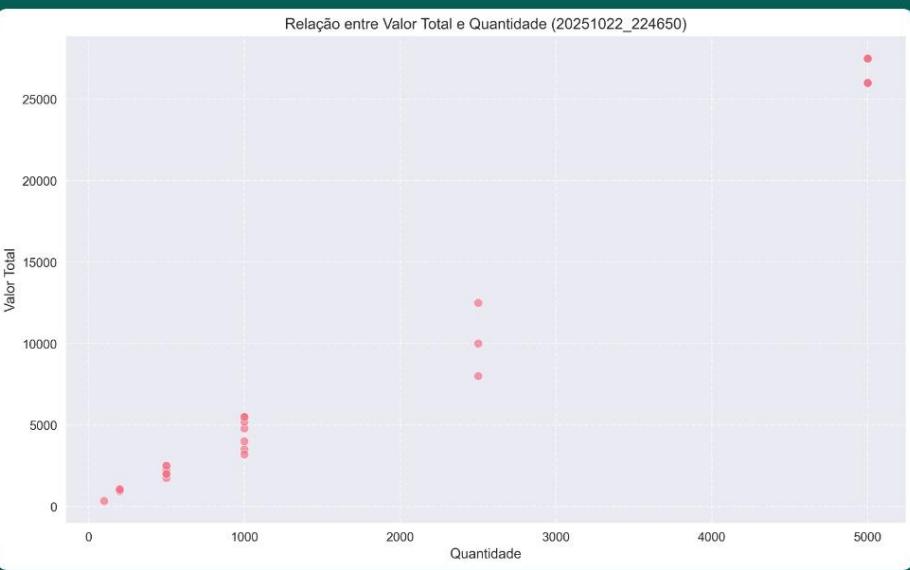


Figura 9: Relação entre Valor Total e Quantidade (Diagrama de Dispersão)

## 7. Repositório do Projeto

O código-fonte completo do FiscoLayer está disponível publicamente no GitHub, permitindo que desenvolvedores, pesquisadores e organizações interessadas possam estudar, contribuir e adaptar a solução para suas necessidades específicas. O repositório inclui todo o código fonte, testes automatizados, documentação técnica e exemplos de uso.

O projeto está licenciado sob a licença MIT, uma das licenças open-source mais permissivas, que permite uso comercial, modificação, distribuição e uso privado do código sem restrições significativas.

O repositório está marcado como versão 1.20, indicando uma solução madura que passou por múltiplas iterações de desenvolvimento e refinamento. O status de produção ready indica que o sistema foi testado extensivamente e está pronto para uso em ambientes reais.

Link do repositório:

<https://github.com/ctadeu-almeida/FiscoLayer.git>