

# Relatório

## 1. Framework escolhido:

Os agentes foram desenvolvidos em Python, utilizando o LangChain como framework principal para orquestração e integração entre componentes. Para a interface de interação com o usuário, foi empregado o Streamlit, que possibilitou a construção de uma aplicação web simples e interativa. Já a comunicação com o modelo de linguagem foi realizada por meio da API do Gemini, que serviu como LLM (Large Language Model) para geração de respostas e processamento de linguagem natural.

## 2. Como a solução foi estruturada:

A solução foi desenvolvida com o Streamlit como interface principal de interação com o usuário. O sistema é composto por diferentes agentes especializados, organizados da seguinte forma:

Agente Orquestrador: responsável por coordenar o fluxo da aplicação e encaminhar as solicitações para os agentes adequados.

Agente de Prompt: analisa a solicitação do usuário e define o tipo de resposta mais apropriado. Esse agente pode:

Gerar diretamente uma resposta em texto quando o conteúdo for simples;

Elaborar um prompt de programação a ser enviado ao agente programador, quando for necessário criar ou executar código;

Produzir um prompt para geração de gráficos, quando a resposta demandar visualizações de dados.

Além disso, a aplicação mantém o histórico da conversa, permitindo um contexto contínuo e uma interação mais natural e eficiente com o usuário.

## 3. Perguntas e respostas

Pergunta 1: Pode me dizer sobre o que se trata esse arquivo?

Resposta: Este arquivo parece conter dados de transações financeiras, possivelmente transações de cartão de crédito, com o objetivo de detectar fraudes. As colunas V1 a V28 são componentes principais resultantes de uma transformação PCA (Análise de Componentes Principais), que visa preservar a privacidade dos dados. A coluna 'Time' indica o tempo decorrido desde a primeira transação no conjunto de dados. 'Amount' representa o valor da transação e 'Class' é a variável alvo, onde 1 indica uma transação fraudulenta e 0 indica uma transação legítima.

Pergunta 2: Quais os tipos de dados do arquivo?

Resposta: Time float64 V1 float64 V2 float64 V3 float64 V4 float64 V5 float64 V6 float64 V7 float64 V8 float64 V9 float64 V10 float64 V11 float64 V12 float64 V13 float64 V14 float64 V15 float64 V16 float64 V17 float64 V18 float64 V19 float64 V20 float64 V21 float64 V22 float64 V23 float64 V24 float64 V25 float64 V26 float64 V27 float64 V28 float64 Amount float64 Class int64 dtype: object.

Pergunta 2.1: Pode contextualizar para mim?

Resposta: O conjunto de dados contém informações sobre transações financeiras, com colunas representando tempo, valores transformados por PCA (V1-V28), valor da transação (Amount) e uma coluna indicando se a transação é fraudulenta ou não (Class). A maioria das colunas são do tipo float64, com 'Class' sendo do tipo int64.

Pergunta 3: Qual a média e a mediana da coluna amount?

Resposta: Média da coluna 'Amount': 88.35 Mediana da coluna 'Amount': 22.00

Pergunta 4: Qual o desvio padrão da coluna amount?

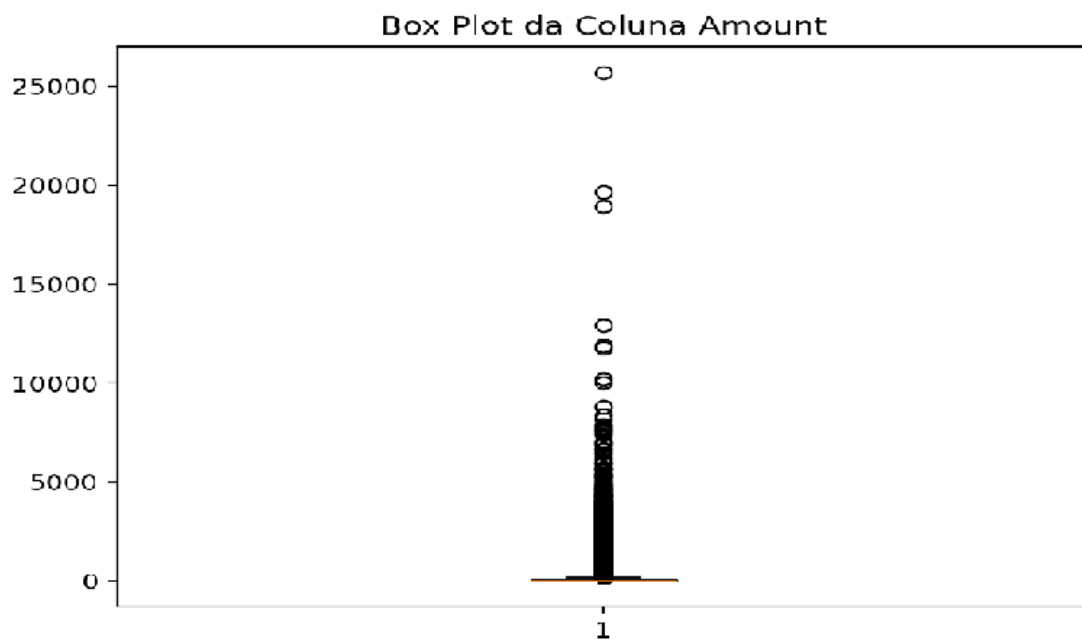
Resposta: O desvio padrão da coluna 'Amount' é: 250.1201092401885

Pergunta 5: Levando em consideração os valores da media, mediana e desvio padrão da coluna Amount, que conclusão posso tirar dessa coluna?

Resposta: A média da coluna 'Amount' (88.35) é significativamente maior que a mediana (22.00), e o desvio padrão (250.12) é alto. Isso sugere que a coluna 'Amount' possui uma distribuição assimétrica à direita, com a presença de outliers de valores altos que puxam a média para cima.

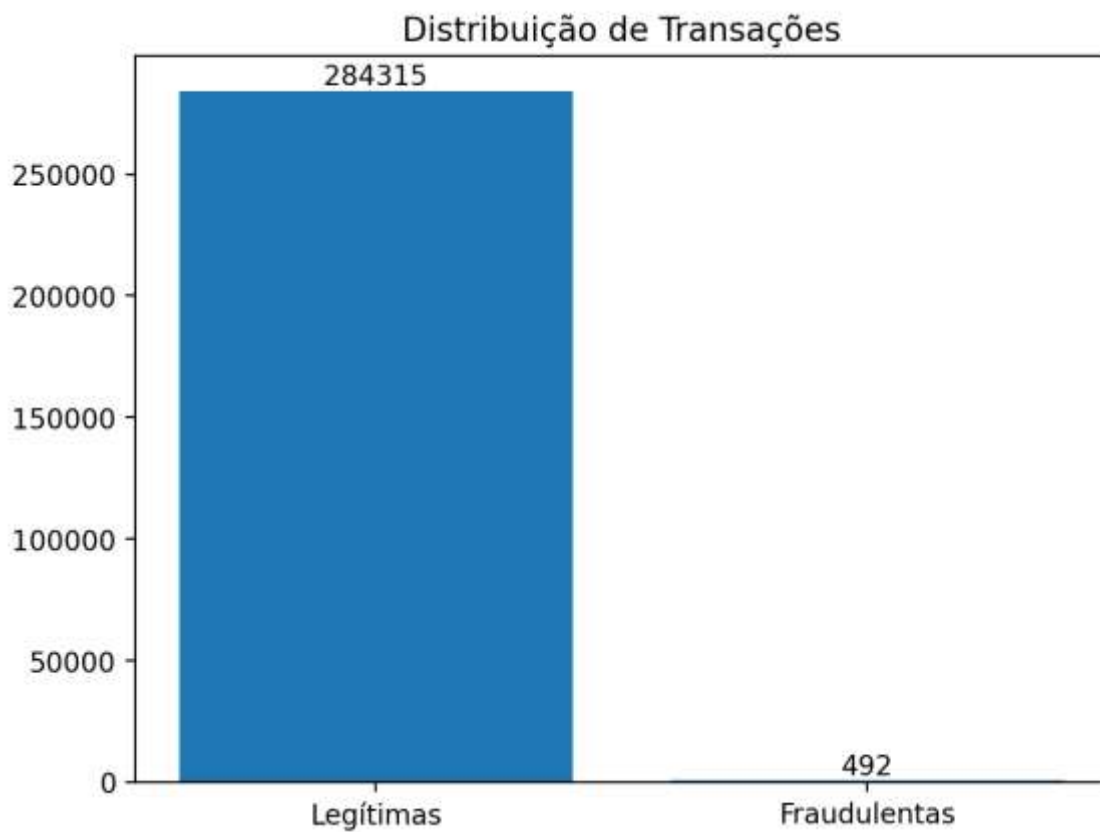
Pergunta 6: Gera para mim um gráfico box plot da coluna amount

Resposta:



Pergunta 7: Gera um gráfico de barras com o total de transações fraudulentas e o total das transações legítimas

Resposta:



Pergunta 8:

Qual as variáveis parecem ter maior ou menor influência sobre as outras?

Resposta: Top 5 Positive Correlations: V7 Amount 0.397311 Amount V7 0.397311 V20 0.339403 V20 Amount 0.339403 Amount V6 0.215981 dtype: float64

Top 5 Negative Correlations: V2 Amount -0.531409 Amount V2 -0.531409 Time V3 -0.419618 V3 Time -0.419618 V5 Amount -0.386356 dtype: float64

Pergunta 9: Que conclusões posso tirar sobre isso?

Resposta: Com base nos dados e análises anteriores, pode-se concluir que a coluna 'Amount' tem uma distribuição assimétrica com outliers, e apresenta correlações notáveis com as colunas V2 e V7. A coluna 'Class' (fraude) tem correlações negativas mais fortes com V3, V10, V12, V14, V17 e V18, e uma correlação positiva com V4 e V11.