

MINERAÇÃO DE DADOS

**Descoberta de Conhecimento em
Bases de Dados (KDD)**

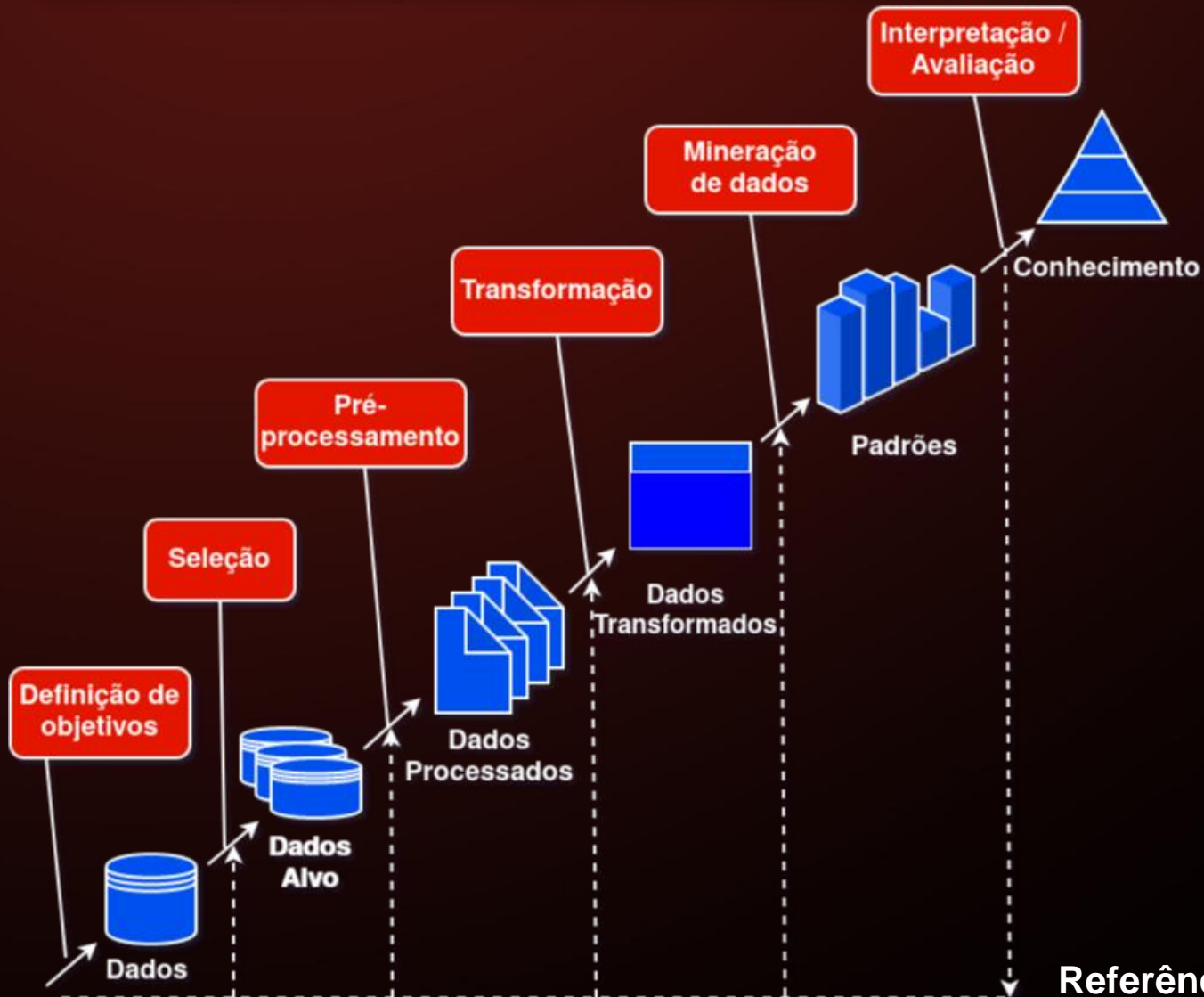


DESCOBERTA DE CONHECIMENTO EM BASES DE DADOS (KDD)

Knowledge Discovery in Databases (KDD)

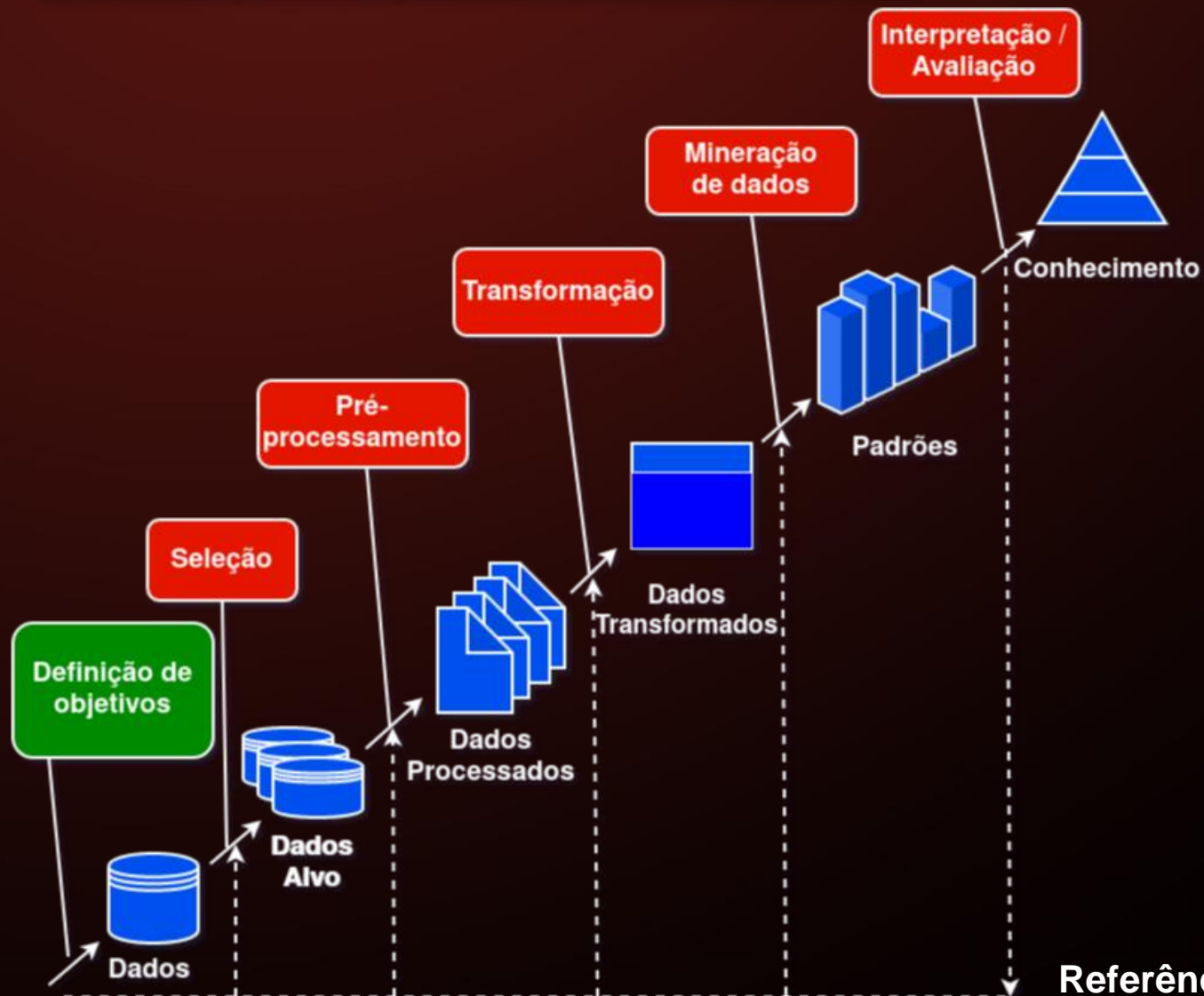
- **Processo de transformação de volumes de dados em conhecimento**
- **Etapas iterativas e interativas**

PROCESSO DE KDD



Referência: Schneider [1]

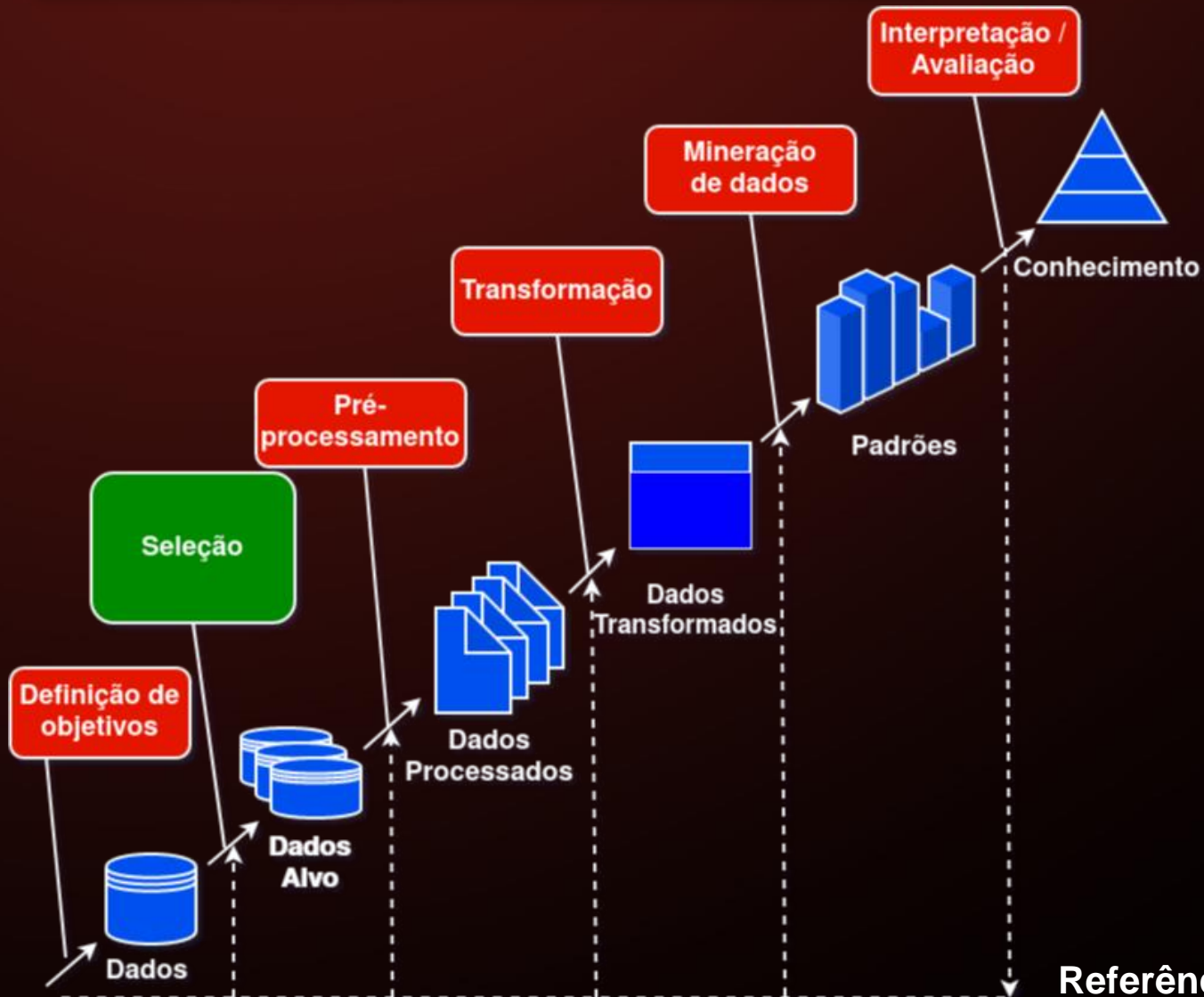
PROCESSO DE KDD



Entender o domínio

Referência: Schneider [1]

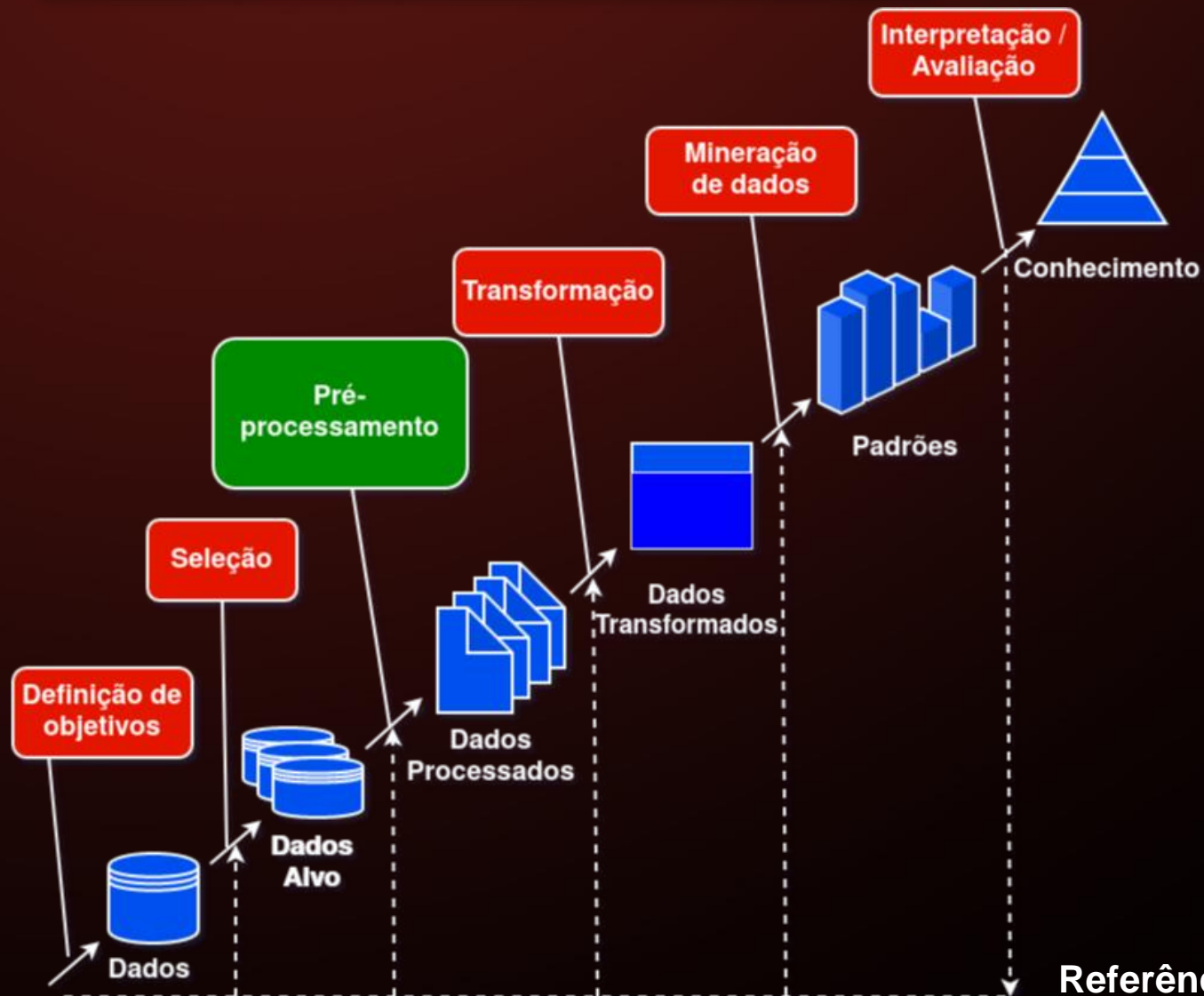
PROCESSO DE KDD



Identificar fontes de dados

Referência: Schneider [1]

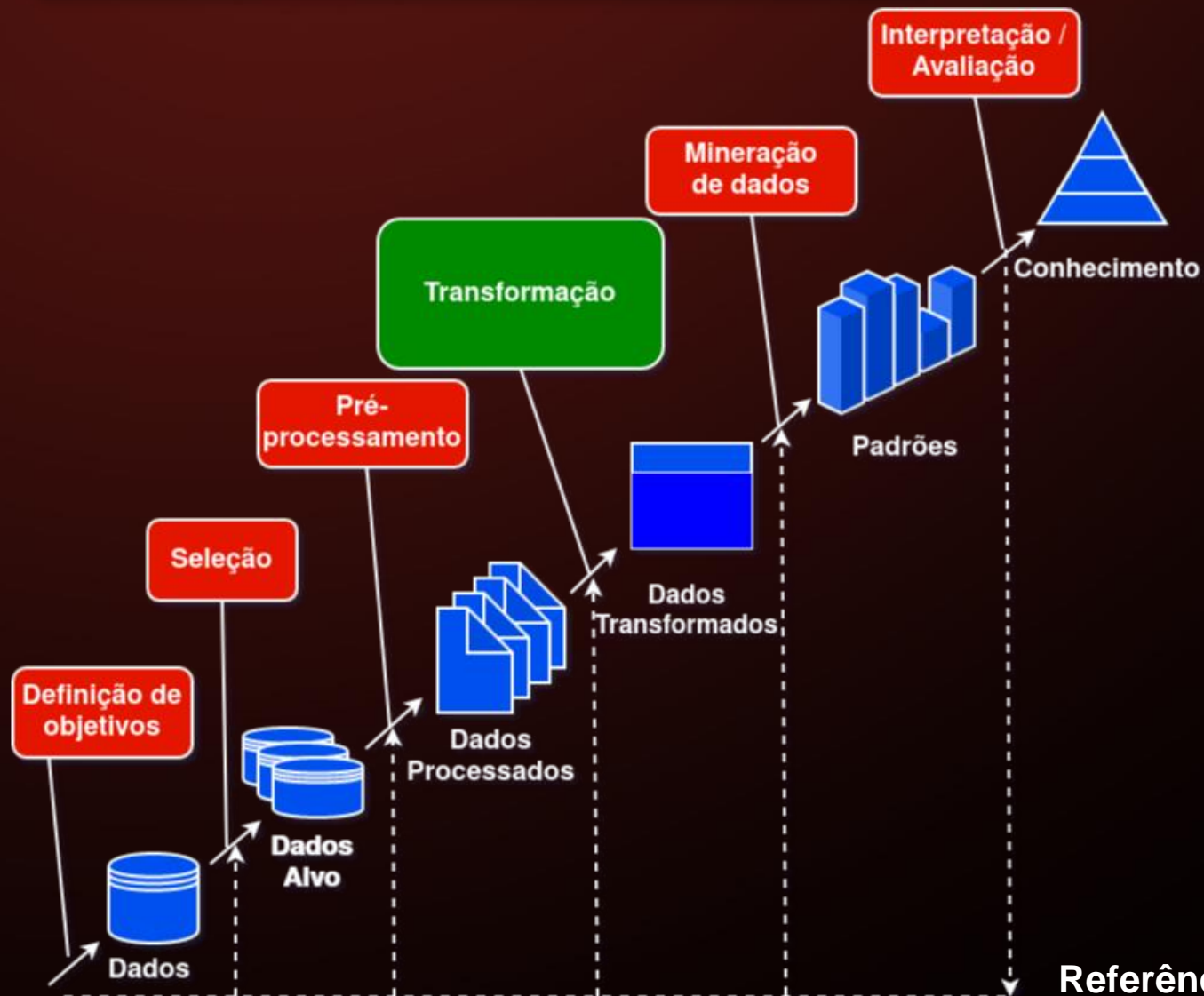
PROCESSO DE KDD



Limpeza dos dados

Referência: Schneider [1]

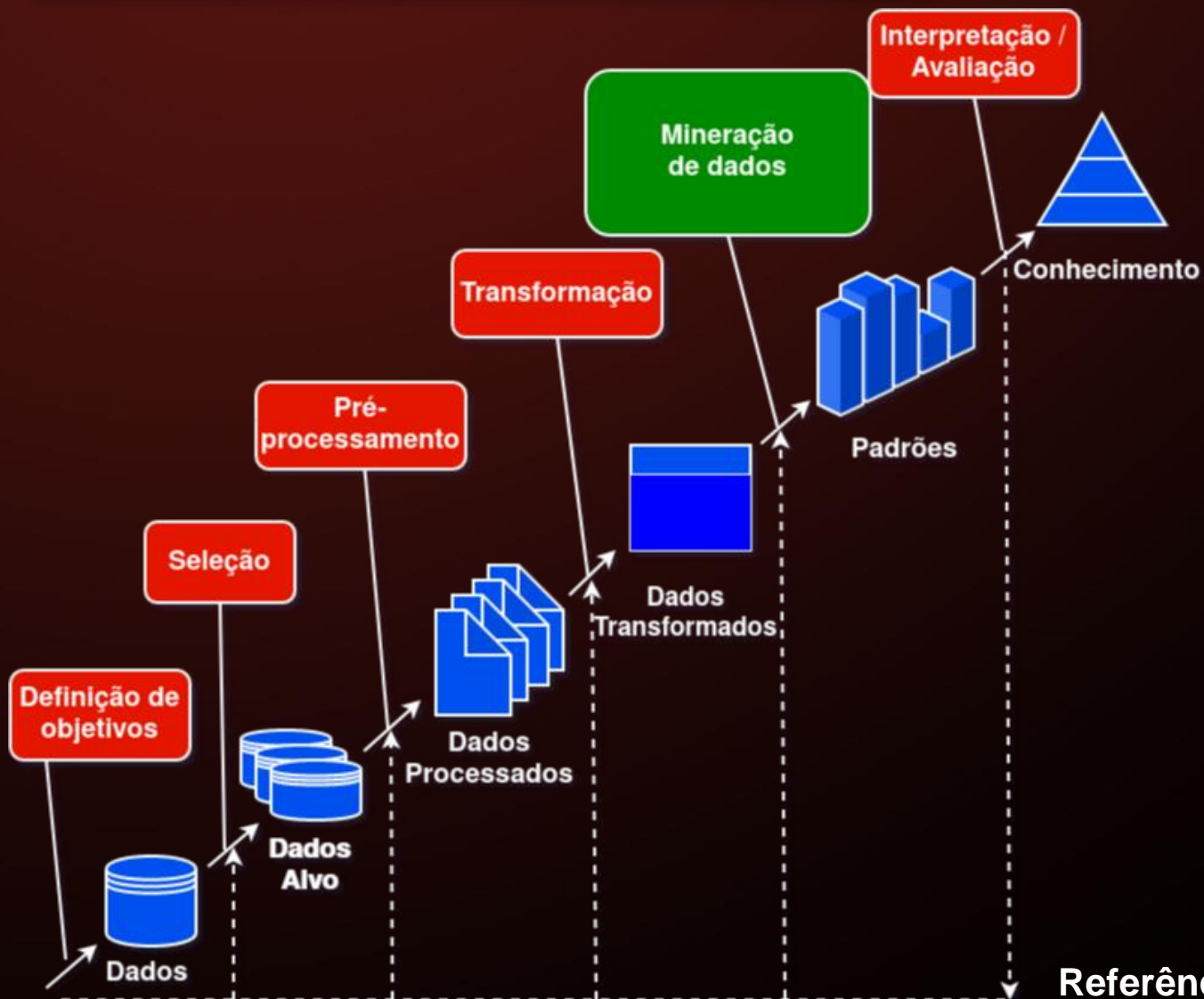
PROCESSO DE KDD



Formatação para uso /
armazenamento

Referência: Schneider [1]

PROCESSO DE KDD

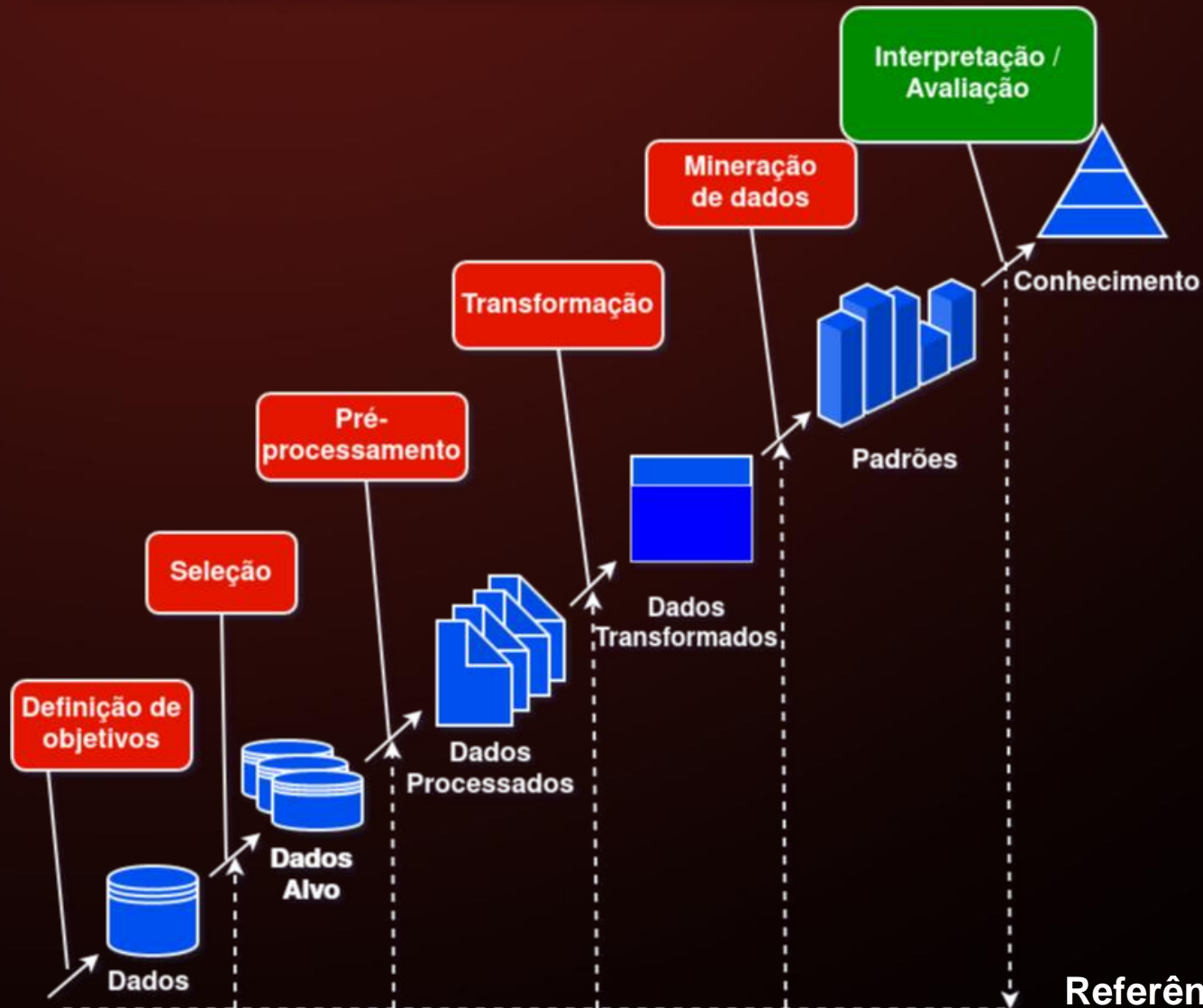


Núcleo do processo de KDD

Identificar padrões nos dados

Referência: Schneider [1]

PROCESSO DE KDD

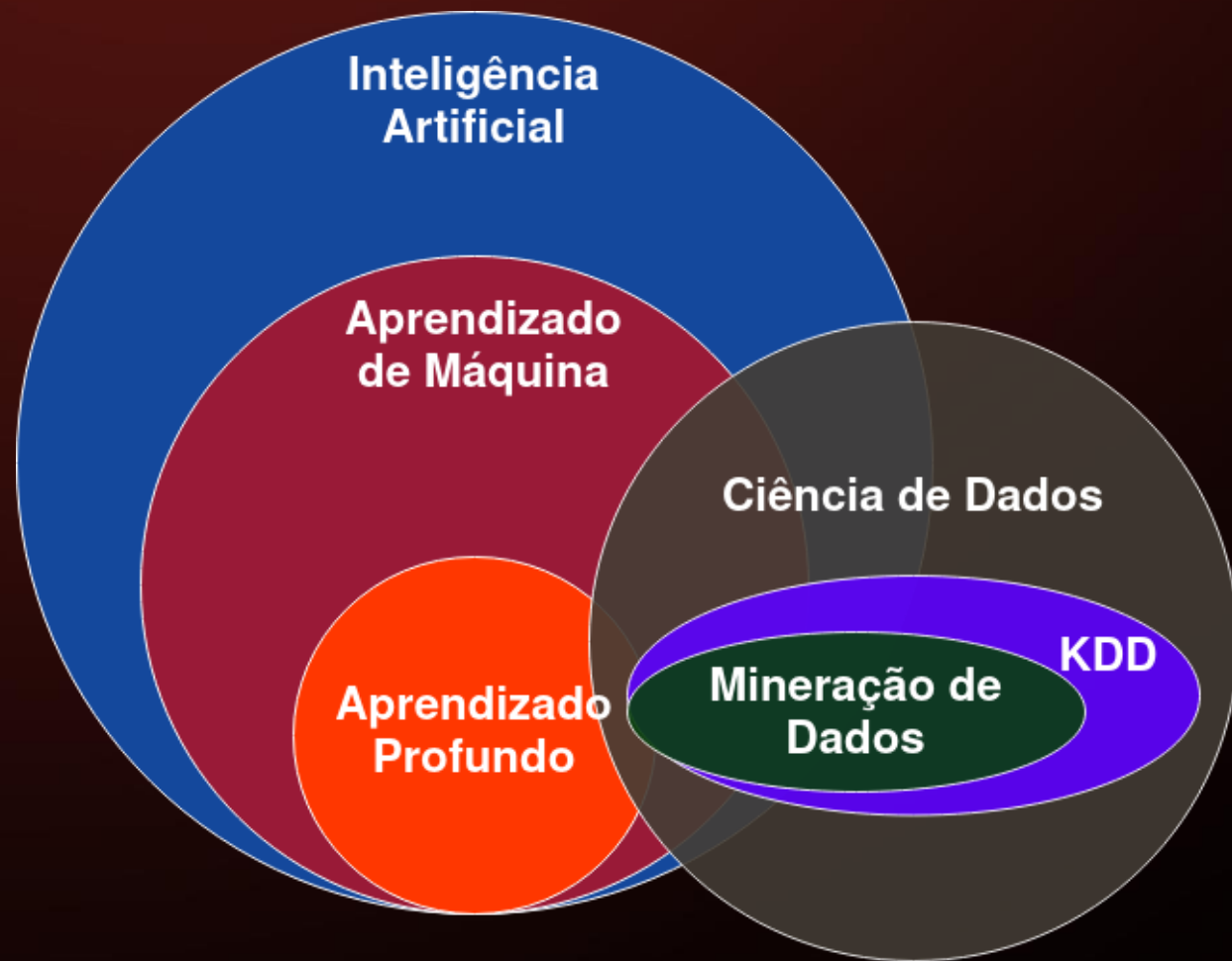


Identificar
conhecimentos

Avaliar resultados

Referência: Schneider [1]

TERMINOLOGIA



TERMINOLOGIA

- **KDD:**
 - **Processo de geração de conhecimento**
- **Mineração de Dados:**
 - **Técnicas para identificar relações ou padrões nos dados**
 - **Parte do KDD**

TERMINOLOGIA

- **Inteligência Artificial:**
 - Criação de programas inteligentes
 - Inspiração na biologia
 - Sistemas especialistas
- **Aprendizado de Máquina:**
 - Programas capazes de melhorar por experiência
 - Métodos usados em mineração de dados

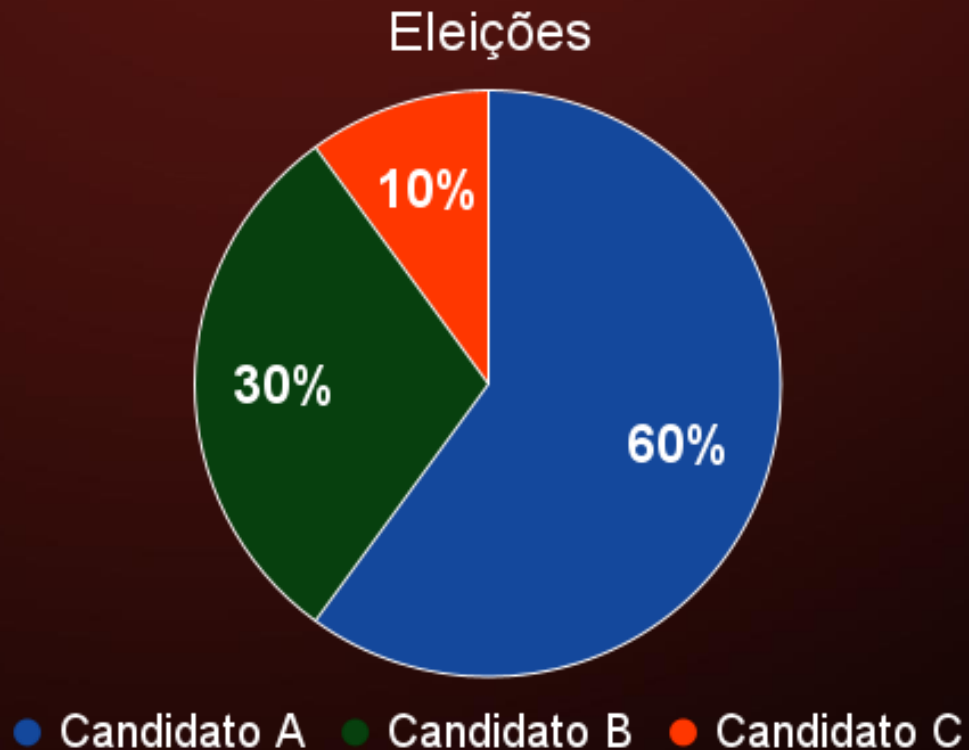
TAREFAS DA MINERAÇÃO DE DADOS

- **Tipos de informação:**
 - **Descritivas:**
 - **Propriedades dos dados**
 - **Conhecimento sobre os dados**
 - **Preditivas:**
 - **Inferências**
 - **Relações**
 - **Geração de conhecimento**

PRINCIPAIS TAREFAS

- **Análise descritiva de dados**
- **Predição: classificação e estimação**
- **Análise de grupos**
- **Associação**
- **Detecção de anomalias**

TAREFAS: ANÁLISE DESCRITIVA DE DADOS



- Entender detalhes da pesquisa:
 - Idade, região, classe social
 - Média, desvio padrão

TAREFAS: PREDIÇÃO

CLASSIFICAÇÃO / ESTIMAÇÃO

- **Análise de sentimentos:**
 - **Postagens em redes sociais**
 - **Classificar sentimentos dos usuários**



TAREFAS: AGRUPAMENTO

- **Análise de crédito**
 - **Identificar perfis de clientes:**
 - **Renda**
 - **Idade**
 - **Histórico de pagamentos**



Fonte: johnhain (pixabay)

TAREFAS: ASSOCIAÇÃO

- Comércio eletrônico
 - Identificar produtos relacionados:
 - Frequência de compra conjunta
 - Promoções
 - Sugestões



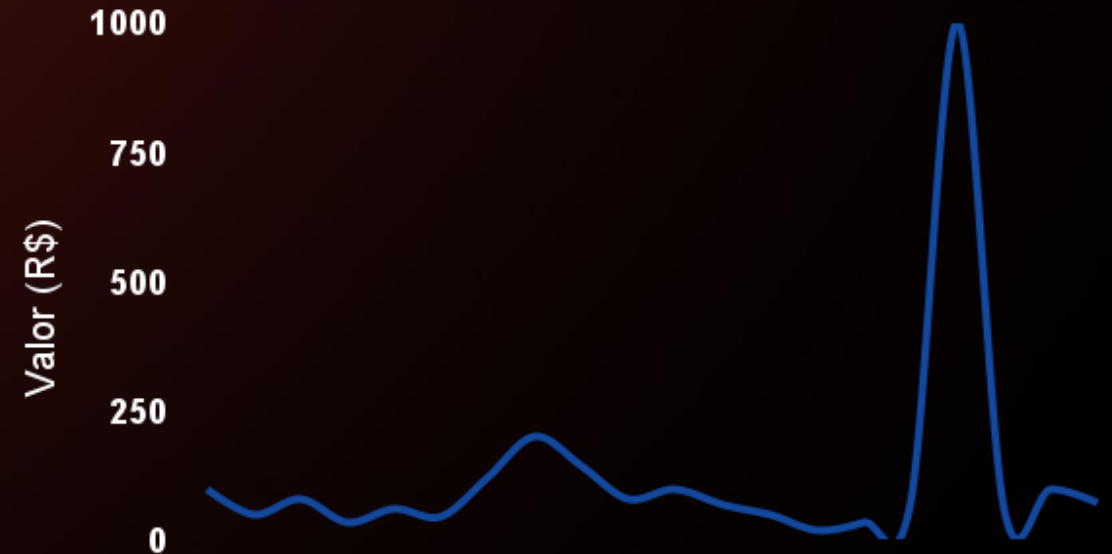
Fonte: kandlaker (pixabay)

TAREFAS: DETECÇÃO DE ANOMALIAS

- **Fraudes em cartão de crédito**
 - Comportamento de consumo do cliente:
 - Horários
 - Valores
 - Local

Resultado: avisos, bloqueio

Padrão de compras



REFERÊNCIAS

Introdução à Mineração de Dados: Conceitos Básicos, Algoritmos e Aplicações: Cap. 1: Introdução à Mineração de Dados. Leandro Nunes de Castro e Daniel Gomes Ferrari. Editora Saraiva, 2016.

Introdução a Big Data e Internet das Coisas (IoT): Descoberta de Conhecimento com Big Data. Izabelly Soares de Moraes et al. Sagah, 2018.

[1] SCHNEIDER, L. F. Aplicação do processo de descoberta de conhecimento em dados do poder judiciário do estado do Rio Grande do Sul. 2003. 103 f. Dissertação (Mestrado em Informática) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003. Disponível em: <<https://lume.ufrgs.br/handle/10183/8968>>. Acesso em: 4 jan. 2019.