

# **ECLAT e Regras de Associação**

**Autor:** Herick Brito  
Juary Junior  
Lucas Lan

# Agenda

O objetivo dessa apresentação é .... demonstrar a aplicação do algoritmo ECLAT na identificação de padrões de compra no setor de vestuário, mostrando o processo, os resultados e as possíveis aplicações práticas.

## **1. Contexto. e Objetivos do Projeto :**

Identificar padrões de compra no setor de vestuário

## **2. Tratamento e Preparação dos Dados :**

Limpeza e preparação dos dados

## **3. Algoritmo ECLAT – Funcionamento e Métricas :**

Aplicação do ECLAT com suporte mínimo e métricas

## **4. Estrutura de Implementação e Métodos ::**

Etapas de execução em python

## **5. Regras de associação e sistema de recomendação:**

Principais combinações encontradas

## **6. Aplicações Práticas e Resultados:**

Padrões úteis e conclusões do estudo

# Contexto e Objetivos



- Contexto do Projeto
  - • Dataset de transações de vendas de vestuário
  - • Múltiplas categorias de produtos (19 categorias)
  - • Necessidade de identificar padrões de compra
- Objetivos
  - • Descobrir produtos comprados frequentemente juntos
  - • Gerar regras de associação para recomendação
  - • Auxiliar estratégias de marketing e layout de loja

# Tratamento e Preparação dos Dados

- **Classe: LimpadorVestuarios**
- 1. Normalização de Texto
  - • Remoção de acentos e caracteres especiais
  - • Padronização para minúsculas
  - • Filtro de marcas e termos irrelevantes
- 2. Categorização de Produtos
  - • 19 categorias principais mapeadas
  - • Conversão de descrições para categorias estruturadas





# Algoritmo ECLAT

- ECLAT - Equivalence Class Clustering and bottom-up Lattice Traversal
- Como Funciona?
  - **1. Cria TID-list: mapeia cada item aos IDs de transações**
  - **2. Busca recursiva: combina itemsets por interseção de TIDs**
  - 3. Filtra por suporte mínimo (0.1% neste projeto)
  - 4. Gera itemsets de tamanho crescente (até 3 itens)
- **Vantagens: Mais eficiente que Apriori em grandes volumes**

# Métricas e Parâmetros

- Parâmetros Configurados
  - • Suporte Mínimo: 0.1% (0.001)
    - Frequência mínima de aparição do itemset
  - • Confiança Mínima: 40% (0.4)
    - Probabilidade de Y ocorrer dado que X ocorreu
  - • Lift Mínimo: 1.1
    - Indica associação positiva ( $> 1$  = correlação)

# Estrutura de Implementação

- Arquitetura Modular
  -  tratamento\_dados.py
    - Classe LimpadorVestuarios - normalização e categorização
  -  mineracao\_regras.py
    - Classe MineradorECLAT - algoritmo e geração de regras
  -  avaliacao\_resultados.py
    - Funções de análise e geração de insights
  -  main.py
    - Pipeline completo - análise por faixas de suporte

Mineração de Dados com ECLAT

# Classe MineradorECLAT – Métodos

- **Métodos Principais:**






- • `_gerar_tidlist()`
  - Cria mapeamento item → transações (estrutura vertical)
- • `_explorar_combinacoes()`
  - Busca recursiva de itemsets frequentes
- • `encontrar_itemsets()`
  - Função principal de mineração com suporte mínimo
- • `gerar_regras()`
  - Cria regras de associação a partir dos itemsets
- • `recomendar_itens()`
  - Sistema de recomendação baseado nas regras



# Regras de Associação

- O que são Regras de Associação?
- **Formato: [Antecedente] → [Consequente]**
- Se o cliente compra X, então provavelmente comprará Y
- **Exemplos de Regras Encontradas:**
  - • [camisa] → [short]
  - • [calcinha] → [sutia]
  - • [pijama] → [calcinha]
  - • [short, camisa] → [calca]
- **Cada regra possui: suporte, confiança e lift**




# Análise por Faixas de Suporte

- Estratégia de Classificação
- Itemsets são organizados por faixas de suporte:
  -  0.1% - 0.2%: Padrões menos frequentes
  -  0.2% - 0.3%: Padrões ocasionais
  -  0.3% - 0.4%: Padrões recorrentes
  -  0.4% - 0.5%: Padrões frequentes
  -  0.5%+: Padrões muito frequentes
- Benefício: Identificar oportunidades em diferentes níveis

# Sistema de Recomendação

- Como Funciona o Sistema?
  - 1. Cliente adiciona item(ns) ao carrinho
  - 2. Sistema busca regras onde o carrinho é antecedente
  - **3. Calcula score para cada recomendação:  $\text{lift} \times \text{confiança}$**
  - 4. Retorna top N recomendações ordenadas por score
  
- Exemplo de Uso:
  - **Carrinho: [camisa]**
  - → Recomendações: short (score: 2.8), calça (2.1), regata (1.9)
  
  - **Carrinho: [calcinha]**
  - → Recomendações: sutia (score: 3.2), pijama (2.5), meia (1.8)

# Aplicações Práticas

- Onde utilizar os resultados?
-  E-commerce
  - • Seção 'Quem comprou isso também comprou...'
  - • Recomendações personalizadas no carrinho
  - • Cross-selling automático
-  Layout de Loja Física
  - • Posicionar produtos relacionados próximos
  - • Organizar prateleiras por padrões de compra
-  Marketing e Vendas
  - • Promoções cruzadas (compre X, ganhe desconto em Y)
  - • Kits e combos baseados em dados reais
  - • Campanhas segmentadas por perfil de compra

# Fluxo de Execução Completo

- Pipeline do main.py:

- **1** Carregamento e Limpeza
  - LimpadorVestuarios processa o CSV
- **2** Mineração ECLAT
  - MineradorECLAT busca itemsets frequentes
- **3** Geração de Regras de Associação
  - Extrai regras com métricas (suporte, confiança, lift)
- **4** Análise por Faixas de Suporte
  - Classifica itemsets em faixas (0.1% a 0.5%+)
- **5** Relatórios e Insights
  - Exibe resultados e gera recomendações

Mineração de Dados com ECLAT

# Resultados Esperados

- O que o sistema entrega?
- ✓ Itemsets Frequentes
  - Conjuntos de produtos comprados juntos, organizados por
  - faixas de suporte (0.1% a 0.5%+)
- ✓ Regras de Associação
  - Relações do tipo 'se A, então B' com métricas completas
  - (suporte, confiança e lift)
- ✓ Sistema de Recomendação Funcional
  - Top N sugestões baseadas no carrinho atual, com scores
- ✓ Insights de Negócio
  - Distribuição de itemsets e análise estratificada

# Resultado final no terminal

```

🚩 Itens/categorias mais frequentes (por ocorrência):
- camisa: 2069 ocorrências (14.36% do total)
- meia: 1481 ocorrências (10.28% do total)
- calca: 1463 ocorrências (10.15% do total)
- short: 1368 ocorrências (9.49% do total)
- calcinha: 1188 ocorrências (8.24% do total)

Total de produtos (instâncias) contabilizadas: 14413

antecedente,consequente,suporte,confianca,lift

Com base no maior lift, regra mais forte
"('roupa_cama', 'toalha'),('fralda_pano'),",0.0030695023021267264,0.8235294117647058,18.4123414
07151096
"('regata', 'roupa_cama'),('fralda_pano'),",0.00131550098662574,0.7499999999999999,16.768382352941174
"('macacao', 'roupa_cama'),('fralda_pano'),",0.0015347511510633632,0.7,15.650490196078431
"('body', 'roupa_cama'),('fralda_pano'),",0.0010962508221881166,0.625,13.973651960784315
"('fralda_pano', 'mijao'),('body'),",0.0028502521376891033,0.4482758620689655,13.276533811016568
"('camisa', 'roupa_cama'),('fralda_pano'),",0.0015347511510633632,0.5833333333333334,13.042075163398694
"('mijao', 'toalha'),('body'),",0.0017540013155009867,0.42105263157894735,12.470266575529731
"('mijao', 'roupa_cama'),('fralda_pano'),",0.0010962508221881166,0.5555555555555556,12.421023965141613
"('body', 'toalha'),('fralda_pano'),",0.0015347511510633632,0.5384615384615384,12.038838612368023
"('regata', 'roupa_cama'),('macacao'),",0.0010962508221881166,0.625,11.977415966386555
Análise concluída com sucesso.

```

Mineração de Dados com ECLAT

# Considerações Finais

- **Resultados Alcançados:**

- ☒ Implementação completa do algoritmo ECLAT
- ☒ Sistema de tratamento de dados robusto
- ☒ Geração automática de regras de associação
- ☒ Sistema de recomendação funcional
- ☒ Análise estratificada por faixas de suporte

- **Próximos Passos:**

- ☐ Integração com plataforma de e-commerce
- ☐ Testes A/B para validar recomendações
- ☐ Otimização de parâmetros (grid search)
- ☐ Análise temporal e sazonalidade



# Referências

- ZAKI, M. J. Scalable algorithms for association mining. IEEE Transactions on Knowledge and Data Engineering, v. 12, n. 3, p. 372-390, 2000.
- HAN, J.; KAMBER, M.; PEI, J. Data Mining: Concepts and Techniques. 3rd ed. Morgan Kaufmann, 2011.
- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. In: Proceedings of VLDB, 1994.
- Dataset: Transações de vendas de vestuário (dataset interno)