

Análise de Dados com Base em Processamento Massivo em Paralelo

DW + Similaridade

Guilherme Muzzi da Rocha
gui.muzzi.rocha@gmail.com



Conteúdo

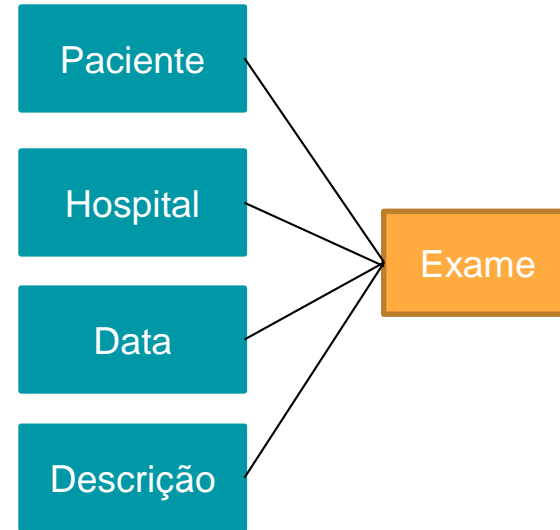
- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

Motivação

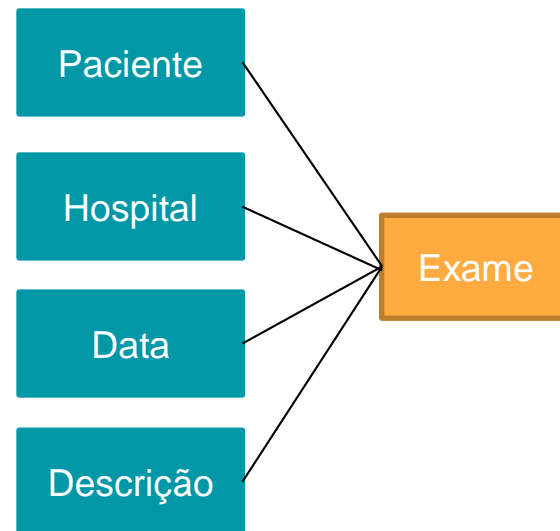
- Consultas OLAP
 - DW - área médica



Motivação

- Consultas OLAP

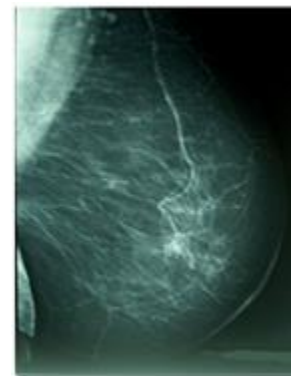
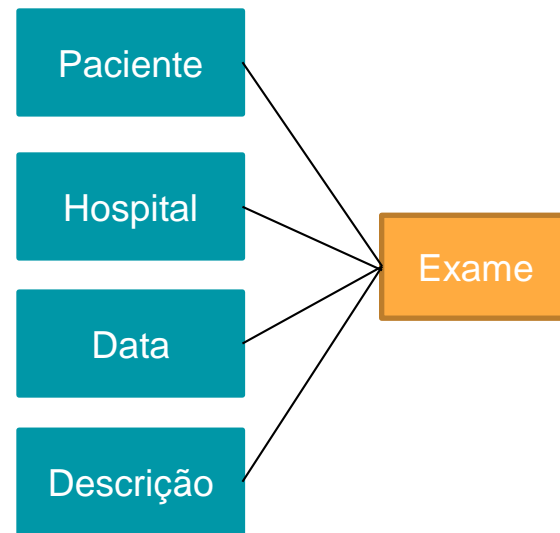
- DW - área médica
- “Quantos exames com suspeita de **câncer de mama** foram realizados em pacientes do **sexo feminino** no último **outubro rosa**?”



Motivação

- Consultas OLAP

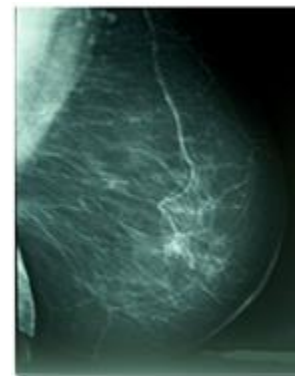
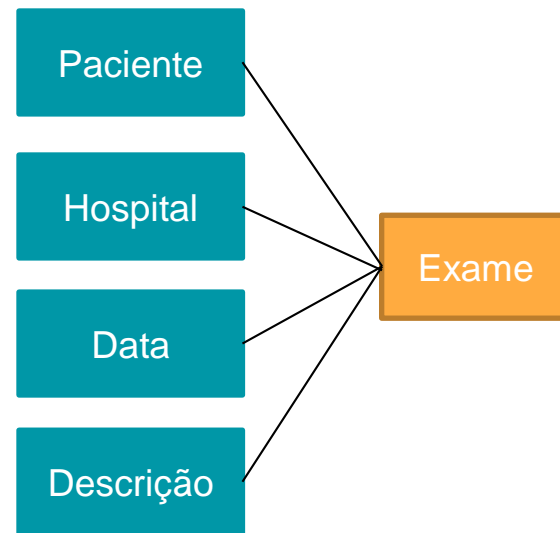
- DW - área médica
- “Quantos exames com suspeita de **câncer de mama** foram realizados em pacientes do **sexo feminino** no último **outubro rosa**?”



Motivação

- Consultas OLAP

- DW - área médica
- “Quantos exames com suspeita de **câncer de mama** foram realizados em pacientes do **sexo feminino** no último **outubro rosa**?”
- Quantos exames com suspeita de **câncer de mama** e que foram realizados no último **outubro rosa** são **SIMILARES** à determinado exame considerando pacientes do **sexo feminino**?

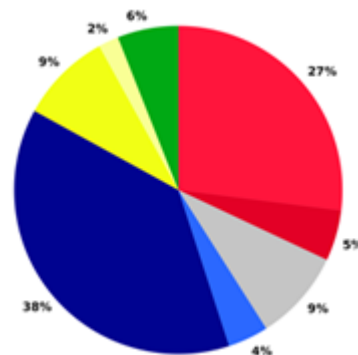


Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

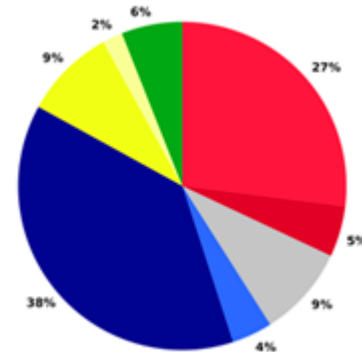
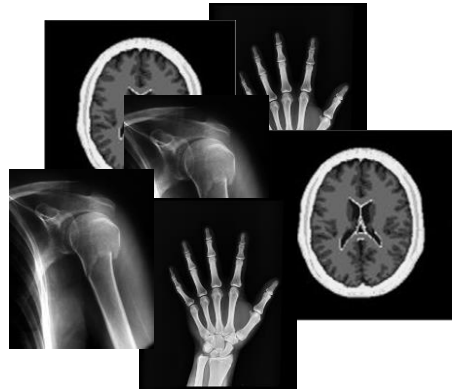
Similaridade

- Objetos complexos



Similaridade

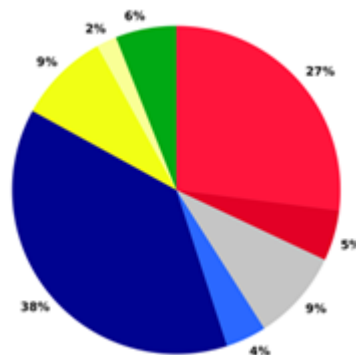
- Objetos complexos



Cor, textura, forma

Similaridade

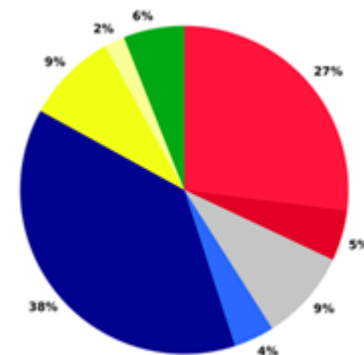
- Objetos complexos



Lat, Long

Similaridade

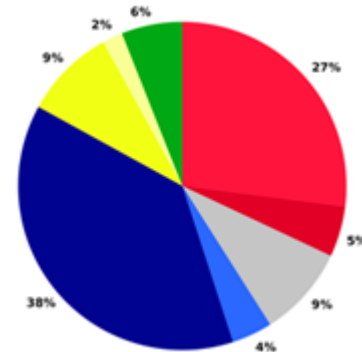
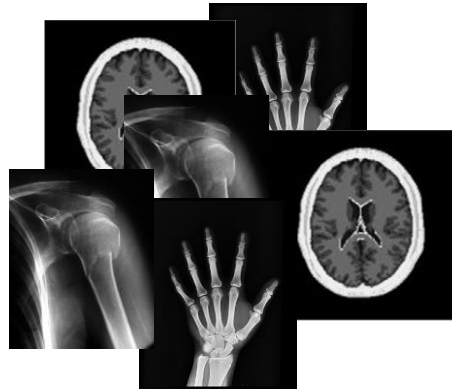
- Objetos complexos



Faixa etária

Similaridade

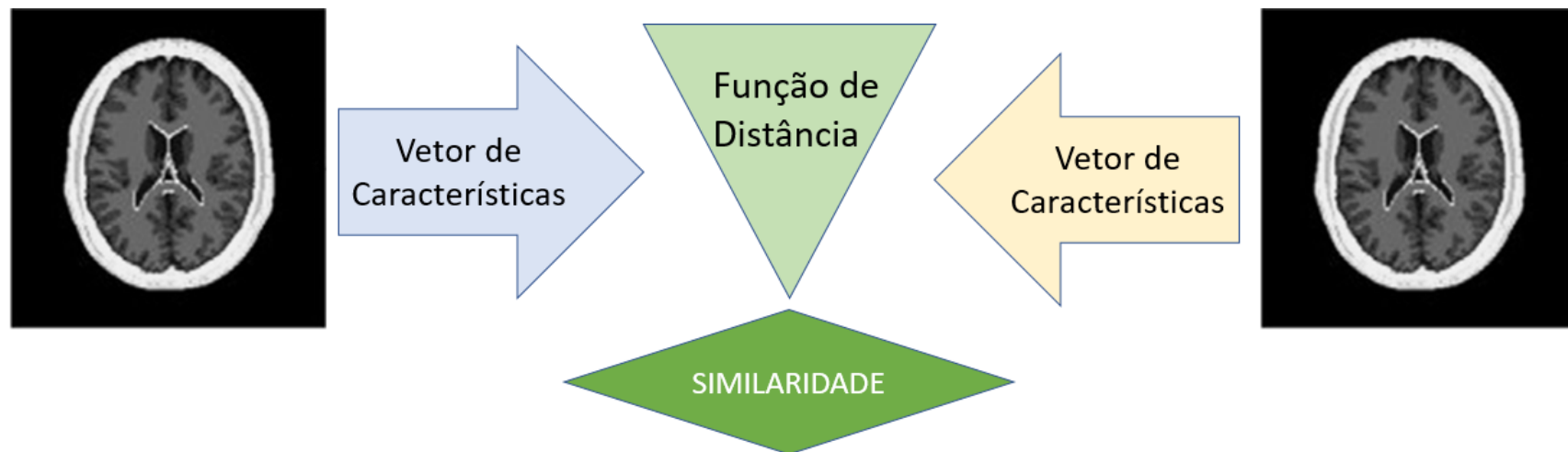
- Objetos complexos



Energia, valência

Similaridade

- Similaridade entre dois objetos complexos



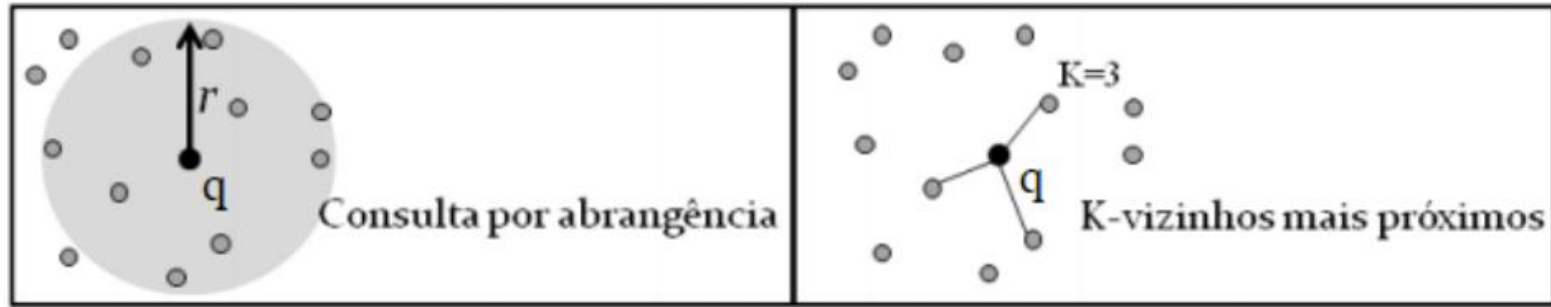
Similaridade

- Fluxo de definições

Conceitos	Definido por	Ex: CBIR Médico
Extração de características	Especialistas da aplicação	Médico
Função de distância	Especialistas do domínio de dados	Profissional de processamento de imagens
Ambiente para executar buscas por similaridade	Especialistas em recuperação de dados	Engenheiro de dados

Similaridade

- Operações de similaridade



Similaridade

- Prática

- Dataset

- Características de 2017 músicas disponibilizadas pelo spotify (Kaggle)

- Energia, “dançabilidade”, “ao vivo”, instrumental, valência, ...



Similaridade

- Prática

- Dataset

- Características de 2017 músicas disponibilizadas pelo spotify (Kaggle)
 - Energia, “dançabilidade”, “ao vivo”, instrumental, valência, ...
 - ***Quais músicas são mais similares à determinada música?***
 - ***Quais músicas possuem o máximo/mínimo de energia e “dançabilidade”?***



Similaridade

- Prática

- Dataset

- Características de 2017 músicas disponibilizadas pelo spotify (Kaggle)
 - Energia, “dançabilidade”, “ao vivo”, instrumental, valência, ...
 - Quais músicas são mais similares à determinada música?
 - Quais músicas possuem o máximo/mínimo de energia e “dançabilidade”?

- **HANDS ON!!!**



Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

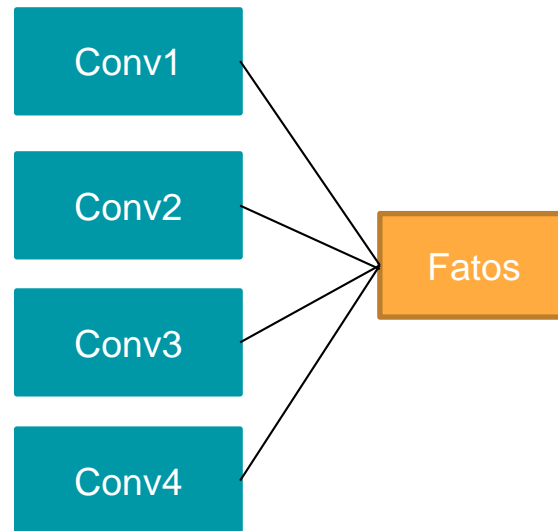
DW + Similaridade

- Tabela de fatos

Fatos

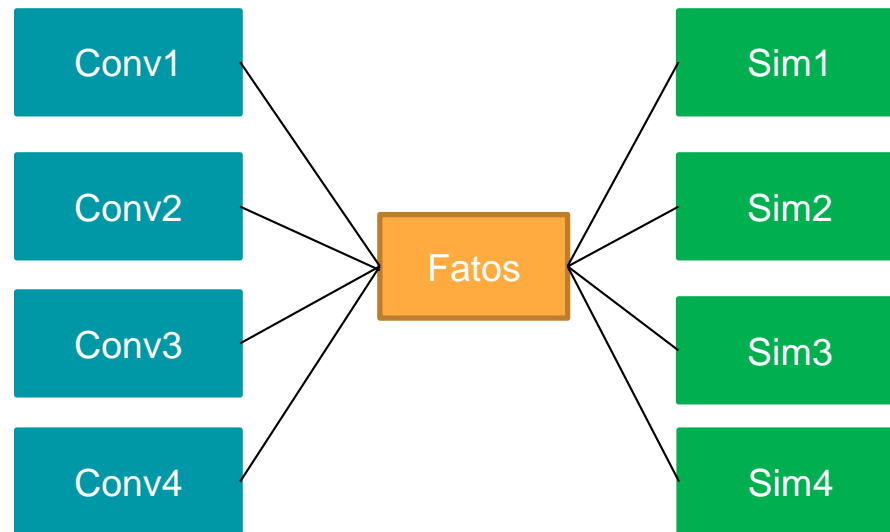
DW + Similaridade

- Tabela de fatos
- Tabelas de dimensão
 - Convencionais



DW + Similaridade

- Tabela de fatos
- Tabelas de dimensão
 - Convencionais
 - Similaridade



DW + Similaridade

- DW de imagens médicas
 - Fatos: **Exame**

Exame

DW + Similaridade

- DW de imagens médicas

- Fatos: **Exame**

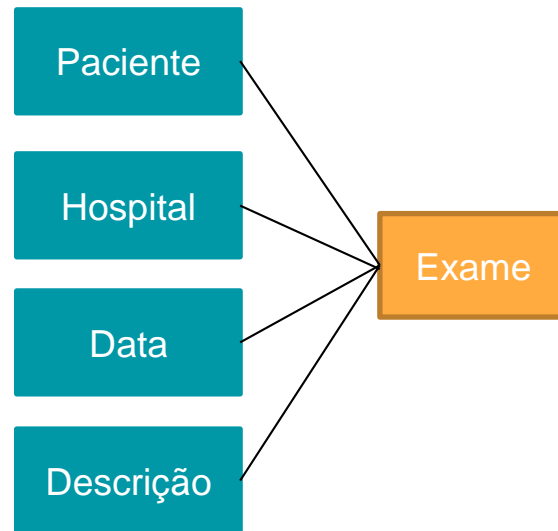
- Convencional

- **Paciente**

- **Hospital**

- **Data**

- **Descrição**



DW + Similaridade

- DW de imagens médicas

- Fatos: **Exame**

- Convencional

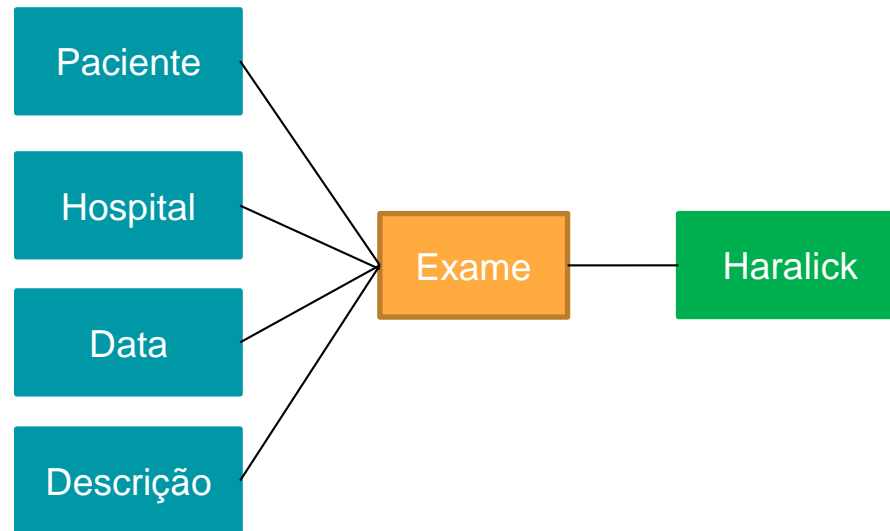
- **Paciente**

- **Hospital**

- **Data**

- **Descrição**

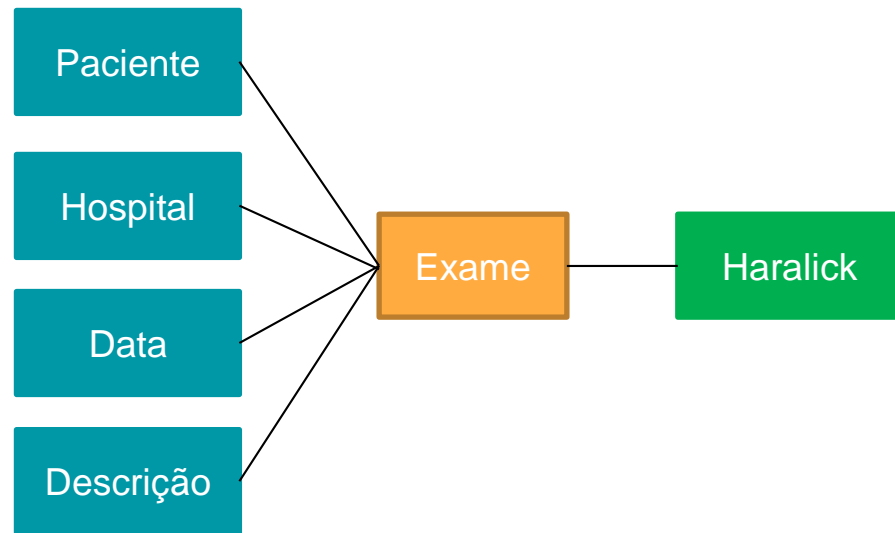
- Similaridade: **Haralick**



DW + Similaridade

- DW de imagens médicas

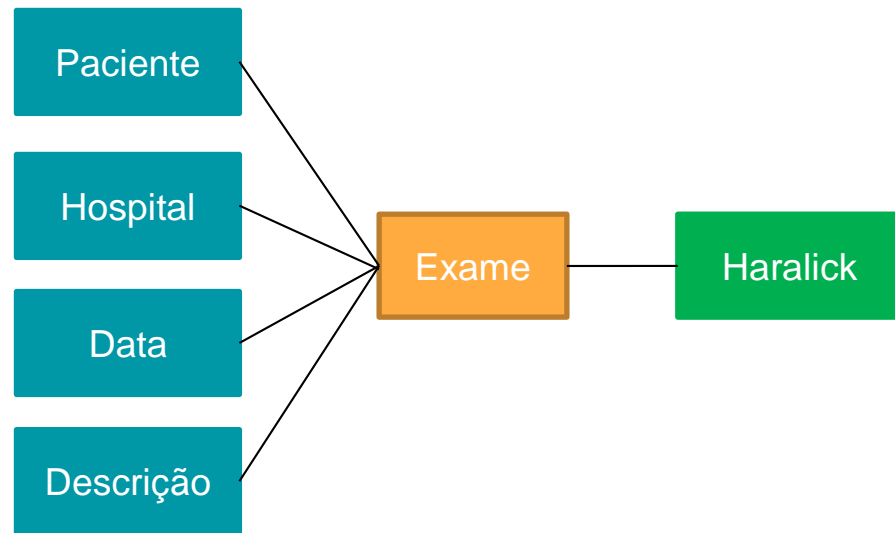
- “Quantos exames com suspeita de **câncer de mama** foram realizados em pacientes do **sexo feminino**?”
- Quantos exames com suspeita de **câncer de mama** são **SIMILARES** à determinado exame considerando pacientes do **sexo feminino**?



DW + Similaridade

- DW de imagens médicas

- “Quantos exames com suspeita de **câncer de mama** foram realizados em pacientes do **sexo feminino**?”
- Quantos exames com suspeita de **câncer de mama** são **SIMILARES** à determinado exame considerando pacientes do **sexo feminino**?



- **HANDS ON**

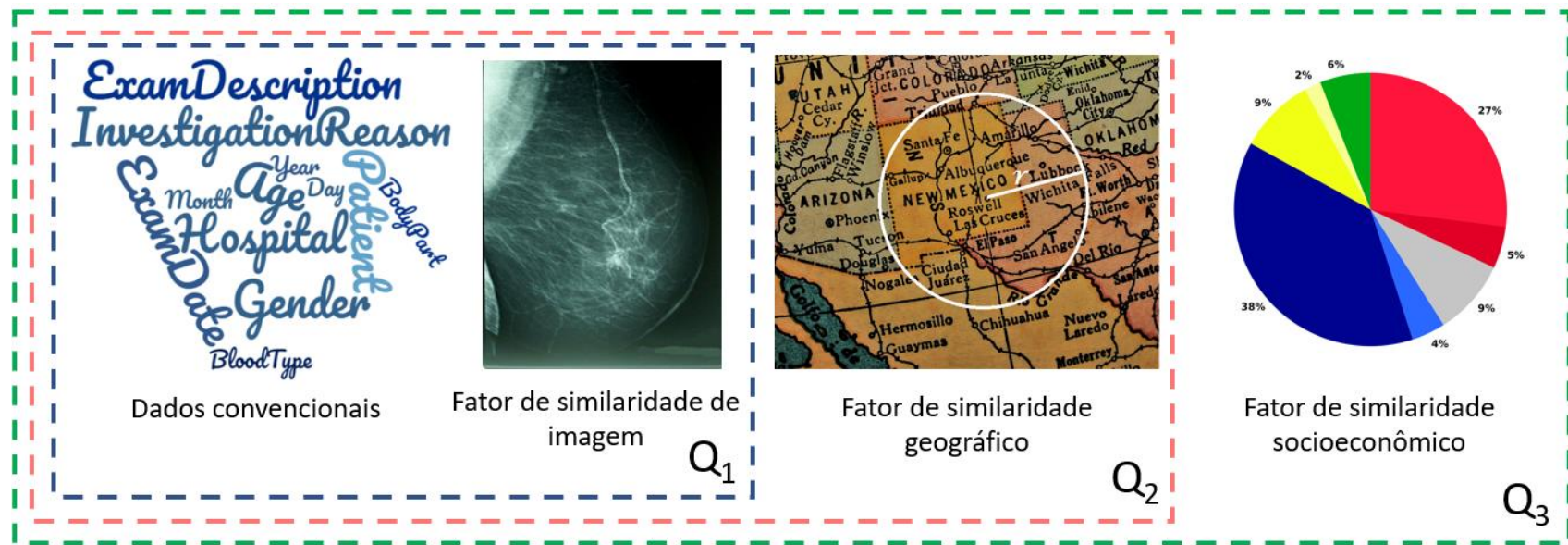


Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

DW + Similaridade

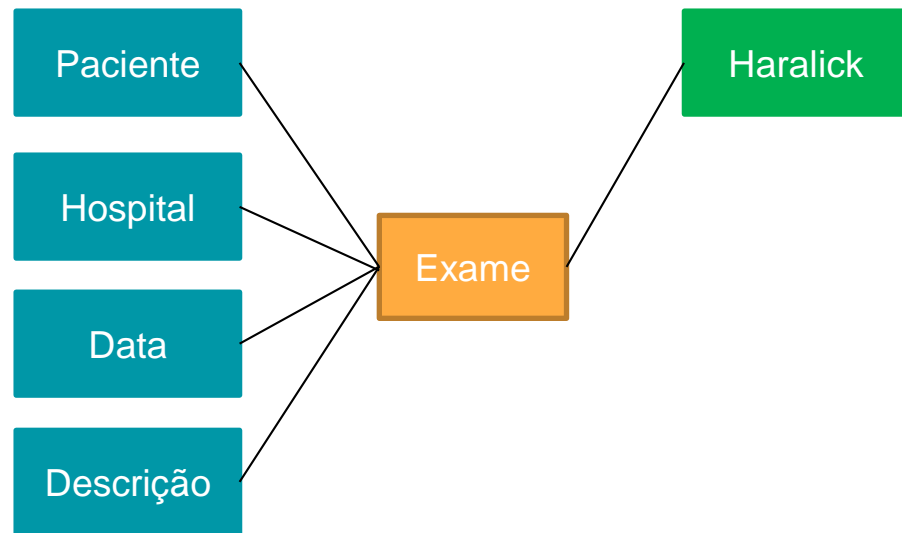
- DW com diversos fatores de similaridade



DW + Similaridade

- DW com diversos fatores de similaridade

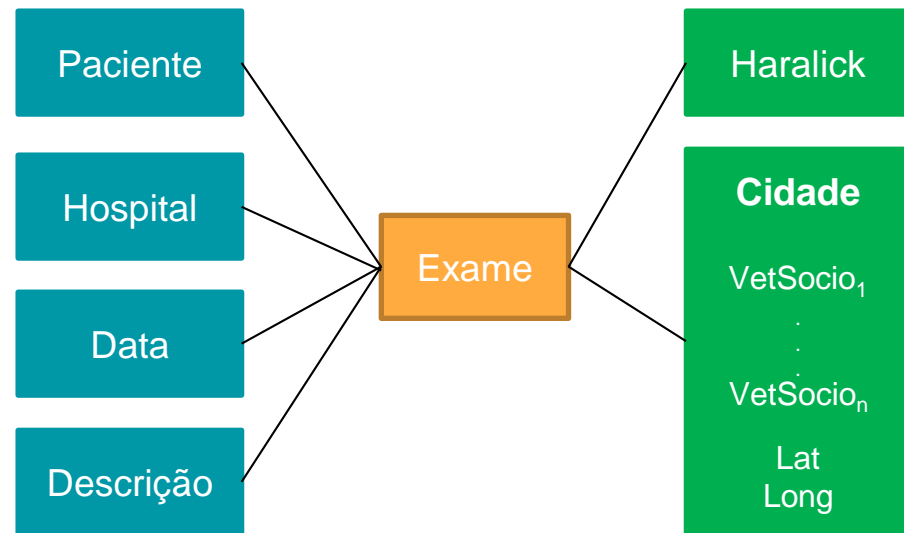
- Fatos
- Convencional
- Similaridade
 - Imagem: **Haralick**



DW + Similaridade

- DW com diversos fatores de similaridade

- Fatos
- Convencional
- Similaridade
 - Imagem: **Haralick**
 - Geográfico: **Lat, long**
 - Socioeconômico: **Censo**

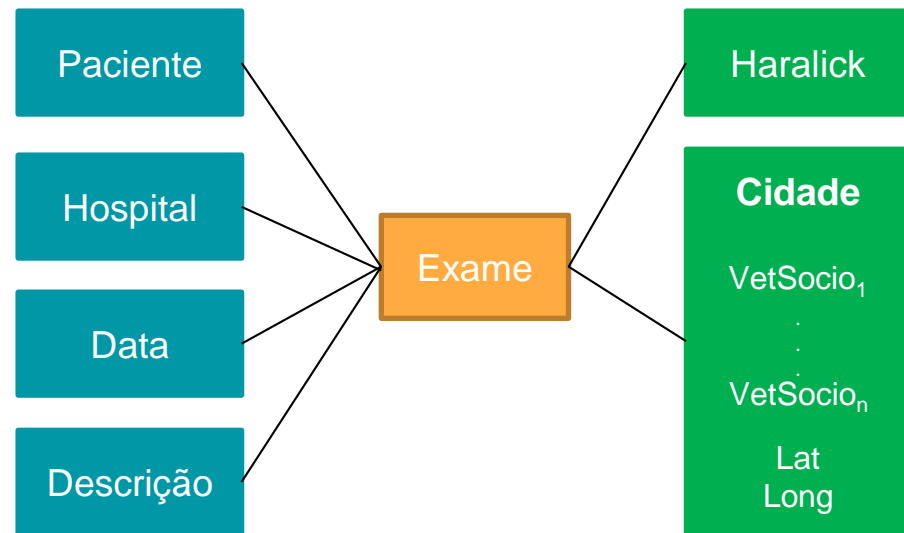


DW + Similaridade

- DW com diversos fatores de similaridade

- Fatos
- Convencional
- Similaridade
 - Imagem: **Haralick**
 - Geográfico: **Lat, long**
 - Socioeconômico: **Censo**

- **HANDS ON**



Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

Conclusão

- Casos de uso
 - Operações dinâmicas
 - Vetor de características
 - Centros de consulta

Conclusão

- Casos de uso
 - Operações dinâmicas
 - Vetor de características
 - Centros de consulta
- Bala de prata?
 - Dificuldades?
 - Estudo de caso

Conclusão

- Casos de uso
 - Operações dinâmicas
 - Vetor de características
 - Centros de consulta
- Bala de prata?
 - Dificuldades?
 - Estudo de caso
- Clusterização x Operações de similaridade

Conteúdo

- Motivação
- Similaridade
 - Conceito
 - Prática – Notebook
- DW + Similaridade
 - DW de imagens
 - Conceito
 - Prática – Notebook
 - DW com diversos fatores de similaridade
 - Fatores de imagem, geográficos e socioeconômicos
 - Prática - Notebook
- Conclusão

DÚVIDAS?