



1

# Data Warehouse e Data Mining

## Sistemas de Apoio à Decisão



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

- **Data Warehouse**

DATAWAREHOUSE (ou Data Warehouse) é um imenso banco de dados que integra diversos outros bancos de dados, daí a sua tradução literária ser “Fábrica” ou “Armazém” de dados.

As diversas informações dos bancos de dados de um Datawarehouse podem servir para dinamizar e trazer muitas vantagens para qualquer empresa, agilizando seus negócios, diminuindo custos e otimizando processos decisórios devido à riqueza de informações que agrega.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

- **Data Warehouse**

Um Datawarehouse pode ser definido também como um super banco de dados que trabalha por traz de outros bancos de dados de ferramentas de Content Management, e-Business, Business Intelligence, Workflow, CRM, Gestão Empresarial etc, integrando informações que são de extrema relevância em processos decisórios e de estratégia dentro de uma empresa.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

- **Data Mining**

Os Data Mining (DM) são as ferramentas de busca de dados em um sistema amplo de banco de dados, de um Datawarehouse. De que adianta uma empresa agregar dados de todas suas sedes, seus departamentos e clientes se não souber buscar as informações necessárias para otimizar seus serviços e agilizar suas mudanças e inovações? Data Mining é a ferramenta, ou são as diversas ferramentas que buscam informações relevantes dentro de um Datawarehouse: estatísticas, números, relatórios e diversos dados necessários para aperfeiçoar diversos processos empresariais.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Características do Datawarehouse

- É um banco de dados que extrai informações de toda a empresa, dos setores de produção, vendas, recursos humanos etc;
- Seus dados são otimizados para extrair informações, e não para processamento de informações;



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Características do Datawarehouse

- Utiliza ferramentas de mining desenvolvidas para buscar as informações mais relevantes, mas que podem também acessar os dados primitivos de outros bancos de dados em caso de necessidade de se aprofundar em algum nível específico de informação, as ferramentas Data Mining devem ser maleáveis para permitir tais consultas;
- Um Datawarehouse contém informações de diversos outros bancos de dados relacionais, e também de arquivos e documentos diversos.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Características peculiares do Datawarehouse

- **Orientado por temas:** armazena informações dentro de temas específicos e de relevância para a empresa, tais como suas atividades, clientes etc;
- **Integrado:** o Datawarehouse uniformiza informações provenientes de diferentes base de dados;
- **Variante no tempo:** o Datawarehouse tem a características de não atualizar seus dados, e sim acumulálos, e sem perder a sua referência de tempo;



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Características peculiares do Datawarehouse

- **Não volátil:** os dados carregados dentro de um Datawarehouse são imutáveis, ficam lá disponíveis para consulta, e não para processos transacionais;
- **Granulidade:** diz respeito ao nível de organização dos dados dentro de um Datawarehouse, compactando os dados e organizando-os por índices, se torna mais fácil o acesso aos dados relevantes e economiza-se espaço de armazenamento, obtendo-se assim uma baixa granulidade;
- **Particionamento dos dados:** é mais uma maneira de se organizar os dados dentro de um Datawarehouse. Dispõem-se os dados mais detalhados em unidades físicas menores, classificando-os por data, área de negócio ou área geográfica, unidades ou outros critérios.





# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Arquitetura do Datawarehouse

Para que um Datawarehouse seja eficiente, não basta agregar informações provenientes de vários sistemas, é preciso filtrar informações relevantes para um processo decisório e estratégico de uma empresa. Isso depende muito da arquitetura do Datawarehouse e das necessidades da empresa em questão, mas, de um modo geral, a arquitetura de um sistema Datawarehouse segue o padrão genérico como veremos a seguir.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Estrutura Genérica de um Datawarehouse

Um Datawarehouse é composto por diversas camadas:

- **Banco de dados operacionais/ fontes externas:** são os banco de dados operacionais e fontes de arquivos externos que irão povoar os dados do Datawarehouse;
- **Acesso à informação:** softwares e hardware necessários para se extrair as informações do Datawarehouse, camada pela qual os usuários interagem por meio das ferramentas de Data Mining;
- **Acesso aos dados:** camada de ligação entre as ferramentas de Data Mining aos dados primitivos do Datawarehouse;



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Estrutura Genérica de um Datawarehouse

- **Metadados:** são os índices, dicionários de dados e outras ferramentas que classificam e descrevem os dados e onde estão armazenados;
- **Gerenciamento de processos:** camada que faz a atualização do Datawarehouse, que tem a função de coletar dados, classificá-los e resumi-los;
- **Transporte:** camada que gerencia o transporte de dados através da rede;
- **Datawarehouse:** local físico e virtual onde estão os dados e informações que compõem do Datawarehouse propriamente dito. Essa camada pode ser, muitas vezes, composta somente de índices e dicionário de dados.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Estrutura Genérica de um Datawarehouse

Expostas suas características e vantagens em comparação aos sistemas de bancos de dados relacionais e particionados, o Datawarehouse se apresenta como a melhor solução para agregar e organizar essas estruturas de dados para as empresas, especialmente as de grande porte e transnacionais. Contudo, para obter vantagem de negócio por meio de uma estrutura de Datawarehouse, é preciso lançar mão de ferramentas capazes de extrair inteligência dessas largas bases de dados integradas. Na ponta dessas ferramentas está o Data Mining a qual abordaremos a seguir.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining Data Mining

DATA MINING (ou Data Mining) significa "garimpagem" automática de dados, recurso ou recursos que buscam informações armazenadas em compartimentos diversos de um sistema e, dependendo do caso, as cruzando, resultando na obtenção de relatórios, freqüentemente com o objetivo de apoiar decisões corporativas. Qualquer sistema de Datawarehouse (DW) só funciona e pode ser utilizado plenamente, com boas ferramentas de exploração. Com o surgimento do Datawarehouse, a tecnologia de Data Mining (mineração de dados) também ganhou a atenção do mercado.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Data Mining

DATA MINING (ou Data Mining) significa "garimpagem" automática de dados, recurso ou recursos que buscam informações armazenadas em compartimentos diversos de um sistema e, dependendo do caso, as cruzando, resultando na obtenção de relatórios, freqüentemente com o objetivo de apoiar decisões corporativas. Qualquer sistema de Datawarehouse (DW) só funciona e pode ser utilizado plenamente, com boas ferramentas de exploração. Com o surgimento do Datawarehouse, a tecnologia de Data Mining (mineração de dados) também ganhou a atenção do mercado.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Data Mining

Como o Datawarehouse possui bases de dados bem organizadas e consolidadas, as ferramentas de Data Mining ganharam grande importância e utilidade. Essa técnica, orientada à mineração de dados, oferece uma poderosa alternativa para as empresas descobrirem novas oportunidades de negócio e, acima de tudo, traçarem novas estratégias para o futuro.



# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Data Mining

O propósito da análise de dados é descobrir previamente características dos dados, sejam relacionamentos, dependências ou tendências desconhecidas. Tais descobertas tornam-se parte da estrutura informacional em que decisões são formadas. Uma típica ferramenta de análise de dados ajuda os usuários finais na definição do problema, na seleção de dados e a iniciar uma apropriada análise para geração da informação, que ajudará a resolver demandas descobertas nessas análises. Em outras palavras, o usuário final reage a um estímulo externo, a descoberta do problema por ele mesmo. Se o usuário falhar na detecção do problema, nenhuma ação é tomada. A premissa do Data Mining é uma argumentação ativa, isto é, em vez do usuário definir o problema, selecionar os dados e as ferramentas para analisar tais dados, as ferramentas do Data Mining pesquisam automaticamente os(...)





# Sistemas de Apoio à Decisão

## Data Warehouse e Data Mining

### Data Mining

(...) mesmos à procura de anomalias e possíveis relacionamentos, assim detectando problemas que não tinham sido identificados pelo usuário. Ou seja, as ferramentas de Data Mining analisam os dados, descobrem problemas ou oportunidades escondidas nos relacionamentos desses dados, então diagnosticam o comportamento de inúmeros negócios, requerendo a mínima intervenção do usuário, que poderá se dedicar somente à busca por conhecimento assim angariando vantagens competitivas.



# Sistemas de Apoio à Decisão

## Data Wirehouse e Data Mining

### Data Mining

Assim compreendido os recursos potenciais da tecnologia, entende-se que as ferramentas de Data Mining são baseadas em algoritmos que formam blocos de inteligência artificial, redes neurais, regras de indução e lógica de predicados, com isso facilitam e auxiliam o trabalho dos analistas de negócio das empresas, ajudando as mesmas a serem mais competitivas e maximizarem seus lucros.



## Referências

- <http://www.pedroom.com.br/portal/vitae/download/cientificos/02%20Datawarehouse%20e%20Datamine.pdf>