



Análise de Dados

Programa

1. Modelos, técnicas e tecnologias de Business Intelligence
2. Estrutura funcional das Bases de dados multidimensionais ;
3. Processo ETL
4. Análise de dados com ferramentas OLAP
5. Ferramentas de modelação.
6. Sistemas de Processamento Analítico;
7. Algoritmos e estruturas de dados para o processamento analítico de dados;
8. Otimização e reestruturação dinâmica de cubos;
9. Sistemas de Apoio à decisão inteligente
10. Monitorização de indicadores, reporting e dashboarding
11. Plataformas de Business Intelligence
12. Vistas



Análise de Dados

Definição

- **O Que é Análise de Dados?**



Análise de Dados

Definição

- A partir de dados estruturados inferir padrões
- Filtrar informação
- Interpretação de dados
-



Análise de Dados

Definição

- Análise de dados pretende transformar dados em informação,
- Dar-lhes significado,
- Conhecimento
- Resolver problemas.



Análise de Dados

Definição

- **Principal motivação**
- Análise em **Big Data**



Análise de Dados

Motivações

- “Explosão” na **disponibilização de dados**;
- **Fontes de dados** abundantes:
- **Solidificação** das Tecnologias de Bases de Dados;
- Disponibilização de **ferramentas automáticas** de procura e arquivo de informação;
- **Armazenamento digital** de informação promove aumento significativo na quantidade de dados disponíveis;
- **Custo do arquivo** de informação diminui drasticamente.



Análise de Dados

Motivação

- Qualquer organização precisa utilizar toda informação disponível para criar e manter vantagem competitiva.
- Conseguir tomar decisões corretas e rápidas.
- Recurso a Sistemas de suporte à decisão (**Decision Support Systems - DSS**);
 - ferramentas para análise,
 - front-ends que suportem consultas ad hoc,
 - interfaces gráficas apropriadas, etc.
- Um **data warehouse** tem integrar os dados internos e externos de uma organização numa estrutura única permitindo uma melhor utilização, aumentando a sua capacidade de resposta e adaptação.



Análise de Dados

Motivação

- A integração de sistemas
 - **OLAP (On-Line Analytical Processing)** e;
 - **Data mining.**
- Pode ajudar as organizações a descobrir novas formas de competir numa economia globalizada,
 - Potenciando novos e melhores produtos e/ou serviços,
 - Mais rápida do que a concorrência,
 - Sem aumentar o custo dos produtos e/ou serviços.
- Adaptada às características e às expectativas de cada organização;
- Ajudar a descobrir maneiras diferentes de atuar no Mercado;
- Ajudar organização mudar para atender as novas realidades.



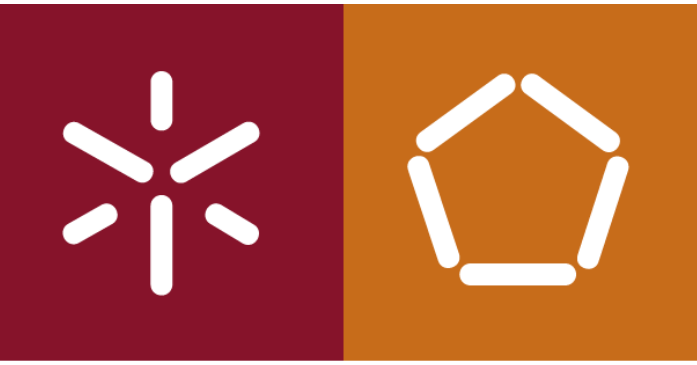
Análise de Dados

Definição

- **Quais são os principais tipos de análise de dados?**

Vamos ter diferentes técnicas e modelos que se destacam pela potencialidade de trazer bons resultados e fáceis de usar.

Vamos considerar duas camadas de análise que ajudam a diferenciar o **Business Intelligence** e o **Analytics**.



Análise de Dados

Definição

A primeira (tradicional), abrange avaliações descritivas, que indicam o que já ocorreu sem emitir julgamentos.

Envolvem análises diagnósticas, que derivam das descritivas. O objetivo é apresentar o que motivou determinado evento a partir da relação entre duas ou mais variáveis.

Nestes casos, são construídos indicadores e é adotado um julgamento de valor.



Análise de Dados

Definição

A **segunda**, considera diferentes tipos de análise prescritivos e preditivos.

São estes os tipos que definem o **Analytics** por serem mais aprofundados.

Isto não significa que as avaliações descritiva e diagnóstica sejam deixadas para segundo plano: as leituras são complementares e permitem compreender melhor os cenários.



Análise de Dados

Definição

Principais tipos de análise existentes

1. Análise preditiva
2. Análise prescritiva
3. Análise descritiva
4. Análise diagnóstica



Análise de Dados

Definição

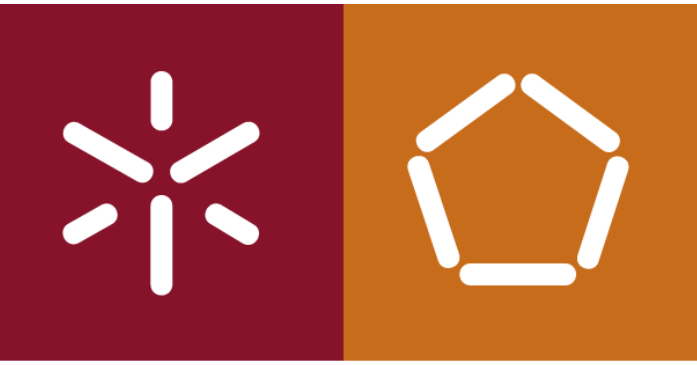
Análise preditiva

Este é o modelo mais conhecido, pois ajuda a prever cenários futuros com base na análise de padrões da base de dados.

Ajuda a tomar decisões mais precisas.

Os métodos usados pela análise preditiva são dados estatísticos e históricos, podem também ser usadas técnicas de datamining e de IA.

- É indicada para projetar comportamentos futuros,
- Avaliar flutuações da economia e
- Tendências de consumo.



Análise de Dados

Definição

Análise prescritiva

Verifica as consequências das ações tomadas,
Saber o que deverá ocorrer ao escolher determinadas atitudes.

Essa camada é a que possui mais valor, pois precisa do elemento humano para se concretizar.

É relevante porque define o caminho a ser tomado para que a ação ocorra conforme o esperado.

Definida a meta, são indicados os caminhos que devem ser percorridos para alcançá-la.



Análise de Dados

Definição

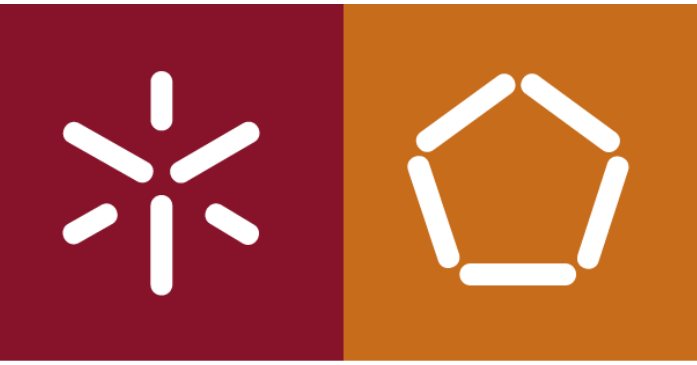
Análise prescritiva

Por isso, a análise prescritiva é considerada a mais complexa,

O profissional deve conhecer técnicas de ciência de dados e ser um especialista no negócio e no ecossistema em que se insere.

Apesar de sua importância, muitas empresas ainda não utilizam a análise prescritiva.

Na saúde, podemos delinear padrões de doenças e verificar como cada atitude impactará sobre esses pacientes.



Análise de Dados

Definição

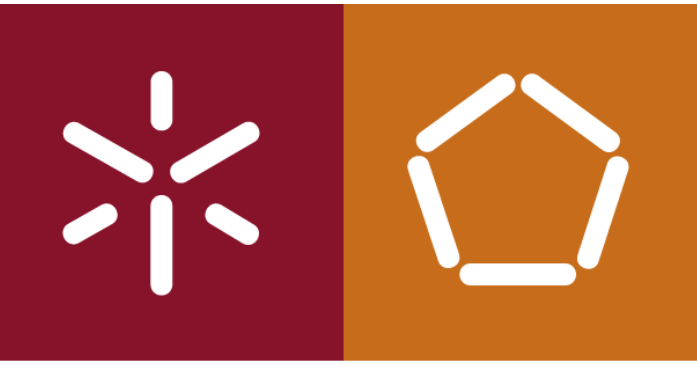
Análise descritiva

O objetivo deste modelo é permitir a análise de eventos em tempo real.

É muito utilizado em situações de análise de crédito. Neste caso, o banco avalia as informações do indivíduo e confere o risco envolvido no processo. Definindo assim por exemplo uma taxa de juro.

Como a análise descritiva não emite julgamento de valor, ela é indicada para visualizar dados e entender o seu impacto no presente, sem os relacionar com o passado ou o futuro.

Ajuda a tomar decisões imediatas com tranquilidade e segurança.



Análise de Dados

Definição

Análise diagnóstica

Aqui a finalidade é compreender as causas de um evento, responder às perguntas:

- Quem?
- Quando?
- Onde?
- Como?
- Por quê?

Analisar o impacto e alcance de uma ação.

A partir disso, podemos definir estratégias para melhorar resultados.

É um modelo muito utilizado em vendas e deve ser complementado com a análise preditiva para reforçar a projeção dos dados.



Análise de Dados

Definição

Para que servem os tipos de análises de dados?

As análises de dados preditivas são as mais utilizadas.

Indica por exemplo as informações demográficas dos consumidores e permite que a empresa saiba exatamente que produto um cliente procura.

A análise preditiva é usada pela Netflix, para perceber que tipos de programas deve recomendar aos seus assinantes.

Com técnicas de Big Data e Machine Learning a Netflix consegue perceber comportamentos e sugerir a série ou o filme indicado para cada cliente.

A análise prescritiva é útil para verificar a eficiência de processos. A Google, usa esta técnica para saber que websites apresentados numa pesquisa são relevantes para os utilizadores e corrigir os resultados apresentados para determinadas palavras-chave.



Análise de Dados

Definição

Para que servem os tipos de análises de dados?

As análises descritivas são muito utilizadas pelos operadores de cartões de crédito. Quando por exemplo um cliente efetua uma operação que ultrapassa o limite de crédito, a operação é autorizada de acordo com seu perfil de consumo e a assiduidade com que realiza pagamentos.

A análise descritiva é tão precisa que oferece uma resposta instantânea para esse pedido e aprova (ou não) um pagamento e a autorização de um crédito extra.

As análises diagnósticas, avaliam a dimensão de uma ação realizada num negócio. Ajuda a comparar métricas de marketing e de vendas, relacionando ambas para entender seus efeitos.



Análise de Dados

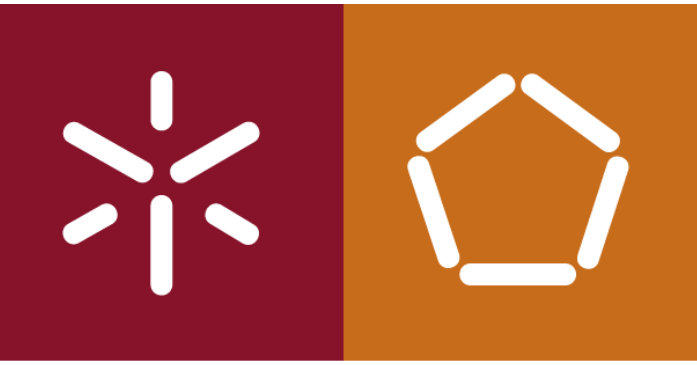
Definição

Como são feitas as análises de dados?

Análises exploratórias

Modelação de dados

Geração de relatórios



Análise de Dados

Definição

Como são feitas as análises de dados?

Análises exploratórias

As análises exploratórias são feitas quando os dados ainda não foram integrados ou podem estar incompletos.

Não são completamente automatizadas porque precisam que alguém verifique pontos fora da curva para que esses dados sejam integrados em sistemas.



Análise de Dados

Definição

Como são feitas as análises de dados?

Modelação de dados

É neste ponto que são definidas as regras para as análises preditivas, descritivas, diagnósticas e prescritivas.



Análise de Dados

Definição

Como são feitas as análises de dados?

Geração de relatórios

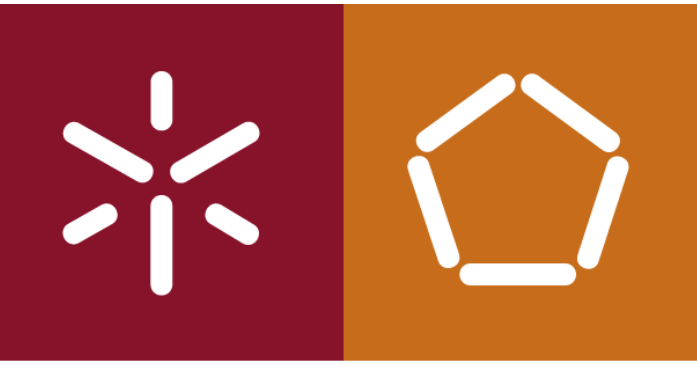


Análise de Dados

Definição

Modelação de dados

O que é Data Warehouse?



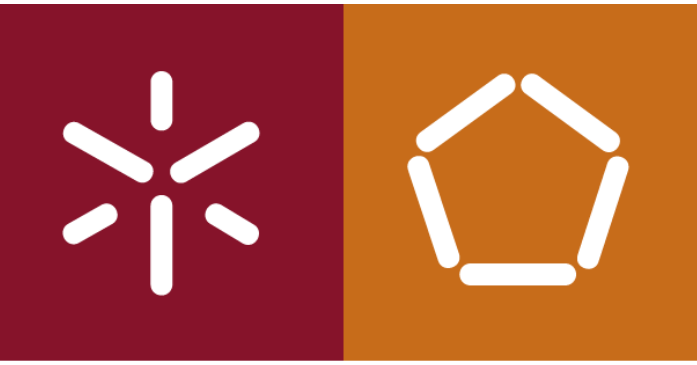
Análise de Dados

Motivação

Data Warehouse

Data Mining

- Um **data warehouse** é uma coleção de dados **orientada por assuntos, integrada, variável no tempo**, que tem por objetivo dar suporte aos processos de tomada de decisão.
- É um base de dados contendo dados extraídos do ambiente de produção da organização, que foram selecionados, depurados e otimizados para processamento de consultas e não para processamento de transações.
- Em geral, um **data warehouse** requer a consolidação de outros recursos de dados além dos armazenados em base de dados relacionais, incluindo informações provenientes de folhas de calculo, documentos, etc.



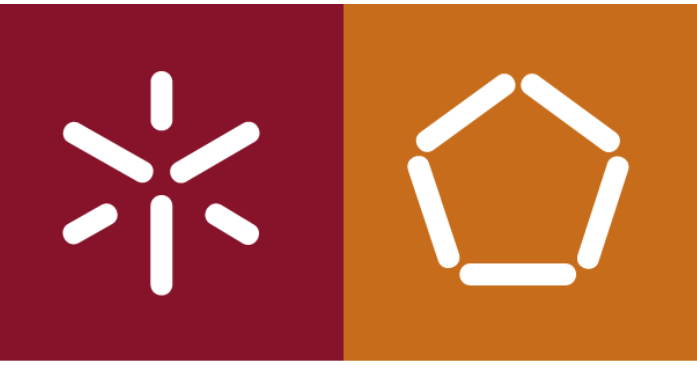
Análise de Dados

Motivação

Data Warehouse

Data Mining

- O objetivo de um data warehouse é fornecer uma "**imagem única da realidade da organização**".
- Sistemas de data warehouse compreendem:
 - programas que extraem dados do ambiente de dados operacionais da organização,
 - uma base de dados que os mantém, e
 - sistemas que fornecem estes dados aos utilizadores.
- Permitem que sistemas mais antigos continuem em operação;
- Consolidam dados inconsistentes dos sistemas mais antigos em conjuntos coerentes;
- Extraem novas informações oriundas das operações correntes;
- Disponibilizam um ambiente para o planeamento e arquitetura de novos sistemas operacionais.



Análise de Dados

Motivação

Data Warehouse

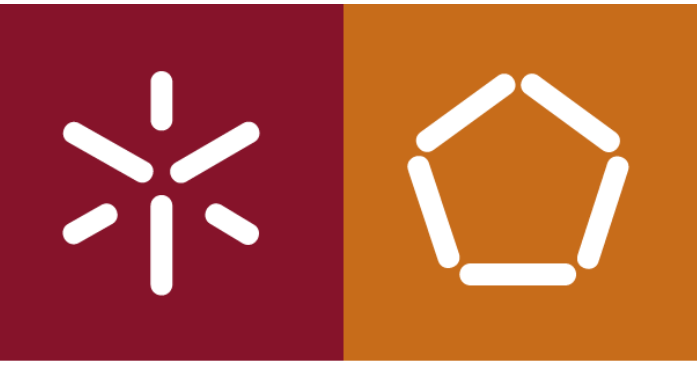
Data Mining

As formas de extração mais comuns no mercado hoje são:

- Ferramentas de consulta e emissão de relatórios;
- EIS (Executive Information Systems);
- Ferramentas OLAP;
- Ferramentas Data mining.

Integração com o ambiente Web, permitindo maior agilidade em consultas estáticas e dinâmicas.

•



Análise de Dados

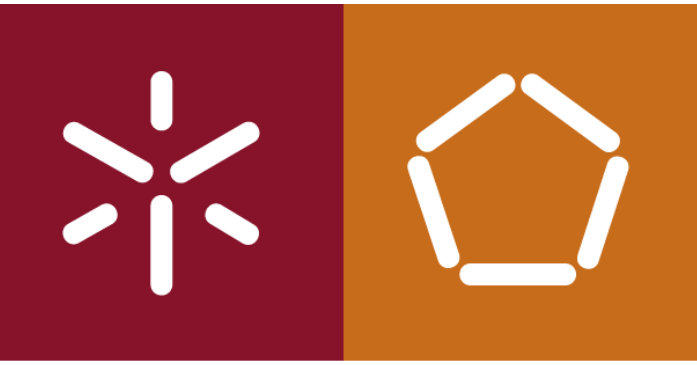
Motivação

Data Warehouse

Data Mining

A diferença básica entre ferramentas **OLAP** e **data mining** está na maneira como a exploração dos dados é abordada.

- **Com ferramentas OLAP a exploração é feita na base da verificação, o analista conhece a questão, elabora uma hipótese e utiliza a ferramenta para confirmá-la.**
- **Com data mining, a questão é total ou parcialmente desconhecida e a ferramenta é utilizada para a busca de conhecimento.**



Análise de Dados

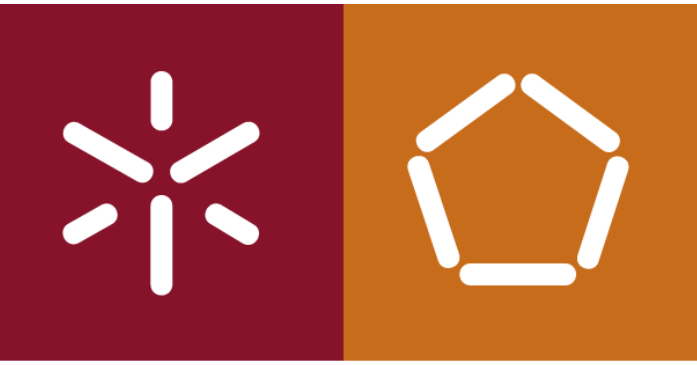
Motivação

Data Warehouse

Data Mining

OLAP (On-Line Analytical Processing)

- Conjunto de tecnologias projetadas para suportar análise e consultas ad hoc.
- Sintetizam informações sobre a organização, através de comparações, visões personalizadas, análise histórica e projeção de dados em vários cenários de "e se...".
- São implementados para ambientes multiutilizador em arquiteturas cliente-servidor e oferece respostas rápidas e consistentes às consultas iterativas executadas, independente do tamanho e complexidade da base de dados.



Análise de Dados

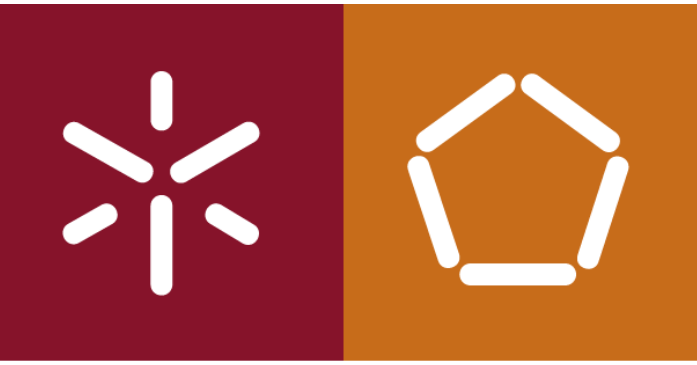
Motivação

Data Warehouse

Data Mining

OLAP (On-Line Analytical Processing)

- A principal característica dos sistemas OLAP é permitir uma visão conceitual multidimensional dos dados de uma organização.
- A visão multidimensional é muito mais útil do que a tradicional visão tabular utilizada nos sistemas de processamento de transação.
- É mais natural, fácil e intuitiva, permitindo a visão em diferentes perspectivas das atividades da organização, facilitando a exploração da informação.
-



Análise de Dados

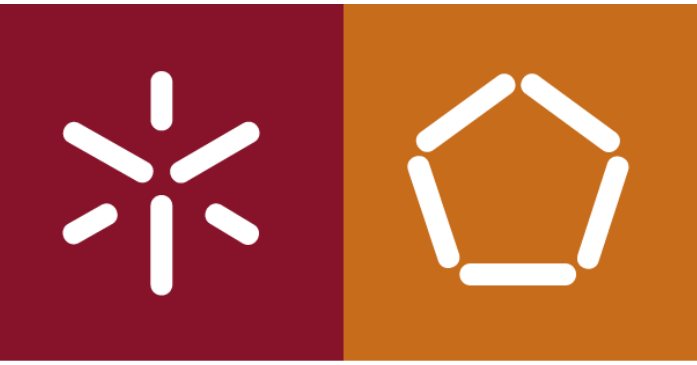
Motivação

Data Warehouse

Data Mining

OLAP (On-Line Analytical Processing)

- A arquitetura OLAP possui três componentes principais:
 - modelo de negócios para análises interativas, uma linguagem gráfica que permita diversas visões e níveis de detalhes dos dados;
 - motor OLAP para processar consultas multidimensionais na procura de um dado-alvo; e
 - mecanismo para armazenar os dados a serem analisados.
- Na base de dados definimos se o pacote é:
 - ROLAP, que faculta um interface com uma base de dados relacional de mercado, ou
 - MOLAP, que se liga a um servidor OLAP, através de uma base de dados multidimensional e dedicada.



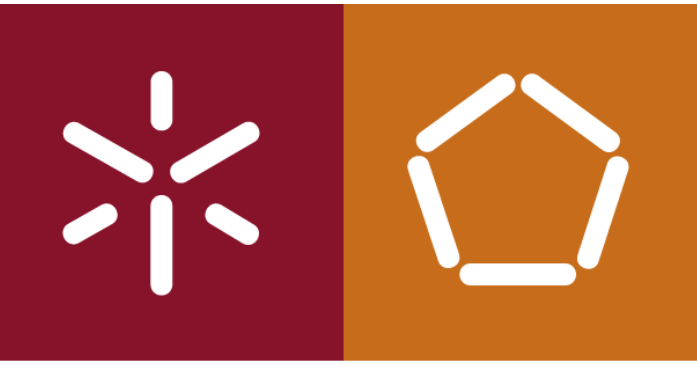
Análise de Dados

Motivação

Data Warehouse

Data Mining

- Quando surgiram os primeiros estudos sobre **data warehouse**, o **data mining** era visto como um subconjunto das atividades associadas ao warehouse.
- Hoje os caminhos do **warehouse** e do **mining** tornaram-se áreas complementares.
- Enquanto o **warehouse** pode ser uma boa fonte de dados para minerar, o **data mining** é reconhecido como uma tarefa genuína, e não mais como uma subtarefa do warehouse.
- O termo data mining ainda provoca certa confusão quanto à sua definição.
- Data mining (**mineração de dados**) é o processo de extrair informação válida, previamente desconhecida e de máxima abrangência a partir de grandes bases de dados.



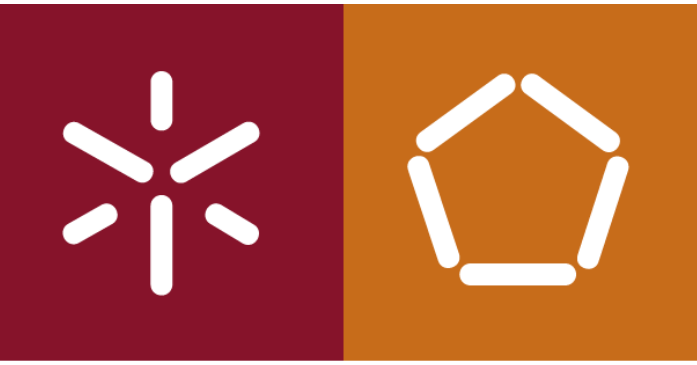
Análise de Dados

Motivação

Data Warehouse

Data Mining

- Data mining é muito mais do que uma simples consulta a uma base de dados, permite aos utilizadores explorar e inferir informação útil, descobrindo relacionamentos escondidos na base de dados.
- Pode ser considerada uma forma de descoberta de conhecimento em bases de dados (**KDD - Knowledge Discovery in Databases**),
- Envolvendo Inteligência Artificial e Bases de Dados.



Análise de Dados

Motivação

Data Warehouse

Data Mining

- Gestão de Produção,
- Gestão Inteligente de Encomendas,
- Gestão de inventários,
- Gestão financeira,
- Análise de risco,
- Gestão de transportes,
- Gestão de frotas,
- Telecomunicação,
- Análise de chamadas,
- Saúde,
- Análise de resultados,
- Marketing,
- Perfil dos consumidores,
- Seguros,
- Detecção de fraudes, etc.



Análise de Dados

Problema

- **“We are drowning in data, but starving for knowledge”**
Data Mining: Concepts and Techniques
Jiawei Han, Micheline Kamber
- **“Necessity is the mother of invention”**
Idem
- **“A necessidade aguça o engenho”**
PT-version



Análise de Dados

Soluções

Business Intelligence

- Data Warehousing Técnicas de OLAP (On-Line Analytical Processing)
 - *Processamento On-line Analítico*, é uma abordagem de fornecimento de respostas rápidas para consultas analíticas de natureza multidimensional. O OLAP faz parte de uma categoria mais abrangente, o **Business Intelligence**, que também inclui **ETL** (Extract, Transform, Load – *Extração, Transformação e Carga*), geração de relatórios relacionais e
- Data Mining/Extração de Conhecimento
 - **Descobrir**, em grandes volumes de dados, sem qualquer formulação prévia de hipóteses, **informações genéricas, relevantes e desconhecidas**, por forma a serem utilizadas em processos de tomada de decisão.