



**FACULDADE DE ADMINISTRAÇÃO E NEGÓCIOS DE  
SERGIPE – FANESE  
CURSO DE PÓS GRADUAÇÃO MBA EM  
ADMINISTRAÇÃO DE BANCO DE DADOS**

**VICTOR MATHEUS DA SILVA**

**Um estudo sobre os componentes do BUSINESS  
INTELLIGENCE e sua importância no planejamento  
estratégico**

**Aracaju – SE  
2016.1**

**VICTOR MATHEUS DA SILVA**

**Um estudo sobre os componentes do BUSINESS INTELLIGENCE e sua importância no planejamento estratégico**

Trabalho de Conclusão de Curso apresentado ao Núcleo de Pós-Graduação e Extensão – NPGE, da Faculdade de Administração de Negócios de Sergipe – FANESE, como requisito para a obtenção do título de Especialista em Administração de Banco de Dados.

---

**Nome completo do Avaliador**

---

**Nome completo do Coordenador de Curso**

---

**Nome completo do Aluno**

**Aprovado (a) com média: \_\_\_\_\_**

**Aracaju (SE), \_\_\_\_ de \_\_\_\_\_ de 2016.**

## SUMÁRIO

|  |           |
|--|-----------|
| <b>RESUMO.....</b>   | <b>4</b>  |
| <b>1 INTRODUÇÃO.....</b>   | <b>5</b>  |
| <b>2 DESENVOLVIMENTO .....</b>                                   | <b>6</b>  |
| <b>2.1 Evolução do Business Intelligence.....</b>                | <b>6</b>  |
| <b>2.2 Componentes do Business Intelligence.....</b>             | <b>7</b>  |
| <b>2.2.1 Data Warehouse.....</b>                                 | <b>8</b>  |
| <b>2.2.2 Data Mart .....</b>                                     | <b>10</b> |
| <b>2.2.3 O processo de ETL .....</b>                             | <b>11</b> |
| <b>2.2.4 OLAP .....</b>  | <b>12</b> |
| <b>2.2.5 Data Mining .....</b>                                   | <b>16</b> |
| <b>2.3 Fatores de Sucesso e Fracasso nos projetos de BI.....</b> | <b>18</b> |
| <b>3 CONCLUSÃO .....</b>   | <b>20</b> |
| <b>4 ABSTRACT.....</b>   | <b>21</b> |
| <b>5 REFERÊNCIAS.....</b>  | <b>22</b> |

## RESUMO

Devido à atual sociedade globalizada e ao dinamismo gerado pelo mercado, o qual vem sofrendo mudanças cada vez mais rápidas e tornando-se cada vez mais exigente, é indispensável a criação de estratégias para que as empresas possam enfrentar a alta concorrência e consigam sobreviver neste mundo competitivo. Para isso, elas devem responder rapidamente às mudanças e aperfeiçoar os seus processos por meio de técnicas de gestão através do Business Intelligence, transformando dados brutos em informação útil para as operações estratégicas e tomada de decisões de modo a trazer ganhos reais para os negócios. A sua implementação, porém, necessita de que a empresa consiga definir fielmente os seus objetivos e o quanto ela está disposta a investir, alinhando o projeto com os interesses estratégicos da corporação. O objetivo deste trabalho reunir estudos sobre os principais componentes do BI explicando cada um deles, além de demonstrar como este poderoso processo atua dentro da organização e a sua importância no planejamento estratégico para as tomadas de decisões, apresentando uma pesquisa qualitativa do tipo bibliográfica.

**PALAVRAS-CHAVE:** Inteligência de Negócios; Planejamento Estratégico; Tomada de Decisão.

## 1. INTRODUÇÃO

Devido à atual sociedade globalizada e ao dinamismo gerado pelo mercado, o qual vem sofrendo mudanças cada vez mais rápidas e com uma alta exigência, torna-se indispensável a criação de estratégias para que as empresas possam enfrentar a alta concorrência e consigam sobreviver neste mundo competitivo. Para isso, as empresas devem responder rapidamente às mudanças e aperfeiçoar os seus processos, por meio de técnicas de gestão.

O *Business Intelligence*, que antes era restrito a pessoas especializadas (pessoal de TI e especialistas em pesquisas de mercado), se expandiu para um número maior de pessoas, graças ao crescimento da Internet, principalmente, com o surgimento do comércio eletrônico. Surge então um novo tipo de cliente, um consumidor que se apresenta virtualmente.

Com o surgimento desse novo mercado consumidor, que cresce espontaneamente, era necessário conhecer e atender às suas necessidades com uma maior agilidade comercial, capacidade de tomada de decisões e refinamento estratégico o mais rápido possível. Todas essas novas necessidades exigiram um maior envolvimento corporativo. O *Business Intelligence*, então, passou a ser visto como uma aplicação estratégica integrada, tornando-se disponível por meio de desktops, servidores da empresa e sistemas web.

Atualmente, todos os tipos de corporações precisam do BI para os mais diversos casos, como por exemplo, para a tomada de decisão, otimização do trabalho e elaboração estratégica. Ele pode ser aplicado em todas as corporações, mas antes é preciso analisar principalmente dois fatores: quanto a empresa está disposta a gastar e o que ela espera obter com a sua implementação. Resumindo, é necessário alinhar o projeto de BI com os interesses estratégicos da corporação.

O papel deste artigo é justamente demonstrar como o BI atua dentro da organização e a sua importância no planejamento estratégico para as tomadas de decisões, sobrevivência e destaque no mercado, reunindo estudos sobre os seus principais componentes e explicando-os, a partir de uma pesquisa qualitativa do tipo bibliográfica.

## 2. REFERENCIAL TEÓRICO

### 2.1 Evolução do *Business Intelligence*

Diferentemente do que a maioria pensa, o conceito de Business Intelligence não é tão recente. Segundo PRIMAK (2008), alguns povos egípcios, por exemplo, já utilizavam deste princípio, utilizando informações vindas da natureza para benefício próprio. Eles observavam e analisavam o comportamento das marés, os períodos de seca e chuva, entre outros, para a tomada de decisões que propiciavam uma ajuda para a melhoria de vida de suas comunidades.

Com o passar do tempo, o mundo mudou, mas a ideia de BI permanece a mesma. Na medida em que o tempo passava, cada vez mais aumentava a necessidade de cruzar informações para a realização de uma gestão estratégica.

Em meados da década de 60, a tecnologia era incerta, se baseando no uso de cartões perfurados, transistores, *mainframes* e linguagens de programação um tanto limitadas. A partir da década de 70, houve grandes avanços tecnológicos, entre eles o surgimento de tecnologias de armazenamento e acesso a dados, chamados DASD (*Direct Access Storage Device*) e SGBD (Sistema Gerenciador de Banco de Dados). Estas tecnologias permitiram a criação de fontes de armazenamento de dados. Neste contexto, surgiram, por exemplo, os sistemas ERPs (*Enterprise Resource Planning* – Planejamento dos Recursos da Empresa) e os CRMs (*Customer Relationship Management* – Gestão de Relacionamento com o cliente).

O surgimento do termo *Business Intelligence* deu no ano de 1989, criado por Howard Dresner, quando era analista do *Gartner Group* (consultoria de pesquisa de mercado na área de tecnologia da informação). Até então, a maioria das grandes empresas contavam com os Centros de Informação (CI), que mantinham o estoque de dados, mas propiciavam pequena disponibilidade de informação. Mesmo assim, eles supriam, até certo ponto, as necessidades de tomadas de decisões, fornecendo para tanto, relatórios e informações gerenciais.

O interesse mais pelas ferramentas de BI veio à tona por volta do final de 96, quando o seu conceito começou a ser largamente utilizado como uma evolução do EIS (*Executive Information Systems* – sistemas de informação executiva), um software

criado no final dos anos 70 por meio dos trabalhos desenvolvidos pelos pesquisadores do MIT.

Com o passar do tempo e por meio de um processo natural de evolução e crescimento, o BI incorporou uma grande quantidade de ferramentas, as soluções DSS (*DecisionSupport System – sistema de suporte à decisão*), planilhas eletrônicas, geradores de consultas e relatórios, *Data Warehouses*, *Data Marts*, *Data Mining*, Ferramentas OLAP, o próprio EIS, entre outras. Todas com o objetivo de auxiliar a tomada de decisão empresarial.

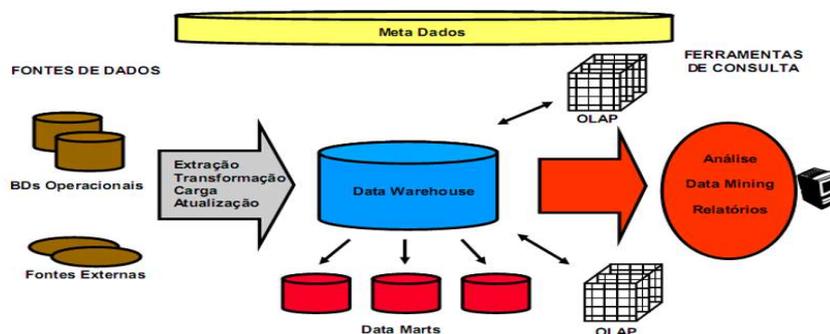
Em meio a todo esse avanço, alguns problemas começam a aparecer, como por exemplo, a dificuldade de extração de informação dos grandes volumes de dados. Muitas informações bastante valiosas ficam armazenadas dentro da própria empresa, porém com acesso bastante difícil, onde não há uma homogeneidade de armazenamento, com múltiplos estoques de dados de lógicas próprias e dificuldades de consolidação.

Atrelando principalmente esta dificuldade com o mercado bastante exigente e competitivo, os gestores da área responsável pela logística empresarial começaram a perceber que somente utilizar sistemas operativos como ERPs não era suficiente. Era necessário realizar uma análise mais estratégica da informação, reunindo os dados numa base única e trabalhá-los de forma que possibilitassem a realização de várias análises sob vários ângulos. Por isso, as empresas que implantam sistemas ERP começaram a incorporar módulos de soluções de BI cada vez mais sofisticados, para atender às necessidades de gestão.

## **2.2 Componentes do *Business Intelligence***

Como dito anteriormente, o BI não é uma tecnologia única, mas um conjunto de componentes, como o *Data Warehouse*, *Data Mart*, OLAP, *Data Mining*, ETL e ferramentas de visualização, que juntos auxiliam na tomada de decisão, como mostrado na figura 1.

**Figura 1**  
**Componentes do Business Intelligence**



Fonte: <http://www.uniriotec.br/~tanaka/SAIN/01-VisaoGeral.pdf>  
2011

Basicamente, sua implementação consiste em unificar os dados em uma base de dados comum, um *Data Warehouse* ou *Data Mart*, por meio de ferramentas ETL, para posterior visualização de dados de forma estatística, através de gráficos, planilhas, relatórios, entre outros. Pode-se usar também a técnica de *Data Mining* para a descoberta de informações valiosas, que não podem ser encontradas em banco de dados relacionais, sendo também muito importante para as tomadas de decisão. Cada um destes componentes será explicado a seguir.

### 2.2.1 *Data Warehouse*

Segundo Inmon (1996), *Data Warehouse* é uma coleção de dados orientados por assuntos, integrados, variáveis com o tempo e não voláteis, para dar suporte ao processo de tomada de decisão. Em outras palavras, *Data Warehouse* é um banco de dados utilizado para armazenar informações relativas às atividades de uma organização de forma consolidada. Possibilita a análise de grande volume de dados, coletados a partir de sistemas transacionais (OLTP – *Online Transaction Processing*). Possui as seguintes características:

- **Orientado ao assunto:** o *Data Warehouse* deve ser orientado por temas de principal interesse das organizações, os quais podem ser clientes, produtos,

promoções. Sua abordagem deve levar em consideração apenas dados que sejam relevantes ao processo de tomada de decisão.

- **Integrado:** Esta é uma das principais características do *DataWarehouse*, onde se define a apresentação única dos dados provenientes dos vários outros sistemas que irão compor a sua base. Os dados devem ser sempre integrados, com nenhuma exceção. Neste processo, muitos problemas podem ocorrer e necessitam ser resolvidos. Estes problemas geralmente estão relacionados a valores homônimos, sinônimos, conflitos de chave e domínio, além do problema da forma de apresentação dos dados. Para exemplificar um caso, pode ocorrer que um determinado sistema de informação (A) do *DataWarehouse* represente o sexo da pessoa com os valores M = “masculino” e F = “feminino”. Outro sistema de informação (B) representa os mesmos valores com os campos numéricos 1 e 2. Quando os dados destes sistemas A e B forem integrados no *DataWarehouse*, eles devem possuir a mesma representação, como mostra a Figura 2.

**Figura 2**  
**Integração dos dados**

| <b>SISTEMA</b> | <b>AMBIENTE OPERACIONAL</b> | <b>DATA WAREHOUSE</b> |
|----------------|-----------------------------|-----------------------|
| A              | Sexo: M, F                  | M, F                  |
| B              | Sexo: 1, 2                  | M, F                  |

**Fonte: Autor**  
**2016**

- **Não Volátil:** Os dados inseridos no *DataWarehouse* não são atualizados ou modificados em momento algum. Uma vez que eles são inseridos, são usados somente para acesso.
- **Variante no Tempo:** O dado em um DW refere-se a um momento específico. Isto nada mais é do que manter o histórico dos dados durante um período de tempo muito superior ao dos sistemas transacionais. Um dado em produção em um sistema transacional é atualizado de acordo com a mudança de estado do objeto em questão. Isso reflete, em geral, o estado do objeto no momento do acesso. Em um *DataWarehouse*, a cada ocorrência de uma mudança, uma nova entrada é criada para marcar esta mudança. Os dados podem ser vistos de

forma temporal através de dimensões de tempo, como dia, mês, ano, semestre, entre outros.

Os dados históricos do *DataWarehouse* possibilitam uma melhor análise de eventos passados, oferecendo suporte às tomadas de decisões presentes e a previsão de eventos futuros. Representam dados sobre um horizonte de tempo distante de 5 até 10 anos, geralmente.

Atualmente, por sua capacidade de sumarizar e analisar grande volume de dados, o *DataWarehouse* passou a ser o núcleo de sistemas de informações gerenciais e apoio à decisão das principais soluções de BI.

### **2.2.2 Data Mart**

Um datamart é um subconjunto de um data warehouse no sentido de que é construído com um escopo menor de informações, sendo segmentado para uma determinada área ou assunto (Gray e Watson, 1998). A sua diferença está apenas com relação ao tamanho do escopo do problema a ser resolvido. Enquanto ele trata de um problema local, um *DataWarehouse* envolve esforço de toda a empresa para que o suporte à decisão atue em todos os níveis da organização. Em outras palavras, um *Data Mart* atende às necessidades de unidades específicas de negócios, ao invés da corporação como um todo. Sua intenção é dividir a centralização dos dados de um *DataWarehouse* para departamentos específicos.

Existem vantagens e desvantagens na sua implementação em relação ao *DataWarehouse*, dependendo de inúmeros fatores, como o tamanho da empresa em questão, número de bases de dados que irão integrar o projeto, interfaces com outros sistemas, quantidades de profissionais na equipe, comprometimento, dedicação e domínio da equipe para a implementação, ferramentas que serão utilizadas, entre outros.

O caminho tradicional sempre foi começar construindo um *DataWarehouse* e a partir dele os *Data Marts*, abordagem denominada *Top-Down*. Porém, percebendo tais vantagens e desvantagens, os arquitetos começaram a fazer o contrário, segmentando

a construção do *DataWarehouse*, começando por departamentos com *Data Marts* até a criação do todo, abordagem denominada *Botton-up*. A principal vantagem no segundo caso está em relação ao tempo de implementação.

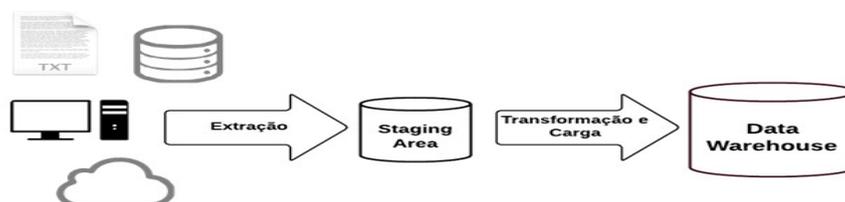
Não há uma melhor solução absoluta, isto depende de vários fatores da organização. Vai depender de cada empresa avaliar a sua demanda e optar pela melhor solução que lhe atenda.

### 2.2.3 O processo de ETL

Para os dados serem carregados no *DataWarehouse*, eles precisam passar pelo processo de ETL (*Extract, Transformand and Load* – Extração, Transformação e Carga), o qual é um conjunto de ferramentas cuja função é extrair os dados provenientes das mais diversas fontes externas, transformando estes dados para atenderem às necessidades de negócios e por fim, realizar a carga dos dados, geralmente em um *Data Warehouse* ou em um *Data Mart*. Segundo COLAÇO (2004), num projeto de construção de um *DataWarehouse*, os procesos ETL consomem mais de 70% do tempo de desenvolvimento.

Pode-se utilizar as ferramentas de ETL para fazer todo o tipo de trabalho de importação, exportação e transformação dos dados para diversas outras bases ou necessidades específicas. A figura 3 ilustra fluxo do ETL e seus principais componentes.

**Figura 3**  
**Etapas do processo ETL**



Fonte: <http://corporate.canaltech.com.br/noticia/business-intelligence/entendendo-o-processo-de-etl-22850>  
2011

- **Extração:** consiste na coleta de dados dos sistemas de origem, chamados de sistemas fontes (em inglês, *data sources*) ou sistemas operacionais, extraíndo-os e transferindo-os para uma área chamada *staging area* (área de transição ou área temporária), onde serão convertidos para um único formato.
- **Transformação:** consiste em padronizar os dados com relação a fatores como: tamanho e tipo, substituição de caracteres estranhos, correção de erros de digitação, duplicidade de informações, substituição de dados não preenchidos ou não informados, padronização de nomenclaturas e unidades de medida, entre outros fatores. Após esta etapa ser bem analisada e homologada por profissionais com função específica de análise de negócios, os dados estão prontos para a carga.
- **Carga:** após os dados serem extraídos das mais diversas fontes, transformados e homogêneos para atender às necessidades de negócio, eles serão carregados no *Data Warehouse* ou *Data Mart*. Este processo pode ser bastante demorado e requer muito processamento. Portanto, procura-se torná-lo mais curto, pois a base será atualizada constantemente, de preferência diariamente, em horários de menor utilização dos sistemas que vão alimentar o *Data Warehouse*, podendo também ser programada para executar mais vezes durante o dia de acordo com a necessidade.

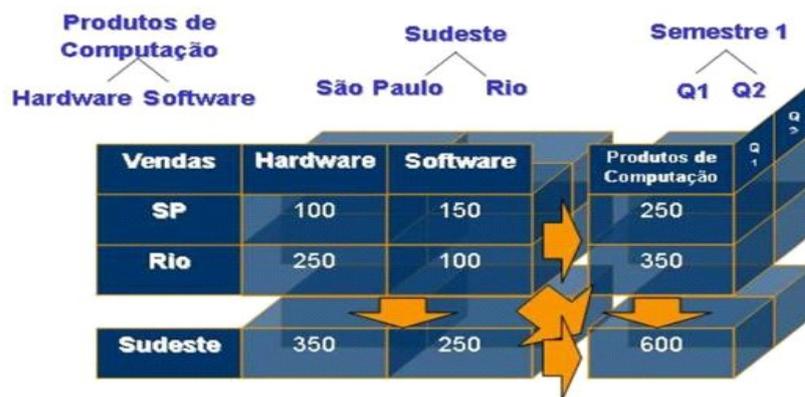
#### 2.2.4 OLAP

Segundo Inmon (1997), OLAP (Processamento Analítico On-line) representa um conjunto de tecnologias projetadas para suportar análises e consultas *ad hoc* (consultas casuais e inéditas). Em outras palavras, é a capacidade de manipular e analisar um grande volume de dados sob múltiplas perspectivas. São utilizadas por gestores em qualquer nível da organização para lhes permitir análises comparativas que facilitem a sua tomada de decisões diárias em sua corporação. OLAP trouxe uma grande capacidade de efetuar cálculos complexos como previsões, percentuais de crescimento e médias diversas considerando-se a variável tempo. É uma ferramenta muito importante no contexto gerencial, ajudando a analisar de forma mais eficiente, a

quantidade de dados crescente armazenada pelas organizações, transformando-os em informação útil (THOMSEM, 2002).

A base de dados de um *DataWarehouse* possui a característica de ser modelada de forma multidimensional. Sua finalidade é fornecer subsídio para realização de análises de dados sob diferentes perspectivas. A visão multidimensional permite o uso mais intuitivo para o processamento analítico pelas ferramentas OLAP. A arquitetura e terminologia desse tipo de modelagem são diferentes das utilizadas para bancos de dados transacionais. Os seus dados são visualizados em forma de um cubo, devido ao fato de ter várias dimensões. Um exemplo de sua visualização é mostrado na figura 4, a qual exibe a quantidade de produtos vendidos por estado ao longo dos semestres.

**Figura 4**  
Visualização do cubo



Fonte: <https://vivianeribeiro1.wordpress.com/2011/07/12/o-que-e-olap>  
2011

O tipo de pesquisa ilustrado na imagem acima se deve ao fato de existirem diversas informações históricas com a possibilidade de serem cruzadas e agrupadas. Alguns termos são utilizados para a modelagem multidimensional:

- **Dimensões:** uma dimensão é uma visão de negócio que faça sentido para a sua análise, como produto, departamento, cliente ou tempo. Elas estabelecem a organização dos dados, determinando possíveis consultas/cruzamentos. Por exemplo, um cubo contendo informações de compras de clientes poderá ser composto pelas dimensões tempo, região, produto, cliente, cenário e medidas.

- **Medidas:** são basicamente os valores a serem analisados, geralmente não fixos. Situam-se na tabela fato. Por exemplo: médias, totais e quantidades.
- **Fatos:** são os dados agrupados pelas tabelas de dimensões, contendo os valores de cada medida para cada combinação das dimensões existentes. É a tabela mais importante do modelo multidimensional.
- **Agregações:** são as totalizações calculadas nos diversos níveis hierárquicos.

Para se ter uma ideia melhor, na Figura 4 é possível fazer vários tipos de análises, como exemplo, a quantidade de vendas de hardware e software no Estado de São Paulo em determinado semestre, a quantidade de produtos de computação vendidos na região Sudeste no primeiro semestre do ano, entre outros.

As ferramentas OLAP, segundo COLAÇO (2004), possuem os seguintes tipos de arquiteturas:

- **ROLAP (*Relational On Line Analytical Processing*):** a consulta é enviada ao servidor de banco de dados relacional e processada no mesmo, mantendo o cubo no servidor. Os dados são armazenados de forma relacional.
- **MOLAP (*Multidimensional On Line Analytical Processing*):** o processamento e armazenamento são realizados em um servidor de banco de dados multidimensional.
- **HOLAP (*Hybrid On Line Analytical Processing*):** é a junção de tecnologias onde há uma combinação entre a estrutura ROLAP e a MOLAP.

Existem ainda, segundo COLAÇO (2004), várias operações de navegação que podem ser feitas em um cubo:

- **Drill Across:** ocorre quando o usuário pula um nível intermediário dentro de uma mesma dimensão. Por exemplo, o usuário passa de uma visualização de compras de um determinado ano e vai direto para a visualização em um determinado mês, sem passar pelo semestre.

- **Roll-up (ou Drill Up):** ocorre quando o usuário diminui o nível de detalhe, aumentando a granularidade. Esta granularidade determina quais os tipos de consultas que podem ser feitas em um *DataWarehouse* ou *Data Mart*, influenciando diretamente na velocidade de acesso às informações e no volume de dados analisados.
- **Drill Down:** é o contrário do *Roll-up*. Ocorre quando o usuário aumenta o nível de detalhe, diminuindo a granularidade.
- **Drill Throught:** semelhante ao *Drill-down*, A diferença que é que o drill-throught envia os dados para o menor nível de detalhe daquela informação, por exemplo, sistemas OLTP para visualização de dados isolados.
- **Slice and Dice:** uma das principais características do OLAP. Tem o objetivo de modificar a posição de uma informação, trocando linhas por colunas de maneira que facilite a compreensão dos usuários, girando o cubo sempre que tiver necessidade.

Os dois principais tipos de modelos multidimensionais são o Modelo Estrela (*Star Schema*) e o Modelo Floco de Neve (*SnowFlake*). No modelo Estrela, todas as tabelas relacionam-se diretamente com a tabela de fatos, assim todas as tabelas de dimensão devem conter todas as descrições que são necessárias para definir uma classe como Cliente, Vendedor ou Tempo. Caso fossem normalizada até a 3FN, seriam necessárias muitas operações de junção para recuperar as informações agrupadas. Por isso, este modelo é mais simples e mais fácil de navegação, porém há um maior desperdício de espaço pelo fato de os dados não estarem normalizados. Possui esta denominação porque a tabela de fato fica ao centro cercada das tabelas de dimensão, assemelhando-se a uma estrela.

O Modelo Floco de Neve é uma extensão do Modelo Estrela, onde as tabelas dimensões também se relacionam com a tabela fato, mas algumas dimensões relacionam-se apenas entre elas. Este tipo de tabela recebe o nome de dimensão auxiliar. Isto ocorre para fins de normalização. Este modelo consegue reduzir o espaço de armazenamento dos dados dimensionais, mas em contrapartida acrescenta várias

outras tabelas, tornando-o mais complexo, dificultando a navegação pelos softwares de análise.

### **2.2.5 Data Mining**

Além das ferramentas OLAP, pode-se também utilizar a técnica de Mineração de Dados (*Data Mining*), a qual consiste em explorar uma grande quantidade de dados (*Data Warehouse* ou *Data Mart*) à procura de padrões consistentes, detectando relacionamentos entre variáveis e descobrindo novos subconjuntos de dados. É a exploração e a análise, por meio automático ou semiautomático, de grandes quantidades de dados, a fim de descobrir padrões e regras significativos (Berry e Linoff, 1997). Estes padrões revelam alguma novidade a serem potencialmente úteis de forma a trazer algum benefício para a tomada de decisões estratégicas.

Algumas diferenças entre OLAP e *Data Mining* devem ser levadas em consideração. Enquanto ferramentas OLAP servem para análise multidimensional a fim de navegação entre os dados, realizando uma análise específica e simulando situações, *Data Mining* é realizada de forma automática ou semiautomática e consegue realçar correlações eventuais para mostrar tendências, descobrindo dados que não são previamente conhecidos.

Um clássico exemplo de sua utilização é o das cervejas e fraudas. Uma das maiores redes de varejo dos Estados Unidos, com base na utilização das técnicas de mineração, descobriu que a venda de fraudas descartáveis estava associada à de cerveja. Em geral, casais jovens ao irem ao supermercado comprar fraudas descartáveis para seus filhos, aproveitavam para comprar seus estoques de cerveja para o final de semana, afinal muitos deles não saíam mais para festas por conta de seus filhos. Isto fez com que os gerentes tomassem a decisão de colocar as prateleiras de cervejas e fraudas lado a lado. O resultado disso foi o estouro de vendas destes dois produtos.

Para aplicação das técnicas de mineração, é preciso passar pelo processo de Descoberta de Conhecimento em Banco de Dados (*Knowledge Discovery in Databases* – KDD), o qual é composto por cinco etapas.

A primeira etapa, chamada de seleção, consiste em selecionar os dados de maior relevância para os negócios que serão inseridos em um *DataWarehouse* ou *Data Mart*. A segunda etapa chama-se pré-processamento e limpeza, consiste na realização de limpeza de dados, eliminando as redundâncias e inconsistências e recuperando dados incompletos. Na terceira etapa, chama de transformação, os dados após serem selecionados, limpos e pré-processados, são armazenados e formatados adequadamente para que os algoritmos de mineração possam ser aplicados.

A quarta etapa do KDD consiste na aplicação propriamente dita das técnicas de mineração. Existem vários algoritmos para isto, com por exemplo os algoritmos Apriori, um dos mais tradicionais e utilizados, Árvore de padrão de crescimento frequente, Árvores de decisão ID3, Redes Neurais, Redes Bayesianas, k-Médias, k-Medoides, Regressão Linear, Regressão Exponencial, Apriorii-All e PrefixScan.

Os principais tipos de informações obtidas com a Mineração de Dados podem ser definidos da seguinte forma:

- **Associações:** são ocorrências ligadas a um único evento. Por exemplo: estudos sobre modelos de compra em supermercados pode revelar que, na compra de pão, compra-se também manteiga em 75% das vezes: mas, quando há uma promoção, a manteiga é comprada em 90% das vezes. Com isso, os gerentes podem tomar decisões mais precisas, pois aprenderam acerca da rentabilidade de uma promoção.
- **Sequências:** são eventos ligados ao longo do tempo. Com isso, pode-se descobrir, por exemplo, que quando se compra um apartamento, em 70% por cento das vezes se adquire um novo fogão no período de duas semanas; e em 40% das vezes, compra-se também um micro-ondas um mês após a compra do apartamento.
- **Classificação:** consiste em reconhecer modelos que descrevem o grupo ao qual um item pertence por meio de exame de itens já classificados e pela definição de um conjunto de regras. Por exemplo, a classificação pode ajudar a descobrir as características de clientes de operadoras de cartão de crédito que poderão vir a abandoná-las e oferecer um modelo para resgatá-los.

- **Aglomeraco:** semelhante a classificao, com a diferena de que ainda no foram definidos grupos. Consegue descobrir diferentes grupos em uma massa de dados. Por exemplo, descobrimento de grupos de clientes de um banco com base em investimentos pessoais.

A quinta e ltima etapa consiste na interpretao e avaliao pelos gestores dos resultados obtidos atravs das tcnicas de minerao que podem ser de diversos tipos, como listagens, saídas grficas, tabelas de resumo ou visualizaes. Atravs destes resultados, os gestores podero definir suas aes estratgicas.

### 2.3 Fatores de Sucesso e Fracasso nos projetos de BI

Iniciar um projeto de BI no  uma tarefa fcil e este  um dos maiores desafios das empresas atualmente.  necessrio seguir a risca vrios fatores para obter sucesso em sua implementao. Muitas empresas ainda no possuem nada relacionado ao BI, existem tambm muitas que j realizaram projetos de implementao ou at j utilizam suas tcnicas, mas ele  projetado indevidamente ou de forma ineficiente, resultando muitas vezes no fracasso de seus projetos.

Segundo DEL PINO (2014), os dez principais motivos para o fracasso do BI so descritos abaixo:

- **Requisitos pouco claros:** muitas empresas usam as aplicaes de BI apenas para confirmar o que faziam anteriormente no Excel, pelo fato de no saberem exatamente o que querem.
- **Dados incorretos ou incompletos:** dados com pouca qualidade so uma causa frequente de grandes problemas nos projetos de BI, por mais persuasivo que o design de aplicaes de BI possa ser.
- **Usurios finais envolvidos tardiamente:**  indispensvel que se inclu colaboradores das unidades de negcio que iro trabalhar com a aplicao final j nas fases iniciais do projeto de BI.
- **Resultados apresentveis apenas aps dois anos:** quando a equipe apresenta os seus primeiros resultados aps dois anos,  altamente provvel que estes se desviem significativamente das suas expectativas iniciais.

- **Falta de gestão da mudança:** a falta de uma gestão de mudanças resulta rapidamente em custos adicionais e em atrasos na conclusão do projeto.
- **Cumprimento e segurança negligenciados:** as equipes de projeto raramente têm em conta as normas, as regras e os conceitos de segurança, desde o início, para integrar futuras mudanças o mais facilmente possível.
- **Documentação pobre sobre o ambiente da aplicação:** não é raro que os projetos mais abrangentes de BI revelem que a documentação existente sobre a aplicação está incorreta ou desatualizada. Como resultado desta negligência, surgem custos adicionais e, normalmente, um atraso no projeto.
- **Recursos de hardware cotados de forma incorreta:** muitas vezes, os requisitos de hardware são subestimados, resultando em um desempenho pobre e usuários finais insatisfeitos. Há também empresas que são generosas na determinação das suas necessidades de hardware, o que deixa os recursos inativos e leva a custos contínuos consideráveis.
- **Funcionários centrados no Excel:** muitos funcionários ainda não estão preparados para perder os velhos hábitos, apoiando-se exclusivamente no Excel, subestimando a prática que será necessária para tornar a aplicação um sucesso.
- **Um orçamento inadequado:** o custo de um projeto de BI, que irá proporcionar transparência aos processos de negócio e fornecer dados para uma gestão eficaz, não pode ser coberto por fundo.

Ainda segundo DEL PINO (2014), empresas que aprendem a partir dos erros dos outros são as que melhor planejam suas aplicações de BI. Os problemas devem ser identificados a tempo para as chances de sucesso serem boas.

### 3. CONCLUSÃO

O uso das tecnologias de *Business Intelligence* está cada vez mais se tornando algo indispensável para sobrevivência no mercado nos dias de hoje. A tendência é de que as empresas busquem cada vez mais a sua implantação e consolidação em seus ambientes de trabalho, afinal muitas delas, ainda que utilizem softwares de BI, não possuem uma maturidade suficiente para aplicação em suas estratégias de gestão.

Diante disso, conclui-se que as empresas devem quebrar seus paradigmas e agir de maneira proativa, com um maior planejamento, criando estratégias de implantação e determinando fielmente seus objetivos para os projetos de BI. O artigo reuniu estudos sobre os seus principais componentes, seguindo um ordem de tempo em relação à sua implementação para se ter uma maior ideia de como o BI é composto e os benefícios trazidos por ele.

## ABSTRACT

*Due to the current globalized society and the dynamism generated by the market, which is undergoing increasingly rapid change and becoming increasingly demanding, it is essential to create strategies that companies can face high competition and can survive in this world competitive. For this, they must respond quickly to change and improve their processes through management techniques through Business Intelligence, transforming raw data into useful information for strategic operations and making decisions in order to bring real gains for business. Its implementation, however, requires that the company can accurately define your goals and how much it is willing to invest, aligning the project with the strategic interests of the corporation. The objective of this work together studies on the key components of the BI explaining each, and demonstrate how this powerful process operates within the organization and its importance in strategic planning for decision-making, with a qualitative research of bibliographical.*

**KEYWORDS: Business Intelligence; Strategic Planning; Decision Taking.**

#### 4. REFERÊNCIAS

BERRY, M.J.A.; LINOFF, G. **Data mining techniques**. John Wiley & Sons, Inc. 1997.

COLAÇO, Methanias Jr. **Projetando Sistemas de Apoio à Decisão Baseados em Data Warehouse**. Rio de Janeiro: Axcel Books do Brasil Editora, 2004.

DEL PINO, Manuel. **Dez principais motivos para o fracasso do BI**. <http://cio.com.br/tecnologia/2014/02/24/dez-principais-motivos-para-o-fracasso-do-bi/>. Acesso em 06 de junho de 2016.

(ELIAS, 2011) ELIAS, D. **Entendendo o processo ETL**. <http://corporate.canaltech.com.br/noticia/business-intelligence/entendendo-o-processo-de-etl-22850>. Acesso em 26 de abril de 2016.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. **Advances in Knowledge Discovery & Data Mining**. 1 ed. American Association for Artificial Intelligence, Menlo Park, Califórnia, 1996.

FINKELSTEIN, Sydney e SANFORD, Sade H. **Learning from corporate mistakes: The rise and fall of Iridium**. Organizational Dynamics, 2000. Disponível em <<http://mba.tuck.dartmouth.edu/pages/faculty/syd.finkelstein/articles/Iridium.pdf>>. Acesso em 05 de maio de 2016.

GRAY, P. e WATSON, H. J. **Decision support in the Data warehouse**. The Data Warehousing Institute Series, New Jersey, Prentice Hall PTR, 1998.

INMON, W. H. **Building the Data Warehouse**. New York: John Wiley & Sons Inc, 1996

MINTZBERG, Henry et al. **Strategy Safari: A Guided Tour Through The Wilds of Strategic Management**. New York: The Free Press, 1998.

PRASS, Fernando Sarturi. **Um visão geral sobre as fases do Knowledge Discovery in Databases (KDD)**. 2012.

PRIMAK, Fábio Vinicius. **DECISÕES COM B.I. (BUSINESS INTELLIGENCE)**. Ed. Ciência Moderna, 2008.

(RIBEIRO, 2011) RIBEIRO, V. **O que é OLAP?**. <https://vivianeribeiro1.wordpress.com/2011/07/12/o-que-e-olap>. Acesso em 03 de maio de 2016.

(TANAKA, 2011) TANAKA, A. K. **Sistemas de Apoio à Inteligência do Negócio: Arquitetura de Data Warehouse**. <http://www.uniriotec.br/~tanaka/SAIN/01-VisaoGeral.pdf>. Acesso em 28 de abril de 2016.

THOMSEN, E. OLAP: **Construindo sistemas de informações multidimensionais**. Rio de Janeiro: Campus, 2002.

WITTEN, Ian H. Data Mining – **Practical Machine Learning Tools and Techniques**. Burlington: Elsevier Inc., 2011.