

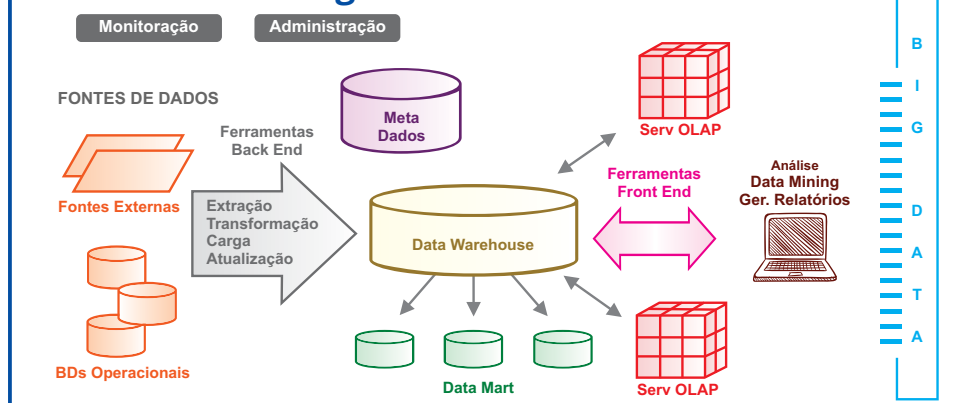
RESUMO BI

Autores:

Grimaldo Oliveira - Consultor de Business Intelligence
grimaldo_lopes@hotmail.com - www.bicomvatapa.blogspot.com - www.aprendavirtual.com

Juracy Almeida - Consultor de Business Intelligence
juracyajr@gmail.com - https://www.linkedin.com/pub/juracy-junior/8/567/9ba/pt

Business Intelligence



BUSINESS INTELLIGENCE: Na tradução livre, significa Inteligência de negócios ou simplesmente "BI". É um conjunto de técnicas e ferramentas que, reunidas, possibilitam ao gestor a tomada de uma decisão com base em métricas ou valores que são extraídos de seus diversos sistemas e consolidados. O BI oferece ao gestor os números para análise, tanto histórica quanto comportamental, de sua empresa além da identificação de tendências de seu público.

Finalidade: Prover suporte decisório de qualidade nas organizações, disponibilizando informações sobre clientes, mercado, negócios e processos com as seguintes características:

- Informações confiáveis, padronizadas, unificadas, com fácil e rápido acesso.
- Composição de análises diferenciadas.
- Visualização intuitiva das informações.

Técnicas e conceitos envolvidos num BI:

- Data Warehouse
- OLTP x OLAP
- Metadados
- Dimensões
- Fatos
- Data Mining
- Big Data

DATA WAREHOUSE (DW): Também conhecido como "armazém de dados", é um banco de dados que **integra e consolida** os diversos sistemas e fontes de dados (arquivos de texto, planilhas eletrônicas, banco de dados, imagens, mensagens, qualquer fonte de informação que possa ser extraída) de uma organização para **apoio na tomada de decisão**. Possui as seguintes características:

- **Não Volátil:** Os dados que são inseridos na base do Data Warehouse **não serão mais atualizados**. Ou seja, como o dado já tem como origem outro sistema onde foi todo tratado, neste ambiente ele não sofre mais atualização;
- **Integrado:** Os dados inseridos no DW têm como origem diversos sistemas. Às vezes, cada um trata da mesma informação de diferentes maneiras. Exemplo: a informação do sexo, no sistema A é guardado como 'Homem'/'Mulher' e no sistema B, como 'Masculino'/'Feminino'. Quando for para o DW, há a **integração e padronização**, assumindo e transformando para um único valor como, por exemplo, 'Masculino'/'Feminino'. Assim, os dados, quando oriundos do sistema A, são transformados para este padrão quando inseridos no DW;

- **Orientado ao Tempo:** Para uma adequada análise do gestor dos dados históricos, um DW é sempre **orientado ao tempo**. Assim, é possível avaliar a evolução dos dados e também realizar comparações em tempos diferentes. Por exemplo, comparar a média de vendas realizadas nos Estados do Sul no primeiro Trimestre de 2014 e 2015, permitindo alguma ação do gestor com base na resposta deste cenário;

- **Orientado a Assunto:** Um DW, entre os seus objetivos, busca a consolidação de informações dos diversos assuntos ou departamentos da organização em um único banco de dados. Por isso, este sempre vai ser construído **orientado a assuntos**. Para ser qualificado como um DW, este deve ser formado e oferecer diversos assuntos. Quando o armazém de dados tratar apenas de um assunto da organização, este é qualificado como um **Datamart**.

Caso surjam outros "**Datamarts**", então teremos disponível um **Data Warehouse**.

GRANULARIDADE: Refere-se ao nível de detalhe que a informação chega no DW. Assim, é correto afirmar que, quanto **menor** for o nível de detalhe, **mais alto** será o nível de granularidade dos dados e **maior** será a possibilidade do sistema responder a qualquer consulta. Do mesmo jeito que, quanto maior for o nível de detalhe, **mais baixo** será o nível de granularidade dos dados e **menor** será a possibilidade de o sistema responder a qualquer consulta. Exemplo: Em um DW que trata do assunto de Notas Fiscais, se a granularidade chegar até os itens da nota, perceba que o volume de dados será imenso. Já se o "grão" for maior chegando apenas no número da nota, teremos um menor volume de dados.

OLTP X OLAP

- **OLTP:** "**Online Transaction Processing**" se refere aos sistemas operacionais ou transacionais da organização. Ou seja, são aqueles sistemas que controlam os processos e as operações da empresa. Nestes, os dados são voláteis, ou seja, os dados são inseridos, alterados ou excluídos a todo o momento.

FINALIDADE PRINCIPAL: Permitir a execução operacional do negócio.

Entretanto, no OLTP, a execução de consultas e análises gerenciais é lenta.

- **OLAP:** "**Online Analytical Processing**" se trata da capacidade de analisar os grandes volumes de dados armazenados no DW apoiada por ferramentas que permitem as consultas com respostas rápidas e funções de análises gerenciais do negócio.

Métodos de armazenamento dos dados OLAP:

- **ROLAP:** A organização dos dados no banco de dados segue a modelagem RELACIONAL. Indicado para DW;
- **MOLAP:** Os dados são organizados seguindo a modelagem multidimensional (tópico tratado mais a frente). Mais indicado para DATA-MARTS;
- **HOLAP:** Combinação (Híbrido) do ROLAP e MOLAP.

Características OLAP:

- **Dimensões:** Estrutura que armazena qualquer visão ou descrição do negócio, como região, produtos, tempo, fornecedores;
- **Cubos:** Estrutura que armazena as métricas pré-processadas em memória, acelerando as consultas e permitindo a visualização de seus valores por diversos ângulos (dimensões);



- **Slice:** Seleção ("fatia") de uma única dimensão de um cubo OLAP;
- **Dice:** Extração de um "subcubo" a partir do original com duas ou mais dimensões;
- **Pivot Table:** Possibilidade de inverter as posições dos dados de linhas para colunas e vice-versa e, assim, realizar a análise de diferentes perspectivas.
- **DRILL:** É a Navegação pelo dado. Esta navegação normalmente ocorre dentro de um cubo através das dimensões.
 - **DRILL UP(ROLL-UP):** Navegar do nível de maior detalhe para um nível acima para agregação;
 - **DRILL DOWN:** Navegar do nível mais agregado para o de maior detalhe;
 - **DRILL ACROSS:** É a operação OLAP que permite relacionar fatos diferentes através de dimensões compartilhadas.
 - **DRILL THROUGH:** Quando há uma troca de uma informação analisada para outra.

São alguns exemplos de bancos MOLAP e que possuem estas características acima listadas: Microsoft SQL Server Analysis Services e Oracle OLAP (Essbase).

METADADOS: São dados que descrevem outros dados. Num Data Warehouse, visam descrever todos os objetos e processos envolvidos na sua construção. Sua existência é fundamental tanto para o gestor de negócios quanto para a equipe de manutenção do projeto, promovendo a rápida identificação das fontes e transformações utilizadas, assim como dos cálculos aplicados nas métricas.

É um artefato que deve ser gerado, atualizado e disponibilizado durante toda existência do DW.

STAGING AREA: Staging Area ou área auxiliar serve como ponto único para a carga efetiva no data warehouse, a cada carga seu conteúdo é limpo. Também evita o acesso à produção em caso de recarga durante o dia.

DIMENSÕES: As dimensões são estruturas que descrevem visões importantes dos assuntos tratados no DW. Exemplos de dimensões: departamento, tempo, produto, clientes, funcionários.

Nas análises a partir do DW, as dimensões se tornam cabeçalhos de colunas ou linhas, enquanto as métricas das fatos, os valores entre elas.

Características:

- **Hierarquias:** Uma dimensão pode ser organizada com níveis hierárquicos que representam melhor suas informações a serem analisadas. Um exemplo é uma dimensão sobre Países, que pode ter as hierarquias de REGIÕES, ESTADOS, MUNICÍPIOS;
- **Membros:** É um subconjunto da dimensão que representa os dados daquele nível da hierarquia;
- **Versionamento:** Característica exclusiva das dimensões. É o controle dos dados que foram modificados ao longo do tempo naquela dimensão. Com este atributo, o dado histórico armazenado na fato irá se referir ao descritor correspondente ao tempo pesquisado e não ao mais recente, como ocorreria se não houvesse o versionamento.

FATO: É utilizada para composição de análises e relatórios, armazenando métricas e sendo quantificado quando associado a uma ou mais dimensões. São todos os dados quantitativos que podem ser levantados com o cliente, como: valores faturados (valores em reais), pagamentos dos clientes com as compras (valores em reais), pagamentos aos fornecedores (valores em reais), quantidade de mercadorias no estoque, gastos com compra de mercadorias (valores em reais), dentre outros.

GRÃO DA FATO: Conceito que identifica a unidade de medida das métricas. Influencia diretamente o tamanho de uma tabela Fato e está ligado ao nível de detalhe dos dados. A lógica é a seguinte: **Mais detalhe → Mais dados → Análise mais longa → Informação mais detalhada → Grão baixo.** Já ao contrário: **Menos detalhe → Menos dados → Análise mais rápida → Informação menos detalhada → Grão alto.**

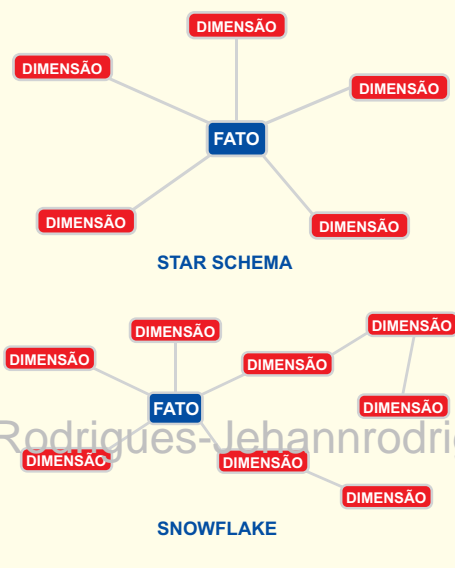
TIPOS DE FATO:

- **Aditivas:** As métricas aditivas são aquelas que podem ser sumarizadas independente das dimensões utilizadas.
- **Semi-Aditivas:** As métricas semi-aditivas são aquelas que podem ser sumarizadas em alguns casos. A sumarização só fará sentido em algumas dimensões.
- **Não-Aditivas:** As métricas não-aditivas são aquelas que não podem ser sumarizadas ao longo das dimensões.

ETL (ETC): Originada do inglês Extract, Transform, Load, significa, respectivamente Extração, Transformação e Carga, processo que os dados são submetidos antes de ingressar ao repositório de dados do Data Warehouse.

MODELAGEM MULTIDIMENSIONAL: Chamada também de modelagem dimensional, trata-se do tipo de modelagem normalmente utilizada para a construção de um Data Warehouse. Caracteriza-se por uma tabela central (FATO) e pela suas tabelas periféricas (Dimensões). Possui intuitividade e alta performance proporcionada por sua estrutura.

- **Star Schema (Modelo Estrela):** Tipo de estrutura utilizada na modelagem multidimensional com formato análogo ao estelar. Caracteriza-se por suas tabelas periféricas. Dimensões são desnormalizadas.
- **Snowflake (Modelo Floco de Neve):** Tipo de estrutura utilizada na modelagem multidimensional, onde algumas das tabelas periféricas (Dimensões) possuem normalização de dados.



FERRAMENTAS EM BI:

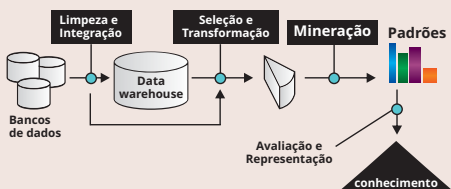
Gratuitas: Pentaho Business Intelligence, Talend Integration (ETL), JasperSoft, Alfresco, dentre outras.

Proprietárias: OBIEE (OLAP), ODI (ETL), Microstrategy (OLAP), Qlikview (OLAP), Tableau, Yellowfin, dentre outras.

DATA MINING: Também conhecido como Mineração de Dados, é o processo de extração de informações, sem conhecimento prévio em base de dados na busca por padrões nos dados.

Padrão: São unidades de informação que se repetem, ou, então, são sequências de informações que dispõe de uma estrutura que se repete. EX: sequência: 123XGF123DGR5F123YTHGJ123UJH7123CHGJ9T Padrão: 123

KDD: Processo criado para busca de padrões de dados de forma interativa, composto de etapas sequenciais. Limpeza e integração de dados, seleção de dados, Mineração de Dados e avaliação de resultados.



Técnica: A técnica de mineração consiste na definição de métodos que nos garantam como descobrir os padrões que nos interessam Ex: árvore de decisão, redes neurais, dentre outros.

Tarefa: está intrinsecamente ligada ao que estamos querendo buscar nos dados, que tipo de categoria de padrões possuímos interesse em encontrar. Ex: classificação, regressão, agrupamento (cluster), dentre outros.

- **Classificação:** É representada por um conjunto de decisões que tem por objetivo a classificação de dados baseados nos valores dos atributos; estes atributos serão classificados em uma ou várias classes predefinidas.
- **Regressão:** Determina uma equação matemática que descreve a relação entre duas ou mais variáveis.
- **Agrupamento (cluster):** Esta técnica é usada para segmentar uma base de dados, identificando um conjunto de categorias ou grupos com características similares.

FERRAMENTAS DATA MINING:

- **Many Eyes:** Produto da IBM que produz minerações a partir de interações visuais: Bolhas, Treemap, mapas geográficos, dentre outros.
- **Knime:** Muito parecida com as ferramentas de ETL, esta se propõe a estabelecer um fluxo sequencial para que a mineração de dados seja realizada.
- **R:** É uma linguagem de programação que possui um ambiente de desenvolvimento integrado para a criação de cálculos estatísticos e gráficos.
- **WEKA:** WEKA (Waikato Environment for Knowledge Analysis) está implementado na linguagem Java, com uma característica importante, ser portátil. Possui um formato próprio para a leitura dos dados, conhecido como ARFF; os algoritmos são executados por linha de comando ou interface gráfica.

BIG DATA



Refere-se ao uso de tecnologias que viabilizam o uso de uma grande quantidade de dados a serem armazenados e com uma maior velocidade. Os dados podem ser dos mais variados tipos: Estruturados (Banco de dados) e não-estruturados (vídeos, tuítes, e-mails). O uso do Big Data se baseia no chamado 5 "V": velocidade, volume, variedade, veracidade e valor.

1

Prova: COSEAC - 2009 - DATAPREV - Analista de Tecnologia da Informação - Análise de Informações
Disciplina: Banco de Dados

Sobre a relação DW e OLAP, pode-se afirmar:

I - DW é utilizado para armazenar informações e o OLAP para recuperá-las, ambos são especializados para exercer suas funções de forma eficiente;

II - DW e OLAP são poderosas tecnologias independentes e não complementares;

III - Para a exploração completa do DW, a ferramenta OLAP irá extrair e alavancar as informações contidas nele.

Dos itens acima mencionados, apenas:

- a) I é verdadeiro;
- b) II é verdadeiro;
- c) III é verdadeiro;
- d) I e III são verdadeiros;
- e) II e III são verdadeiros.

2

Prova: CESGRANRIO - 2008 - Petrobrás - Analista de Sistemas Júnior - Processos de Negócios
Disciplina: Banco de Dados | Assuntos: DW - Data Warehouse;

A seguir são feitas algumas afirmações a respeito de data warehouses e ferramentas OLAP.

I - Os usuários finais do data warehouse, em geral, não possuem acesso à Data Staging Area.

II - Drill in, drill out, roll over e roll on são típicas operações disponibilizadas pelas ferramentas de consultas OLAP para navegar pela hierarquia de uma dimensão.

III - As rotinas de ETL muitas vezes originam solicitações de mudanças e melhorias nos sistemas OLTP e outras fontes de dados que alimentam o data warehouse, pois têm o potencial de revelar inconsistências entre os diversos sistemas corporativos.

IV - Um data warehouse, em geral, deve ser projetado para fazer junções entre fatos e dimensões através de chaves naturais, evitando chaves substitutas (surrogate keys), pois estas apenas contribuiriam para aumentar o tamanho e a complexidade do esquema sem nenhum benefício para o usuário final.

Estão corretas APENAS as afirmações:

- a) I e II
- b) I e III
- c) II e III
- d) II e IV
- e) III e IV

3

Prova: CESGRANRIO - 2010 - Petrobrás - Analista de Sistemas Júnior - Engenharia de Software
Disciplina: Banco de Dados | Assuntos: OLAP (On-line Analytical Processing);

Ferramentas de OLAP são aplicações que permitem uma visão dos dados por meio de perspectivas múltiplas, ajudando aos usuários a tomar decisões através da utilização de uma série de operações básicas que permitem a navegação pelas dimensões do hiper cubo de dados definidos no Data Warehouse. **Qual das seguintes descrições reflete corretamente uma propriedade das operações das ferramentas de OLAP?**

- a) A operação de roll up ou drill up diminui o nível de detalhamento da informação e não é limitado pelo grão máximo, pois os dados podem ser agregados mesmo após se chegar a este limite superior.
- b) A operação de drill down permite aumentar o nível de detalhe e não é limitado pelo grão mínimo, pois os dados podem ser desagregados mesmo após se chegar a este limite inferior.
- c) A operação de drill across consiste em navegar através das várias colunas que compõem uma visão criada dentro do hiper cubo, permitindo que o usuário se concentre apenas nos dados que o interessam.
- d) A operação de drill through consiste em realizar um drill down em mais de uma dimensão do hiper cubo, fazendo o relatório ser tão especializado quanto desejado pelo usuário.

e) A operação de dice permite que o usuário elimine do hiper cubo determinados valores das colunas que foram transferidas para sua visão dos dados, garantindo uma visão personalizada e especializada dos dados.

4

Prova: CESGRANRIO - 2010 - Petrobrás - Analista de Sistemas Júnior - Engenharia de Software
Disciplina: Banco de Dados | Assuntos: OLAP (On-line Analytical Processing);

Ferramentas de OLAP fornecem um método de visualização dos dados armazenados em um sistema de data warehouse, de forma a melhorar a velocidade de obtenção e a qualidade das informações usadas pelos executivos no processo de tomada de decisão. Existem vários tipos distintos de ferramentas de OLAP que possuem características próprias e situações em que são mais recomendadas.

Com base nesses conceitos, tem-se que o:

- a) OLTP é uma ferramenta que oferece maior portabilidade para os data warehouses, pois integra-se com linguagens de programação portáteis, independentes do sistema operacional.
- b) MOLAP, devido à sua natureza portátil, só deve ser utilizado em dispositivos móveis, tais como notebooks e celulares inteligentes (smart phones).
- c) MOLAP, devido à sua estrutura multidimensional, é a alternativa mais natural para acesso a dados em formato de hiper cubos.
- d) ROLAP, devido à sua estrutura inerentemente relacional, só pode se conectar diretamente às bases de dados transacionais, não podendo usar um sistema de data warehousing.
- e) ROLAP é a escolha mais adequada para aplicações nas quais o tempo de resposta é crítico, mas é eficiente apenas nos casos em que as chaves da tabela de fatos são iguais às dos sistemas transacionais.

5

Prova: CESPE - 2010 - MPU - Analista de Informática - Banco de Dados
Disciplina: Banco de Dados | Assuntos: DW - Data Warehouse; OLAP (On-line Analytical Processing); Data Mining;

As funções especiais ROLAP (OLAP relacional) e MOLAP (OLAP multidimensional), as extensões SQL e os métodos de junção (join) são ferramentas que podem ser utilizadas para aumentar a eficiência das consultas em bases de dados multidimensionais.

○ Certo ○ Errado

6

Prova: FUNRIO - 2013 - MPOG - Analista de Tecnologia da Informação
Disciplina: Banco de Dados | Assuntos: OLAP (On-line Analytical Processing);

Na modelagem dimensional de dados para Data Warehouse, existem dois tipos de tabelas, representando os fatos contendo os dados granulares e os pontos de entrada específicos chamados de dimensões que descrevem os fatos. A modelagem dimensional facilita as consultas com operações OLAP (de Processamento Analítico On Line, em Inglês). **A operação OLAP que permite relacionar fatos diferentes através de dimensões compartilhadas é denominada:**

- a) pivoteamento
- b) drill down.
- c) drill up.
- d) drill across.
- e) slice and dice.

7

Prova: CESPE - 2013 - MPOG - Tecnologia da Informação
Disciplina: Banco de Dados

Julgue os itens que se seguem, acerca das ferramentas ETL (extract transform load) e OLAP (on-line analytical processing). OLAP é uma tecnologia utilizada para organizar grandes bancos de dados e fornece, para organizações, um método com alta flexibilidade e desempenho para acessar, visualizar e analisar dados corporativos. Os dados podem ser organizados em uma hierarquia que define diferentes níveis de detalhe, na qual o usuário pode navegar para cima(drill up) ou para baixo (drill down) entre níveis.

○ Certo ○ Errado

8

Prova: CESPE - 2014 - TJ-SE - Técnico Judiciário - Programação de Sistemas
Disciplina: Não definido

Acerca de DataMining e de DataWarehouse, julgue o item subsequente. Um DataWarehouse provê uma excelente base para a realização de DataMining, pois os algoritmos de DataMining demandam grandes quantidades de dados em nível detalhado; o DataMining tira vantagem de bases de dados que estejam integradas e limpas; e a infraestrutura necessária para a criação de um DataWarehouse atende às necessidades das operações de DataMining.

○ Certo ○ Errado

9

Prova: CESPE - 2014 - TJ-SE - Técnico Judiciário - Programação de Sistemas
Disciplina: Não definido

Acerca de DataMining e de DataWarehouse, julgue o item subsequente. Os principais processos de DataMining são a identificação de variações embasado em normas, a detecção de relacionamentos, a paginação de memória e o controle de periféricos.

○ Certo ○ Errado

10

Prova: CESPE - 2014 - TJ-SE - Analista Judiciário - Banco de Dados
Disciplina: Banco de Dados

Julgue os próximos itens, com relação a DataMining e ETL. O uso de agrupamento (clustering) em DataMining exige que os registros sejam previamente categorizados, tendo por finalidade aproximar registros similares para prever valores de variáveis.

○ Certo ○ Errado

11

Prova: CESPE - 2011 - Correios - Analista de Correios - Analista de Sistemas - Desenvolvimento de Sistemas
Disciplina: Banco de Dados (TI) | Assuntos: Data Warehouse

Acerca de soluções de suporte a decisão, julgue os itens seguintes. Em um ambiente data warehouse (DW), é possível a análise de grandes volumes de dados, os quais ficam disponíveis para serem alterados e manipulados pelo usuário.

○ Certo ○ Errado

12

Prova: CESGRANRIO - 2010 - Petrobrás - Analista de Sistemas Júnior - Processos de Negócios

Data warehouse tem por característica ser:

- a) orientado a assuntos
- b) temporário
- c) descentralizado
- d) focalizado em aspectos operacionais
- e) normalizado

13

Prova: CESPE - 2009 - INMETRO - Analista Executivo em Metrologia e Qualidade - Desenvolvimento de Sistemas - Parte II

Entre as operações típicas realizadas em uma organização que adota a abordagem de business intelligence, destaca-se o uso de mineração de dados e a construção de datawarehouses, ambas empregando bancos de dados relacionais com elevado grau de normalização.

○ Certo ○ Errado

14

Prova: CESPE - 2007 - TCU - Analista de Controle Externo - Tecnologia da Informação

Quanto ao nível de granularidade dos dados do data warehouse, é correto afirmar que quanto maior for o nível de detalhe, mais alto será o nível de granularidade dos dados e maior será a possibilidade de o sistema responder a qualquer consulta.

○ Certo ○ Errado

15

Prova: CESPE - 2007 - TCU - Analista de Controle Externo - Tecnologia da Informação

O processamento OLAP é mais flexível do que aquele que ocorre no nível estruturado organizacional do data warehouse.

☐ Certo ☐ Errado

16

Prova: CESPE - 2010 - TRT - 21ª Região (RN) - Analista Judiciário - Tecnologia da Informação

O data mining é um processo automático de descoberta de padrões, de conhecimento em bases de dados, que utiliza, entre outros, árvores de decisão e métodos bayesianos como técnicas para classificação de dados.

☐ Certo ☐ Errado

17

Prova: FEPESE - 2010 - SEFAZ-SC - Auditor Fiscal da Receita Estadual - Parte III - Tecnologia da Informação

Assinale a alternativa correta a respeito de Data Warehouse (DW) e modelagem multidimensional.

- a) Um esquema multidimensional, composto por fatos e dimensões, não pode ser modelado em um banco de dados relacional.
- b) Uma dimensão definida em uma modelagem multidimensional deve estar associada a um único fato.
- c) Operações OLAP podem ser executadas tanto para fins analíticos quanto para fins de atualização de dados.
- d) O volume de dados de um DW geralmente é superior ao volume de dados de um BD transacional.
- e) Um Data Mart é um componente de um DW que não possui dados históricos.

18

Prova: CESGRANRIO - 2010 - ELETROBRÁS - Analista de Sistemas - FUNCIONAL SAP-ERP

Os dados armazenados em um data warehouse, que fornecem informações para um ambiente multidimensional para tomada de decisões de uma empresa, não estão normalizados. **Essa redundância de dados, criada propositalmente, tem como objetivo:**

- a) oferecer segurança na obtenção da informação para o ambiente multidimensional.
- b) melhorar o desempenho de consultas complexas sobre o data warehouse.
- c) permitir a distribuição das informações por data marts departamentais.
- d) facilitar a agregação de dados atômicos.
- e) relacionar, com maior acuidade, os dados atômicos com os transacionais.

19

Prova: CESGRANRIO - 2010 - IBGE - Analista de Sistemas - Desenvolvimento de Aplicações

O texto a seguir se refere à modelagem de Data Warehouse. Se na modelagem do Data Warehouse for adotada uma abordagem _____, cada elemento de dados (por exemplo, a venda de um item) será representado em uma relação, chamada tabela de fatos, enquanto que as informações que ajudam a interpretar os valores ao longo de cada dimensão são armazenadas em uma tabela de dimensões, uma para cada dimensão. Esse tipo de esquema de banco de dados é chamado um esquema estrela, em que a tabela de fatos é o centro da estrela e as tabelas de dimensões são os pontos. Quando a abordagem _____ é escolhida, um operador específico que faz a agregação prévia da tabela de fatos ao longo de todos os subconjuntos de dimensões é utilizado e pode aumentar substancialmente a velocidade com que muitas consultas _____ podem ser respondidas.

Considerando a ordem das lacunas, qual sequência de termos completa corretamente o texto acima?

- a) MOLTP, ROLTP, OLTP.
- b) ROLTP, MOLTP, OLTP.
- c) ROLAP, MOLAP, OLAP.
- d) ROLAP, MOLAP, OLTP.
- e) MOLAP, ROLAP, OLAP

20

Prova: CESPE - 2010 - Banco da Amazônia - Técnico Científico - Análise de Sistemas

A tabela de fatos armazena todos os eventos ocorridos na empresa, possibilitando ao gestor a consulta de todo o seu histórico.

☐ Certo ☐ Errado

21

Prova: CESPE - 2010 - Banco da Amazônia - Técnico Científico - Análise de Sistemas

Ao explorar um data warehouse por meio de uma ferramenta OLAP, o analista pode executar as operações de drill down e drill up, que, respectivamente, aumentam e reduzem o nível de granularidade dos dados.

☐ Certo ☐ Errado

22

Prova: CESPE - 2009 - SECONTE-ES - Auditor do Estado - Tecnologia da Informação

Data Warehouse é uma coleção de dados orientados ao assunto, que tem como características visão conceitual multidimensional, operações interdimensionais irrestritas, dimensões e níveis de agregação ilimitados, sendo que os dados são não voláteis e variantes no tempo.

☐ Certo ☐ Errado

23

CESGRANRIO - 2006 - Petrobrás - Analista de Sistemas Pleno - Especialidade - Engenharia de Software

Bases de dados de sistemas de informações gerenciais são projetadas, muitas vezes, em um esquema denominado estrela que:

- a) é uma arquitetura física que permite definir somente estruturas unidimensionais de dados.
- b) é composta por uma tabela central, denominada de dimensão, e várias tabelas periféricas a ela relacionadas, denominadas fatos.
- c) apresenta tabelas periféricas que se encontram desnormalizadas.
- d) faz com que a expansão e a evolução da base de dados necessite de grande atividade de manutenção.
- e) é um esquema onde o número de junções realizadas é relativamente maior que o realizado em bases de dados relacionais convencionais.

24

Prova: CESPE - 2011 - Correios - Analista de Correios - Analista de Sistemas - Produção

Entre as utilidades de um sistema OLAP (online analytical processing) inclui-se o apoio a usuários finais em decisões estratégicas dentro das organizações.

☐ Certo ☐ Errado

25

Prova: CESPE - 2011 - PREVIC - Analista de Tecnologia da Informação

Os esquemas em estrela e em flocos de neve são dois modelos multidimensionais comuns. Dadas as suas características, o modelo em flocos de neve aumenta a redundância de dados e, por isso, aumenta o espaço utilizado em disco.

☐ Certo ☐ Errado

26

Prova: FUNCAB - 2010 - PRODAM-AM - Analista de TI - Desenvolvimento de Sistemas

No tocante aos datawarehouses, a característica desses armazéns de dados que está relacionada ao fato de que "uma vez inseridos, os dados não podem ser alterados, embora possam ser eliminados", é:

- a) organização (orientado por assunto).
- b) cliente/servidor.
- c) não volatilidade
- d) variante de tempo.
- e) relacional.

27

Prova: FCC - 2010 - TRT - 9ª REGIÃO (PR) - Técnico Judiciário - Tecnologia da Informação

Quando o usuário passa da análise da dimensão tempo e passa a analisar a dimensão região, por exemplo, ele está executando a operação OLAP:

- a) drill through.
- b) slice and dice.
- c) drill across.
- d) roll up.
- e) star.

28

Prova: CESPE - 2010 - Banco da Amazônia - Técnico Científico - Análise de Sistemas

Por fazerem uso de um processo de modelagem dimensional, os data warehouses não podem ser instalados em sistemas de banco de dados relacionais.

☐ Certo ☐ Errado

29

Prova: VUNESP - 2010 - CEAGESP - Analista de Sistemas - Desenvolvimento de Aplicações

Quando se desenha uma solução de data warehouse como base de dados para uma solução de BI, em que os dados gerados serão disponibilizados para uma base de dados OLAP, a principal característica dos dados e do processo de data warehousing é que os dados devem:

- a) ser alterados muito raramente.
- b) ser homogêneos.
- c) possuir índices.
- d) estar relacionados por constraints.
- e) pertencer a uma mesma fonte de dados.

30

Prova: CESPE - 2011 - Correios - Analista de Correios - Analista de Sistemas - Produção

Embora com alta capacidade de manipular dados, os sistemas OLAP suportam somente estruturas unidimensionais.

☐ Certo ☐ Errado

31

Prova: FCC - 2011 - TRT - 1ª REGIÃO (RJ) - Analista Judiciário - Tecnologia da Informação

Ao nível de sumarização dos elementos e de detalhes disponíveis nos dados em um DW dá-se o nome de:

- a) relacionamento.
- b) capacidade.
- c) granularidade.
- d) integridade.
- e) arquitetura.

32

Prova: FGV - 2009 - MEC - Desenvolvedor

No que diz respeito ao desenvolvimento de sistemas de informação e da análise de negócios, analise a citação a seguir.

"É um conjunto integrado de dados, extraídos de bancos de dados operacionais, históricos, externos e selecionados, editados e padronizados para recuperação e análise, para fornecer inteligência empresarial e nos negócios no processo de tomada de decisões gerenciais."

Essa citação representa o conceito de:

- a) Data Smart.
- b) Data Mining.
- c) Data Business.
- d) Data Warehouse.
- e) Data Information

33

Prova: FCC - 2009 - PGE-RJ - Técnico Superior de Análise de Sistemas e Métodos

MOLAP é:

- a) utilizado para análise de segurança e usabilidade de dados em bancos relacionais.
- b) um instrumento utilizado no tuning de bancos de dados.
- c) uma ferramenta de monitoração de redes de computadores.
- d) uma ferramenta de proteção de redes de computadores.
- e) um mecanismo utilizado no âmbito dos bancos de dados multidimensionais.

34

Prova: FCC - 2008 - TCE-AL - Analista de Sistemas

A organização dos data warehouse em tabela de fato e tabelas de dimensão relacionadas, é característica:

- a) do esquema estrela.
- b) da mineração de dados.
- c) do roll-up.
- d) do processador analítico on-line.
- e) do drill-down.

35

Prova: FCC - 2008 - TRF - 5ª REGIÃO - Analista Judiciário - Tecnologia da Informação

As consultas no star schema de um data warehouse podem ser feitas em maior ou menor nível de detalhe. Assim uma consulta mais detalhada das informações denomina-se:

- a) drill-down.
- b) data mart.
- c) data mining.
- d) roll-up.
- e) snowflake.

36

Prova: FCC - 2012 - TST - Analista Judiciário - Análise de Sistemas

O processo de ETL em uma Data Warehouse possui várias fases. Em uma destas fases é efetuada a:

- a) extração dos dados dos sistemas de origem.
- b) introdução de novos produtos no mercado.
- c) validação das interfaces de usuário.
- d) criação de diagramas estáticos e comportamentais das classes e atributos.
- e) definição dos custos e prazos.

37

Prova: CESPE - 2012 - MPE-PI - Analista Ministerial - Informática - Cargo 6

Um data warehouse é um repositório de informações de uma fonte única, armazenadas sob diversos esquemas, em um só local, e que fornecem dados consolidados.

☐ Certo ☐ Errado

38

Prova: FCC - 2011 - TRE-PE - Analista Judiciário - Análise de Sistemas

Um processo importante que ocorre em relação à formação de um data warehouse é a obtenção dos dados de uma ou mais bases de dados da origem. Deve ser rigoroso para evitar a deformação e/ou a perda dos dados quando passados da fonte original para o destino. Trata-se de:

- a) MINING.
- b) DATAMART.
- c) MOLAP.
- d) STAR.
- e) ETL.

39

Prova: CESPE - 2011 - Correios - Analista de Correios - Analista de Sistemas - Desenvolvimento de Sistemas

Acerca de soluções de suporte a decisão, julgue os itens seguintes.

Toda estrutura de dados no DW tem um elemento de tempo - como dia, mês ou ano - como referência.

☐ Certo ☐ Errado

40

Prova: CESGRANRIO - 2010 - Petrobrás - Analista de Sistemas Júnior - Processos de Negócios

No contexto de Data Warehouses, o processo de Extração, Transformação e Carga (ETC):

- a) revela-se como uma das etapas importantes do processo de criação do data warehouse, já que sua função é obter automaticamente os conhecimentos necessários para a padronização dos dados em modelos multidimensionais.
- b) produz, ao seu término, uma série de tabelas (chamadas fatos) que se caracterizam por possuírem dados normalizados até a 3ª forma normal.
- c) apresenta, como algumas de suas tarefas, filtragem, integração, conversão, condensação e derivação dos dados de entrada, que podem ser originários de diversas fontes, inclusive externas aos sistemas OLTP da organização.
- d) leva em consideração o modelo conceitual de dados das fontes de dados, que é geralmente expresso como modelo entidade-relacionamento
- e) considera somente os dados provenientes de sistemas OLTP como válidos para o processo e, caso exista a necessidade de consideração de dados externos, estes devem ser importados para os sistemas legados.

41

Prova: CESGRANRIO - 2010 - Petrobrás - Analista de Sistemas Júnior - Processos de Negócios

Técnicas de modelagem de Data Warehouses diferem das tradicionalmente utilizadas em sistemas transacionais. Analisando uma dessas abordagens, o modelo multidimensional estrela, verifica-se que:

- a) possui melhor desempenho, quando comparado ao modelo floco de neve.
- b) representa hierarquias explicitamente.
- c) facilita a tarefa de manutenção das dimensões.
- d) assemelha-se ao modelo de classes conceitual da UML.
- e) é opcional a sua utilização, sendo substituído pelo diagrama de atividades da UML

42

Prova: FCC - 2010 - TRT - 8ª Região (FA e AP) - Analista Judiciário - Tecnologia da Informação

Como melhor relação custo/benefício, em um DW é mais aconselhável:

- a) separar o dado geográfico do dado histórico, uma vez que não têm correlação.
- b) separar a identificação da informação da data a que ela se refere.
- c) armazenar somente a data das transações.
- d) armazenar dados transacionais.
- e) armazenar informações de caráter histórico e estatístico.

43

Prova: CESPE - 2010 - MPU - Analista de Informática - Banco de Dados

A tecnologia de DW tem como objetivos a extração eficiente, o processamento e a apresentação analítica de dados para suporte à decisão gerencial. Essa tecnologia utiliza o online analytical processing (OLAP) para a análise de dados complexos.

☐ Certo ☐ Errado

44

Prova: CESPE - 2013 - SERPRO - Analista - Suporte Técnico

Algumas das principais técnicas utilizadas para a realização de Datamining são: estatística, aprendizado de máquina, datawarehouse e recuperação de informações.

☐ Certo ☐ Errado

45

Prova: CESPE - 2014 - TJ-SE - Analista Judiciário - Análise de Sistemas

DataMining pode ser considerado uma etapa no processo de descoberta de conhecimento em base de dados, consistindo em análise de conjuntos de dados cujo objetivo é descobrir padrões úteis para tomada de decisão.

☐ Certo ☐ Errado

46

Prova: CESPE - 2014 - TJ-SE - Analista Judiciário - Análise de Sistemas

Assim como o DataMining, os DataMarts são voltados para a obtenção de informações estratégicas de maneira automática, ou seja, com o mínimo de intervenção humana a partir da análise de dados oriundos de DataWarehouses.

☐ Certo ☐ Errado

47

Prova: CESPE - 2011 - Correios - Analista de Correios - Analista de Sistemas - Desenvolvimento de Sistemas

Em um ambiente data warehouse (DW), é possível a análise de grandes volumes de dados, os quais ficam disponíveis para serem alterados e manipulados pelo usuário.

☐ Certo ☐ Errado

48

Prova: CESPE - 2014 - ANTAQ - Analista Administrativo - Sistemas e Negócios

Na recuperação e visualização de dados em um Data Warehouse, o drill thought ocorre quando o usuário visualiza a informação contida de uma dimensão para outra dimensão.

☐ Certo ☐ Errado

49

Prova: FCC - 2003 - TRF - 5ª REGIÃO - Técnico Judiciário - Operação de Computador

Data Warehouse é um banco de dados para sistemas de apoio à decisão com o armazenamento dos dados em estruturas:

- a) operacionais
- b) cadastrais
- c) analíticas
- d) transacionais
- e) dimensionais

50

Prova: VUNESP - 2012 - TJ-SP - Analista de Sistemas

Na especificação de uma tabela fato de um modelo multidimensional de um data warehouse, é importante definir o grau de detalhamento de seus valores. A denominação utilizada para tal detalhamento é:

- a) backroom
- b) cesta de mercado
- c) granularidade
- d) parcelamento
- e) snowflake

25 - ERRADO
24 - CERTO
23 - C
22 - CERTO
21 - CERTO
20 - ERRADO
19 - C
18 - B
17 - D
16 - CERTO
15 - CERTO
14 - ERRADO
13 - ERRADO
12 - A
11 - ERRADO
10 - ERRADO
9 - ERRADO
8 - CERTO
7 - CERTO
6 - D
5 - CERTO
4 - C
3 - A
2 - B
1 - D

GABARITO