

Análise da Informação

Manuel Martins

INFORMÁTICA

TECNOLOGIA DA INFORMAÇÃO

BUSINESS INTELLIGENCE

DATA WAREHOUSE

DATA WAREHOUSE

CARACTERÍSTICAS DA ANÁLISE
DAS FERRAMENTAS OLAP

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

GRANULARIDADE $\Rightarrow$ é o nível de **DETALHAMENTO** de uma informação em um DW.

DRILL $\Rightarrow$ é um processo que permite ao usuário AUMENTAR ou DIMINUIR o nível de DETALHAMENTO dos dados.

DRILL DOWN (ROLL DOWN) $\Rightarrow$ **AUMENTA** o nível de **DETALHE** da informação, **DIMINUINDO** a **GRANULARIDADE**.

DRILL UP (ROLL UP) $\Rightarrow$ **DIMINUI** o nível de **DETALHE** da informação e **AUMENTA** a **GRANULARIDADE**.

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

DRILL $\Rightarrow$ é um processo que permite ao usuário **AUMENTAR** ou **DIMINUIR** o nível de **GRANULARIDADE** dos dados.

Por exemplo, se um relatório estiver consolidado por países, fazendo um **DRILL DOWN**, os dados passarão a ser apresentados por estados, cidades, bairros e assim sucessivamente até o menor nível de detalhamento possível.

O processo contrário, o **DRILL UP (ROLL UP)**, faz com que os dados sejam consolidados em níveis superiores de informação.

GRANULARIDADE **ALTA**

ROLL UP

MENOS DETALHE
MENOS ESPAÇO

PAIS

REGIÃO

ESTADO

CIDADE

BAIRRO

MAIS DETALHE
MAIS ESPAÇO

DRILL DOWN

GRANULARIDADE **BAIXA**



CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

PRINCIPAIS TIPOS DE DRILL

DRILL ACROSS $\Rightarrow$ ocorre quando o usuário pula um NÍVEL INTERMEDIÁRIO dentro de uma MESMA DIMENSÃO.

Por exemplo: a dimensão tempo é composta por ano, semestre, trimestre, mês e dia. O usuário estará executando um **DRILL ACROSS** quando ele passar de ANO **DIRETO** para TRIMESTRE ou MÊS (sem passar por semestre).

DRILL ACROSS



PULA um NÍVEL INTERMEDIÁRIO dentro de uma MESMA DIMENSÃO

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

PRINCIPAIS TIPOS DE DRILL

DRILL ACROSS $\Rightarrow$ ocorre quando se executa uma consulta envolvendo mais de uma tabela **FATO** (cubo de dados). Essa operação exige que os cubos tenham pelo menos uma das dimensões em comum. A ideia é consultar as múltiplas tabelas **FATO** e colocar o resultado em um único conjunto de dados.

O DRILL ACROSS está relacionado com a capacidade de “pular” (buscar informações) de um esquema (cubo de dados) para outro, desde que ambos tenham algumas dimensões compartilhadas.

FATO $\Rightarrow$ representa uma **medição** de um **negócio**, por exemplo, a **quantidade** do **produto X vendido** na **loja Y** no **mês Z**.

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

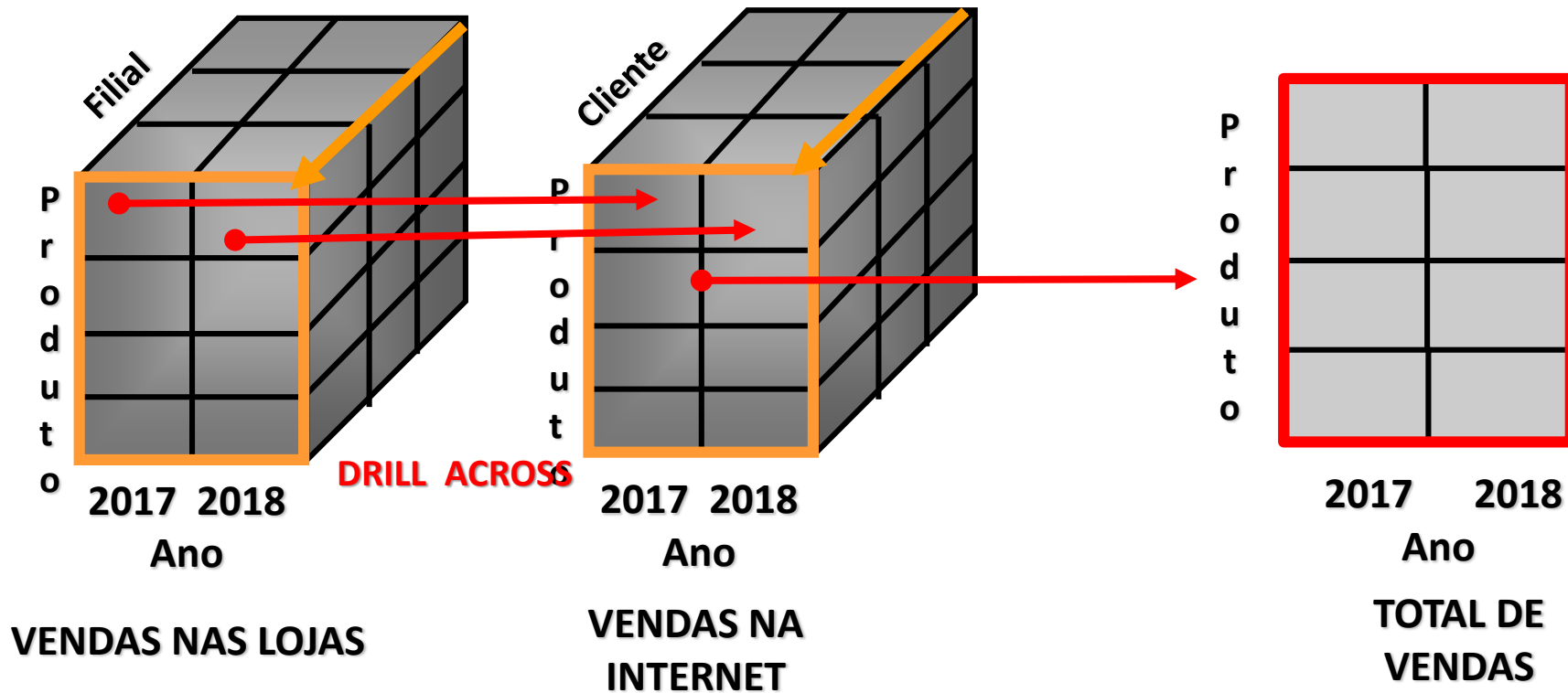
DRILL ACROSS

EXEMPLO - **VENDAS** de produtos em lojas **FÍSICAS** (**Filiais**) e em lojas **VIRTUAIS** (**Internet**) em **dois anos** (2017 e 2018).

Se o gestor deseja o **agregado por produto** em **cada ano** precisa consultar as informações de vendas nas duas tabelas.

O **DRILL ACROSS** permite o tratamento dessas informações que estão correlacionadas, mesmo que estejam em estruturas separadas, mas unidas por uma **dimensão** em **comum** que nesse exemplo será o **ano** e **produto**.

DRILL ACROSS



CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

DRILL ACROSS

- **DRILL ACROSS** - permite o tratamento de informações que estão correlacionadas, mesmo que estejam em estruturas separadas, mas unidas por dimensões em comum.
- É uma espécie de **join dimensional** entre estruturas relacionadas, parecido com o **join** relacional que nos permite buscar informações em outra tabela a partir de campos comuns através de chaves primárias e estrangeiras.
- **DRILL ACROSS** - é equivalente ao **join** relacional, porém entre esquemas dimensionais.

JOIN $\Rightarrow$ operação de **junção** na **Álgebra Relacional** - combina colunas de uma ou mais tabelas em um **Banco de Dados Relacional (BDR)**.

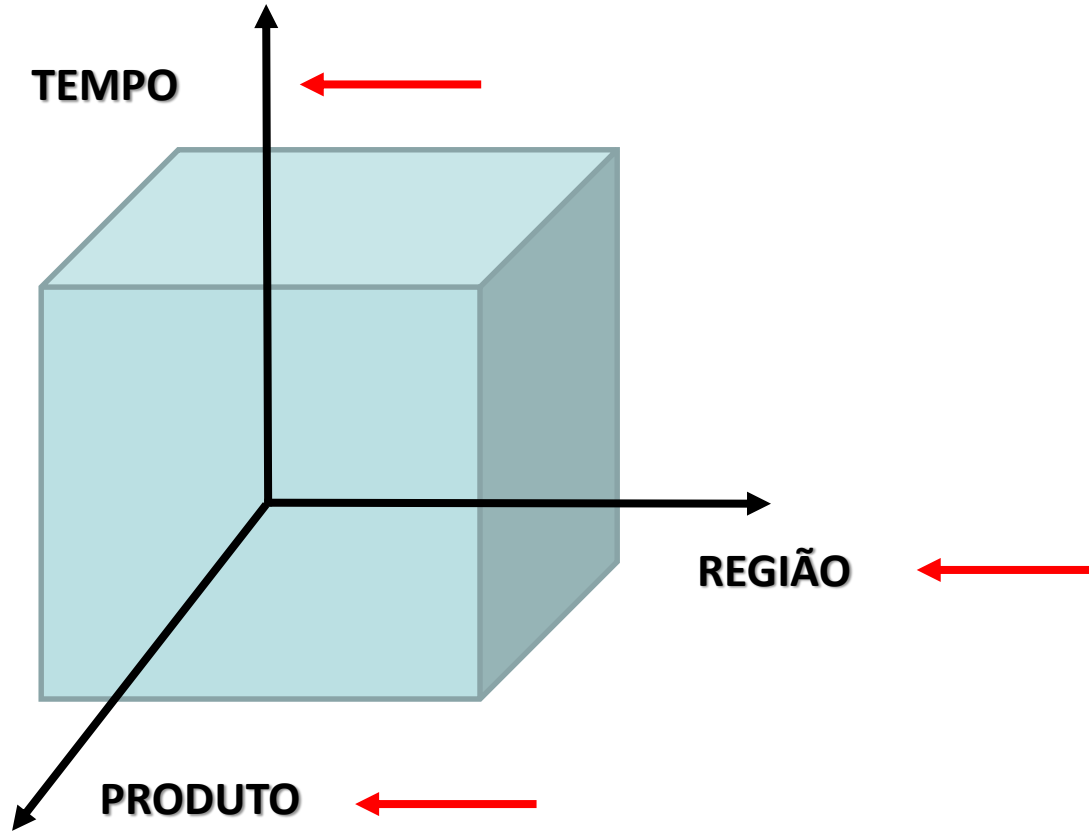


PRINCIPAIS TIPOS DE DRILL

DRILL THROUGHT $\Rightarrow$ ocorre quando o usuário **PASSA** de uma informação contida em **UMA DIMENSÃO PARA** uma **OUTRA**.

Exemplo $\Rightarrow$ o usuário está na dimensão de **TEMPO** e no próximo passo começa a analisar a informação por **REGIÃO**.

DRILL THROUGH



PRINCIPAIS TIPOS DE DRILL

DRILL THROUGHT $\Rightarrow$ é a operação que permite buscar uma informação além do nível de granularidade existente na estrutura dimensional.

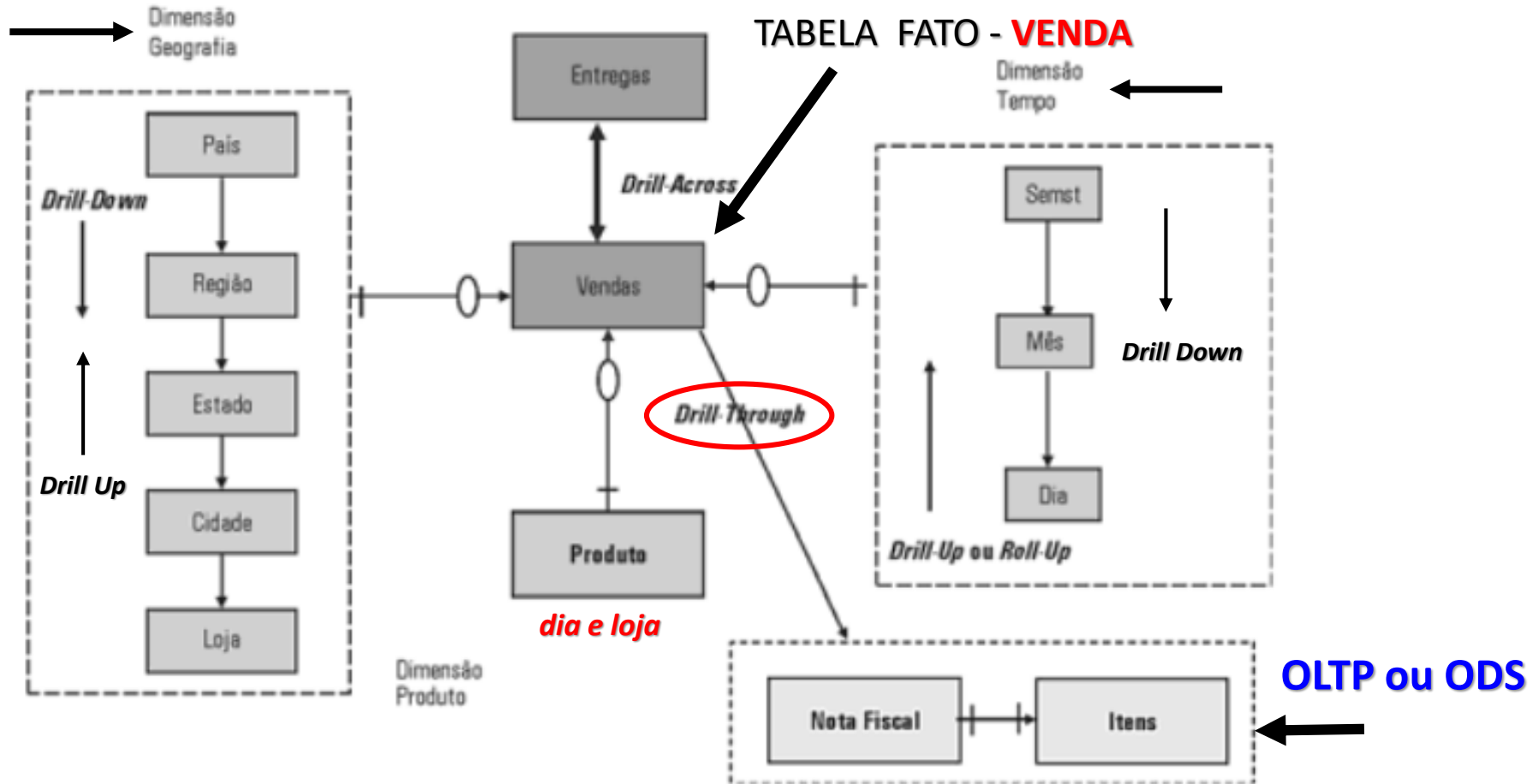
EXEMPLO - DRILL THROUGH

Suponha que você tenha armazenado na tabela FATO o valor das VENDAS com nível de granularidade de PRODUTO por dia e loja. Se você deseja as informações de VENDAS mais detalhadas, por exemplo, por NOTA FISCAL terá que buscar em outro arquivo ou ambiente diferente do DW/DM. Pode ser o próprio sistema transacional de origem (OLTP) ou em um ambiente ODS (Operational Data Store).

O DRILL THROUGH é a operação que permite essa busca além do nível de granularidade existente na estrutura dimensional.

É como se fosse um DRILL-DOWN só que buscando o detalhe da informação desejada em outra estrutura, além do esquema dimensional.

DRILL THROUGH



QUESTÕES DE PROVAS

01- Quando, em um modelo multidimensional, o usuário pular de uma informação contida em uma dimensão para outra como, por exemplo, da dimensão tempo para a dimensão região, a operação executada corresponde a

(A) *roll trough*.

(B) *select*.

(C) *drill trough*.

(D) *roll up*.

(E) *join*.

GABARITO - C

02- Uma informação de uma aplicação BI, de **detalhe menor do que aquele colocado na tabela** e permitido pela sua granularidade, pode ser recuperada por uma operação dimensional

(A) drill-through

(B) drill-across

(C) drill-down

(D) drill-up

(E) roll-up

GABARITO-A

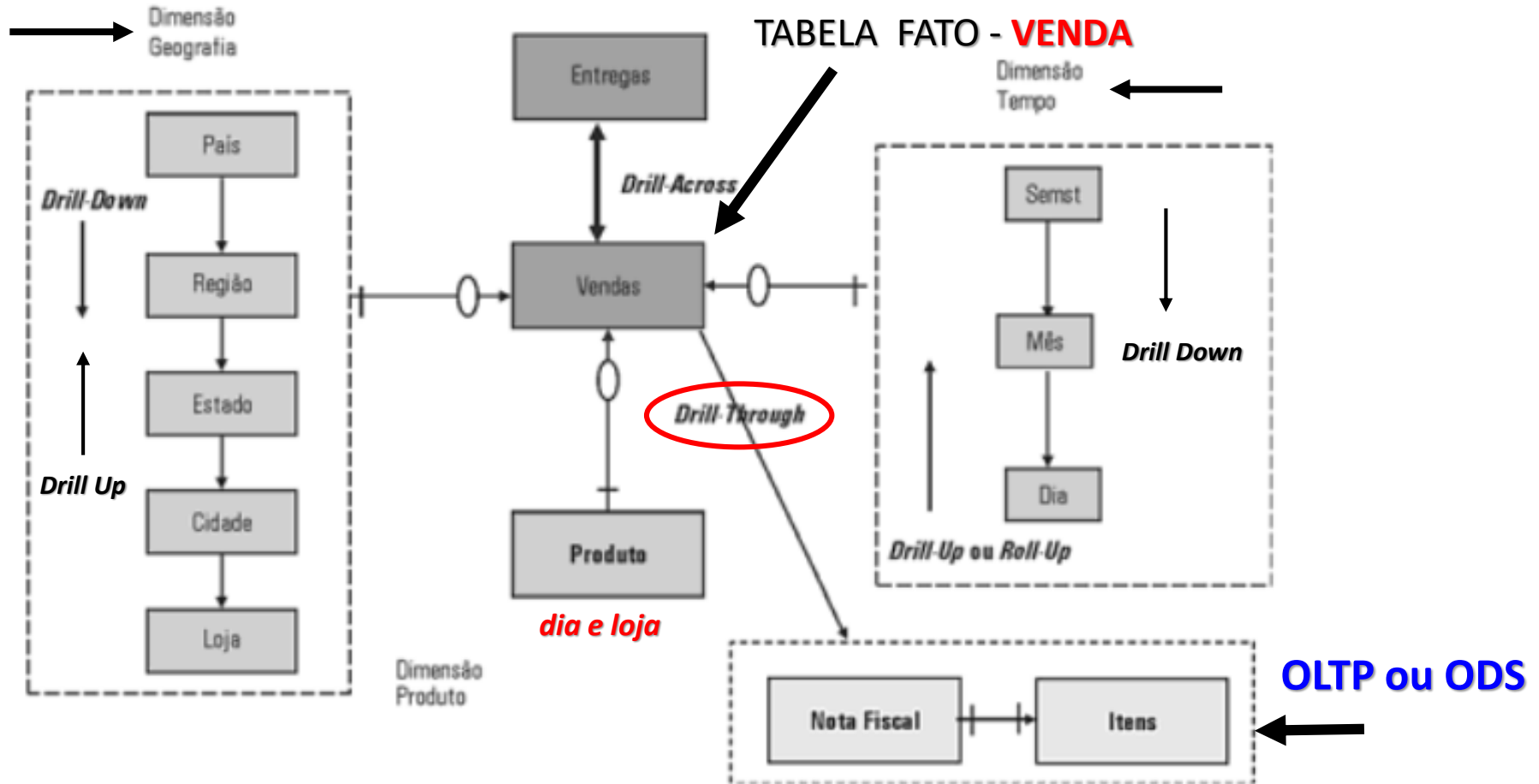
03- Buscar uma informação além do nível de granulidade existente na estrutura dimensional

- (A) é permitido por meio de uma operação *drill-down*.
- (B) é permitido por meio de uma operação *drill-up*.
- (C) não é permitido.
- (D) é permitido por meio de uma operação *drill-across*.
- (E) é permitido por meio de uma operação *drill-through*

GRANULIDADE = GRANULARIDADE

GABARITO - E

DRILL THROUGH



04- Quando, em um modelo multidimensional, o usuário pular um nível intermediário dentro de uma mesma dimensão, como por exemplo, sendo a dimensão tempo composta por ano, semestre, trimestre, mês e dia e o usuário pular de ano para mês, a operação executada corresponde a

- (A) tracking
- (B) hashing
- (C) drill down
- (D) drill across
- (E) union

GABARITO - D

05- Utilizando uma base multidimensional, o usuário passou da análise de informações sob a ótica da dimensão **TEMPO** para a visão sob a dimensão **REGIONAL**. A operação **OLAP** aí realizada foi:

- (A) *roll up*.
- (B) *drill across*.
- (C) *drill through*.
- (D) *slice and dice*.
- (E) *star across*

GABARITO - C

06- A funcionalidade pré-programada de resumir os dados, com generalização crescente, oferecida pelas aplicações por meio das ferramentas de construção de *data warehouses* é denominada

(A) *Roll up*.

ROLL UP = DRILL UP

(B) *Drill down*.

(C) *Pivot*.

(D) *Sorting*.

(E) *Slice and dice*.

GABARITO - A

07- Na análise dos dados de um sistema com o apoio de uma ferramenta OLAP, quando uma informação passa de uma dimensão para outra, inexistindo hierarquia entre elas, ocorre uma operação drill through.

CERTO

08- Na análise dos dados de um sistema com o apoio de uma ferramenta OLAP, quando uma informação passa para um nível de granularidade maior, ocorre uma operação drill up, em que os dados são exibidos com menos detalhes, mas ocupam mais espaço para armazenamento.

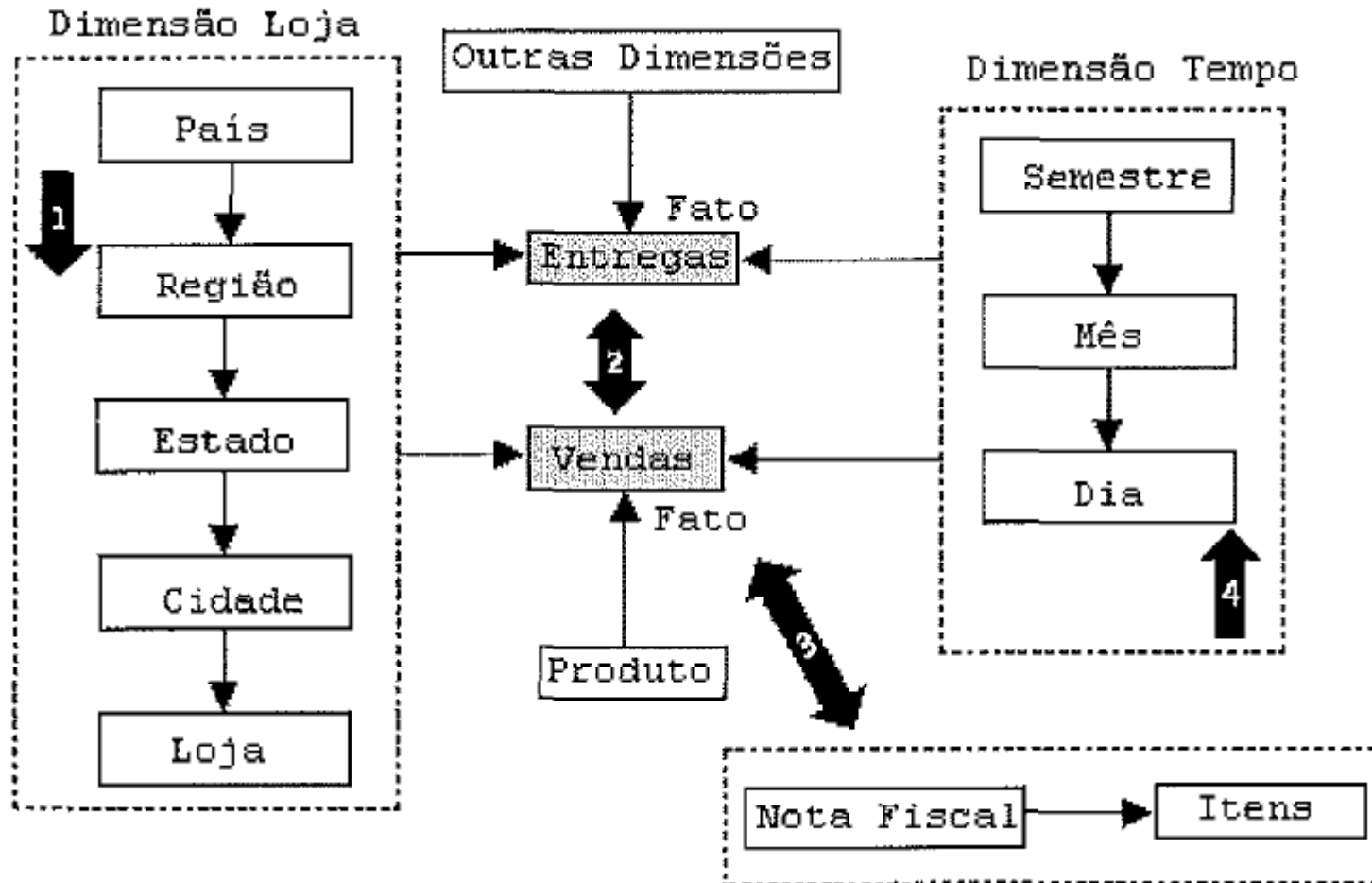
ERRADO

09- Os Data Warehouses existem para facilitar consultas complexas, intensivas em termos de dados, e ad-hoc frequentes. Uma funcionalidade de consulta pré-programada, na qual os dados são resumidos com generalização crescente (semanal para trimestral e, então, para anual), é denominada

- (A) pivô
- (B) drill-down
- (C) roll-up
- (D) slice and dice
- (E) classificação

GABARITO - C

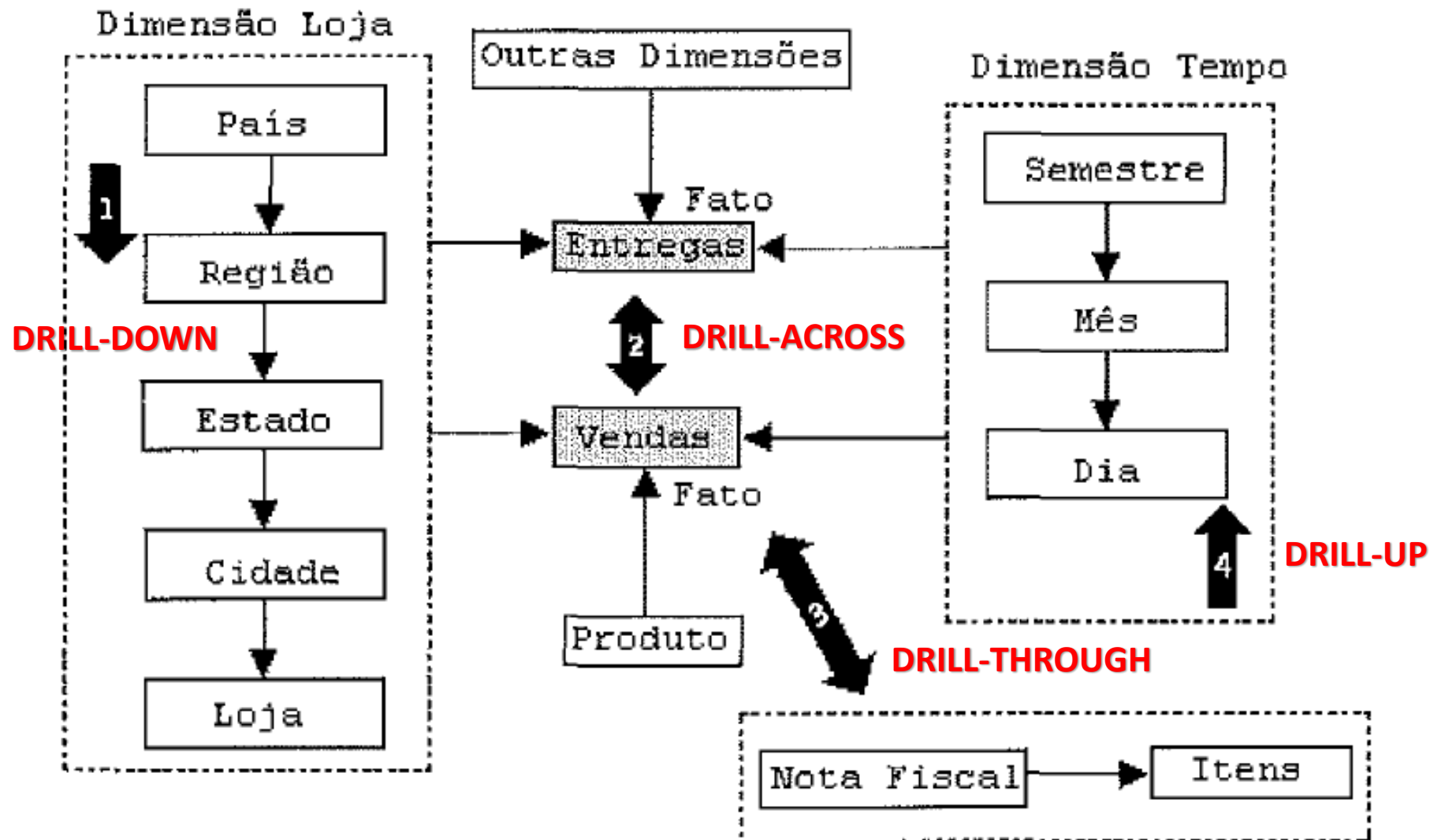
10- Observe o seguinte modelo dimensional de dados de um Data Warehouse.



Em relação ao modelo dimensional acima, é correto afirmar que as setas numeradas representam, respectivamente, os seguintes comandos dimensionais:

- (A) Drill-Down; Drill-Across; Drill-Through; e Drill-Up
- (B) Drill-Across; Drill-Through; Drill-Up; e Drill-Down
- (C) Drill-Through; Drill-Up; Drill-Down; e Drill-Across
- (D) Drill-Up; Drill-Down; Drill-Across; e Drill-Through
- (E) Drill-Through; Drill-Across; Drill-Down; e Drill-Up

GABARITO - A



DATA WAREHOUSE

CARACTERÍSTICAS DA ANÁLISE
DAS FERRAMENTAS OLAP

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

- **GRANULARIDADE**
- DRILL
 - DRILL DOWN (ROLL DOWN)
 - DRILL UP (ROLL UP)
 - DRILL ACROSS
 - DRILL THROUGH

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

SLICE AND DICE $\Rightarrow$ são duas das principais operações das ferramentas OLAP. Como o OLAP recupera o microcubo no banco de dados, essas operações permitem manipular as informações facilitando a compreensão do usuário sobre o problema.

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

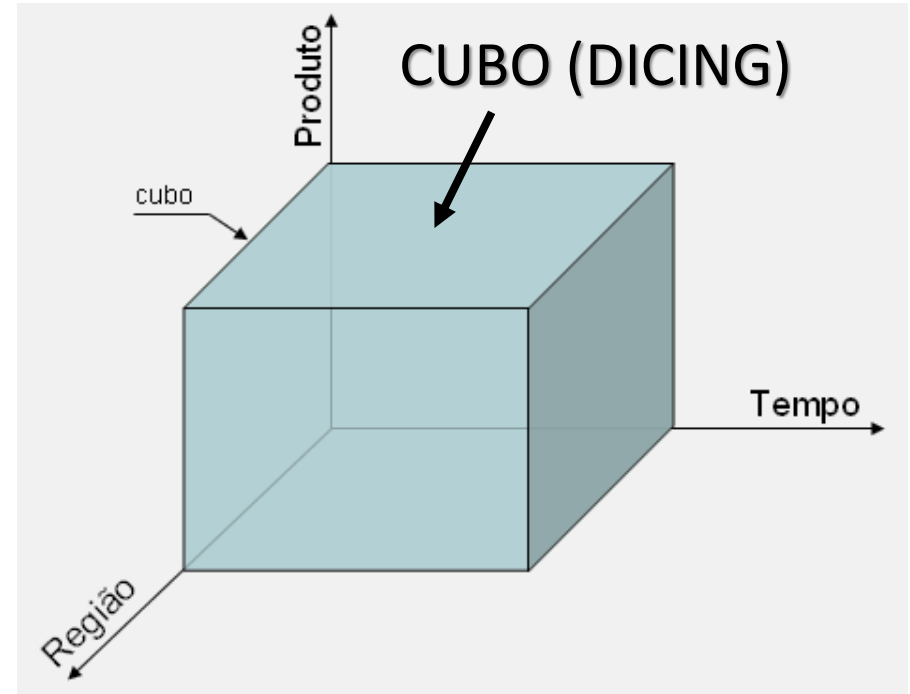
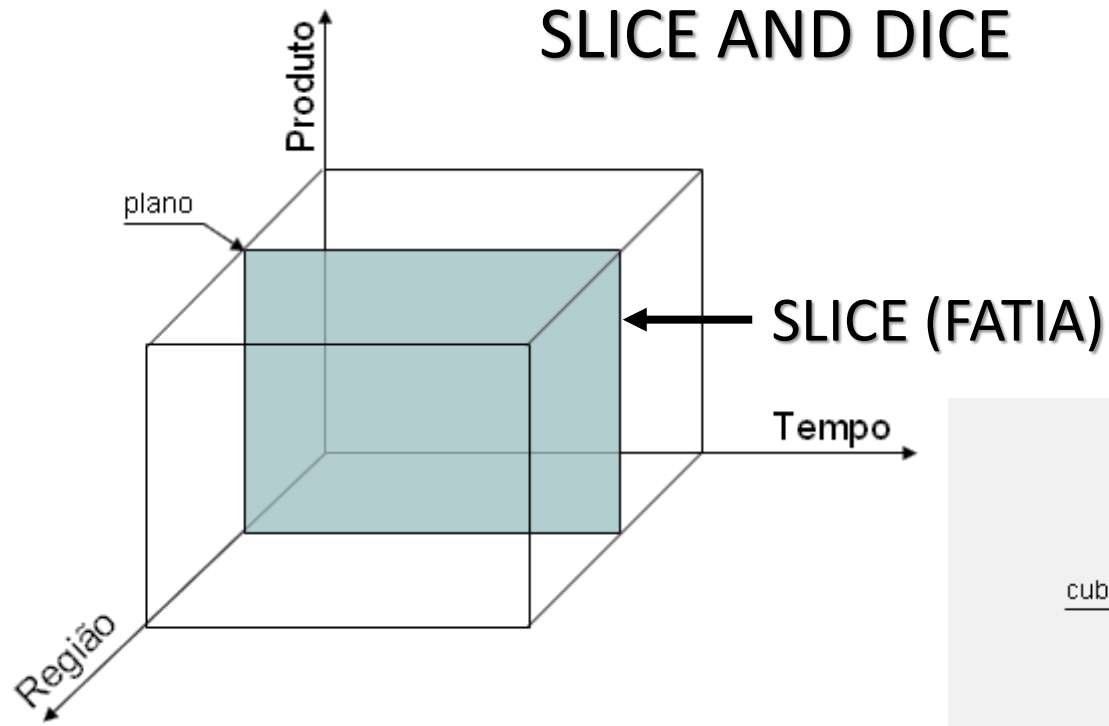
SLICE AND DICE

- SLICE $\Rightarrow$ **fatiar** - seleciona dados de uma das dimensões de um cubo.
- DICE $\Rightarrow$ define e **extraí** um **subcubo** do cubo original executando uma operação de seleção usando duas ou mais dimensões.
- SLICE AND DICE $\Rightarrow$ é o mesmo que **FILTRAR**.

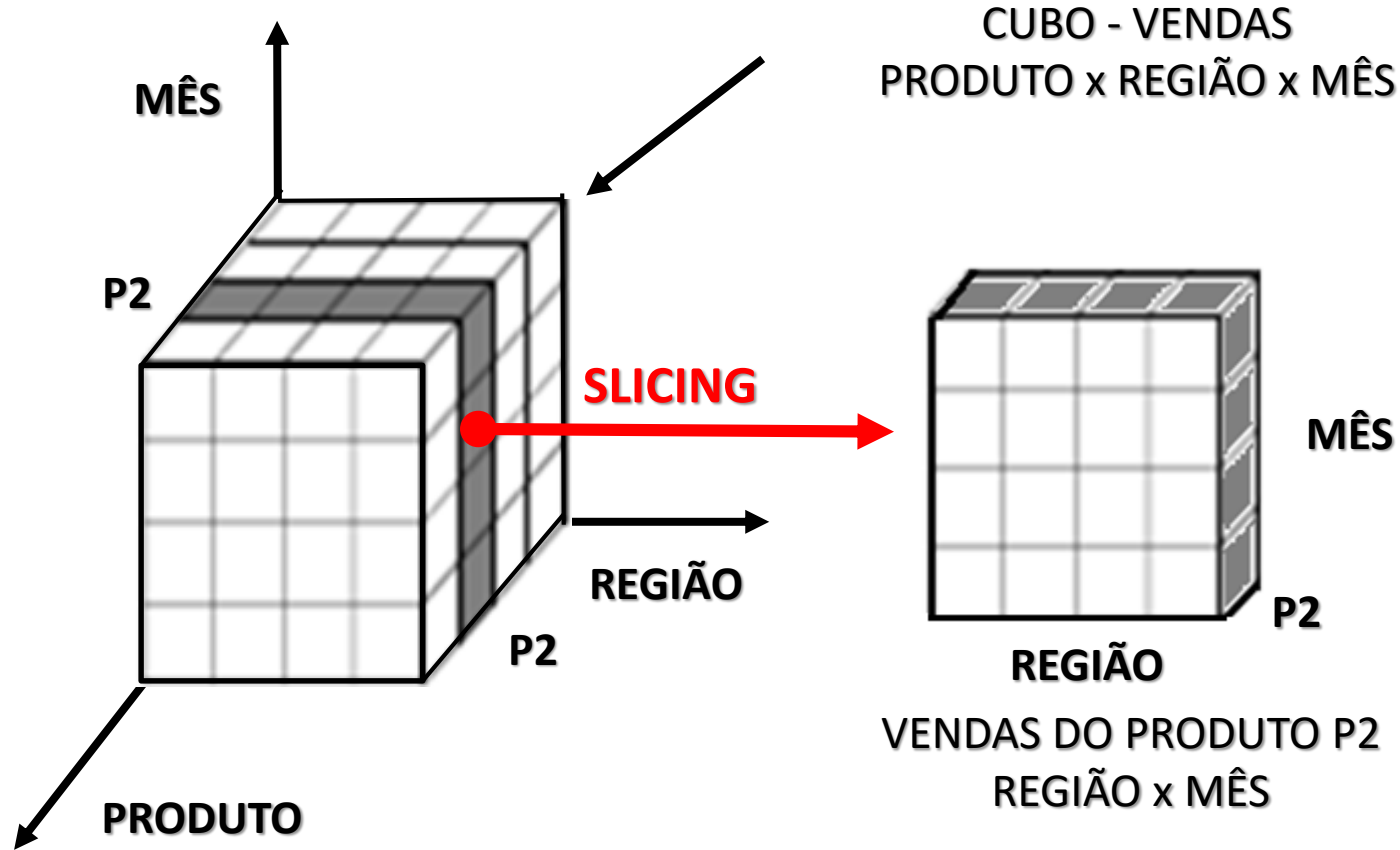
SLICING $\Rightarrow$ FATIAMENTO

DICING $\Rightarrow$ CORTE EM CUBOS

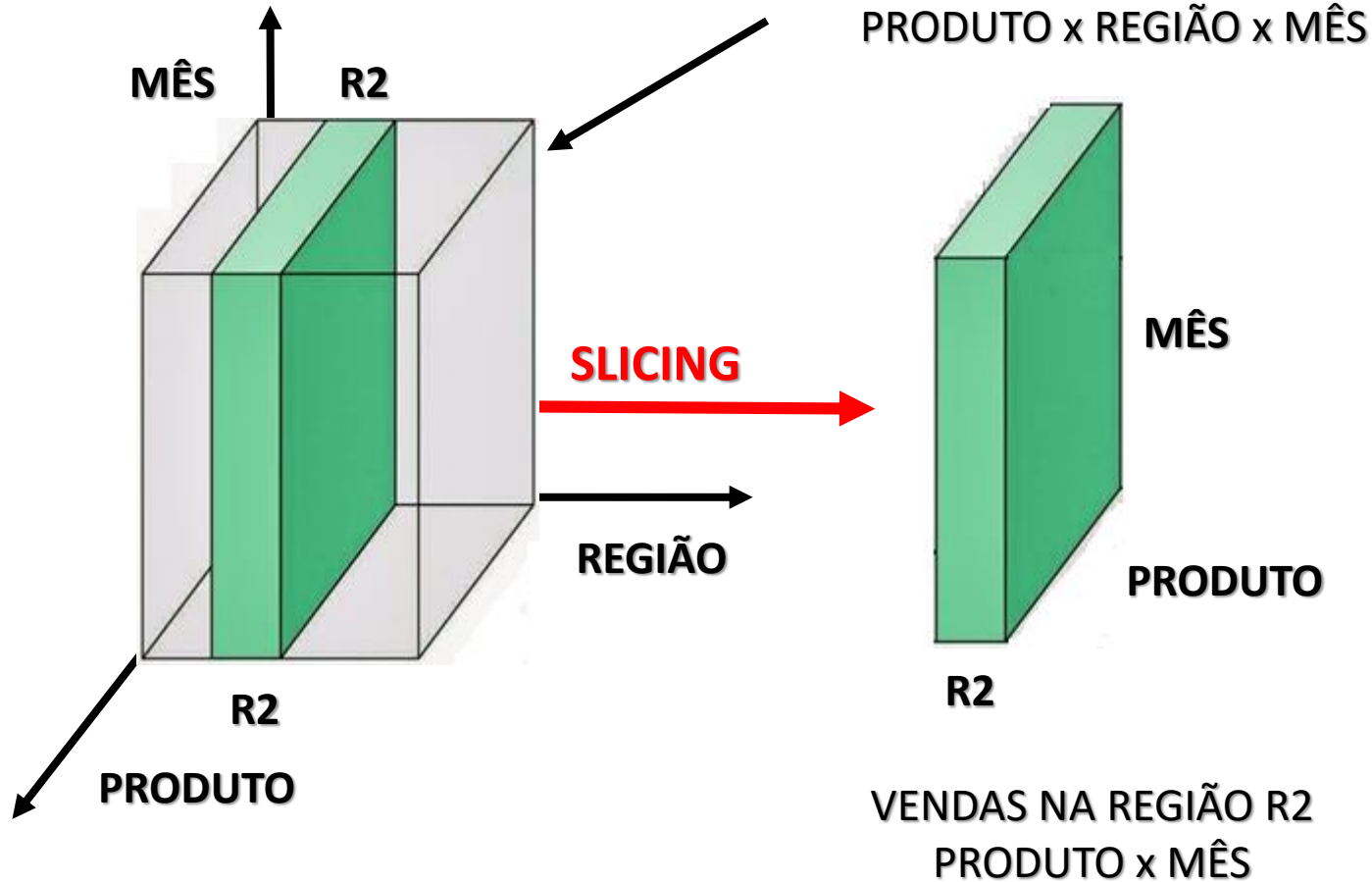
SLICE AND DICE



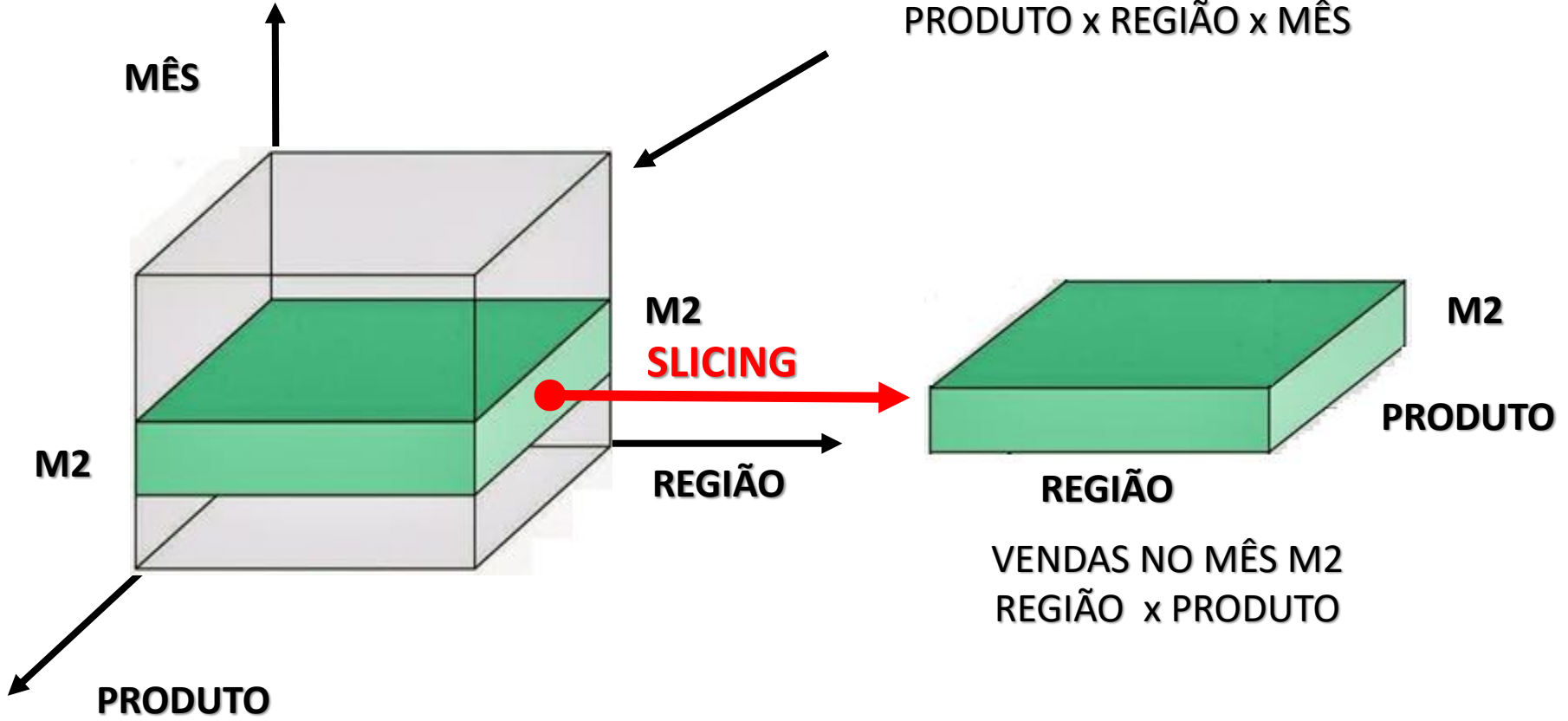
SLICING - FATIAR



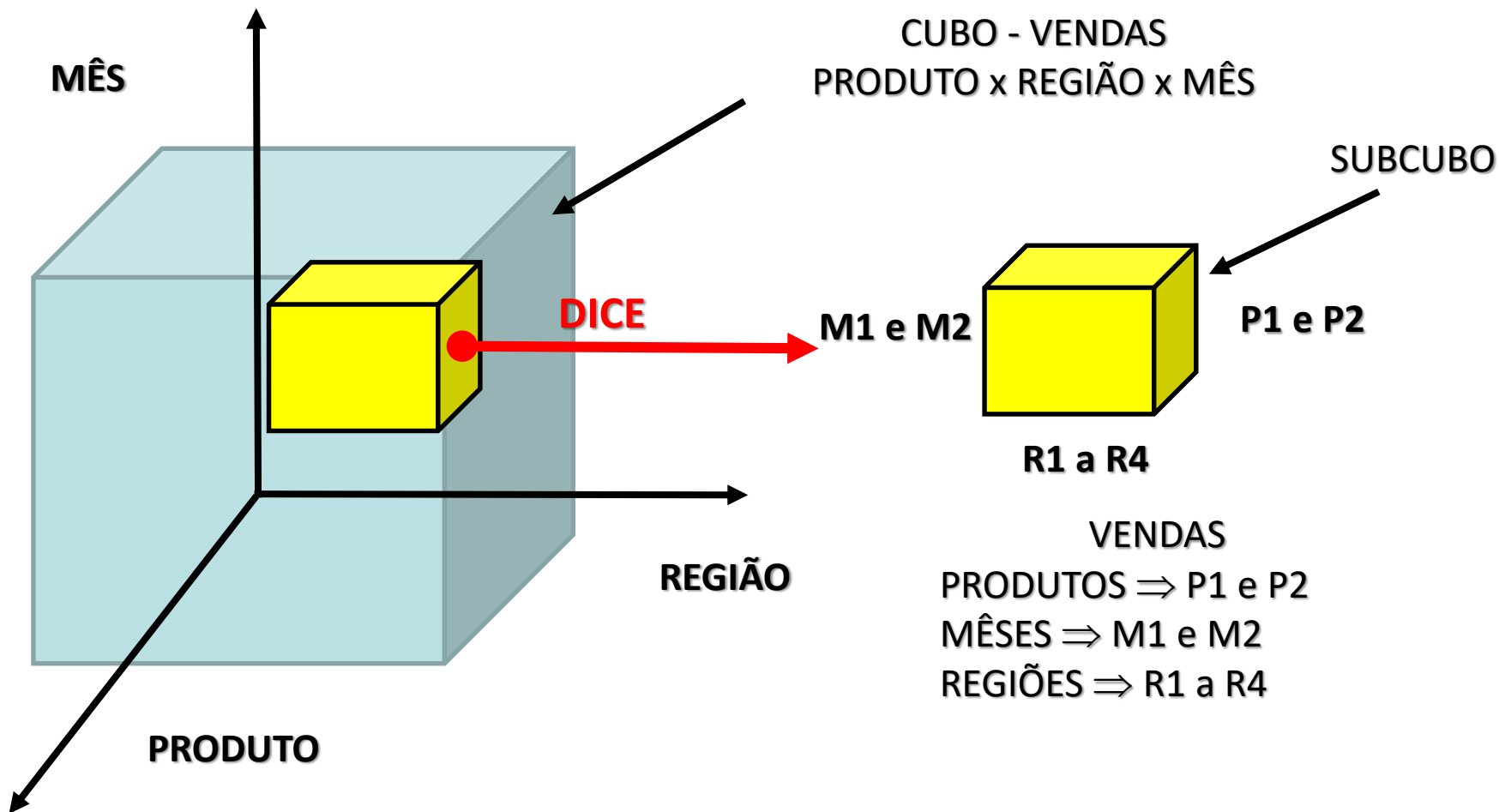
SLICING



SLICING



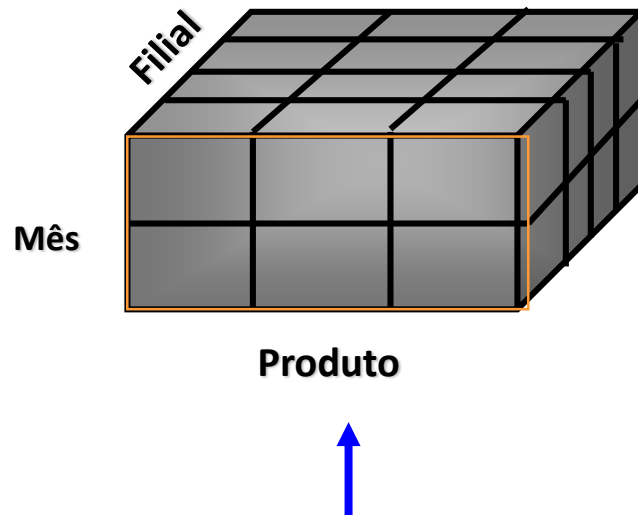
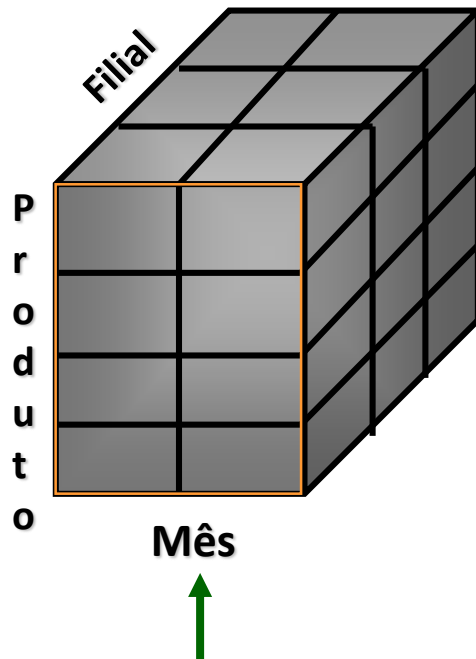
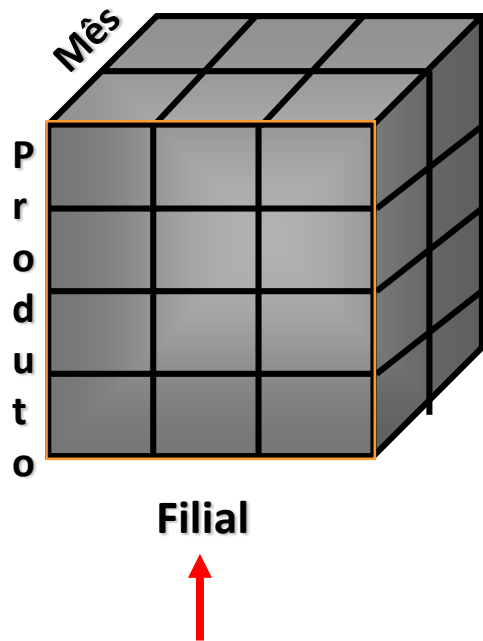
DICE - SUBCUBO



CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

PIVOTEAMENTO (ROTAÇÃO) $\Rightarrow$ Rotaciona os eixos de um cubo de dados, permitindo uma visão alternativa. Ou seja, reorienta a visão multidimensional dos dados, oferecendo diferentes perspectivas de análise dos mesmos.

PIVOTEAMENTO (Rotação)



CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

ALERTAS $\Rightarrow$ são utilizados para indicar situações de destaque em elementos dos relatórios, baseados em condições envolvendo objetos e variáveis. Servem para indicar valores mediante condições, mas não para isolar dados.

RANKING $\Rightarrow$ a opção de ranking permite agrupar resultados por ordem de maiores/menores.

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

FILTROS $\Rightarrow$ os dados selecionados por uma pesquisa (query) podem ser submetidos a condições para a leitura na fonte de dados. Os dados já recuperados pelo usuário podem ser novamente “filtrados” para facilitar análises diretamente no documento.

SORTS (ordenações) $\Rightarrow$ os sorts servem para ordenar uma informação. Esta ordenação pode ser customizada, crescente ou decrescente.

CARACTERÍSTICAS DA ANÁLISE DAS FERRAMENTAS OLAP

BREAKS $\Rightarrow$ servem para separar o relatório em grupos de informações (blocos). Por exemplo: o usuário tem a necessidade de visualizar a informação por cidades, então ele deve solicitar um Break por essa informação. Após esta ação ter sido executada, automaticamente o relatório será agrupado por cidades, somando os valores mensuráveis da pesquisa por cidades.

QUESTÕES DE PROVAS

01- Um Técnico executou uma operação de visualização OLAP que rotacionou os eixos de um determinado cubo, provendo uma visão alternativa dos dados, ou seja, permitindo a visão de suas várias faces. Ele executou a operação

- (A) *slice*.
- (B) *dice*.
- (C) *drill across*.
- (D) *pivot*.
- (E) *roll up*.

02- Um drill down corresponde a ir de um nível mais baixo para um nível mais alto de agregação, ao passo que um drill up permite uma navegação pelas hierarquias em direção contrária.

02- ERRADO

03- Pivoteamento ou rotação é uma técnica para alterar uma hierarquia dimensional para outra em um cubo de dados.

03- CERTO

04- Marque a alternativa que lista operadores analíticos (OLAP) para manipular dados em um cubo.

A) Roll up, slice e star.

B) Roll up, pivot e dice.

C) Star, snowflake e cube.

D) Rank, case-when e list.

E) Slice, group by, e select.

05- Para extrair dados de fontes de dados heterogêneas que irão alimentar um Data Warehouse de forma homogênea e concisa, servindo de base para gerar relatórios e gráficos para apoiar as decisões da gerência da organização, deve-se utilizar um processo conhecido como

- (A) OLAP
- (B) Data Mart
- (C) ETL
- (D) OLTP
- (E) Data Mining

06- Um Técnico executou uma operação de visualização OLAP que rotacionou os eixos de um determinado cubo, provendo uma visão alternativa dos dados, ou seja, permitindo a visão de suas várias faces. Ele executou a operação

(A) *slice*

(B) *dice*

(C) *drill across*

(D) *pivot*

(E) *roll up*

FERRAMENTAS DE UM DATA WAREHOUSE

✓ DATA MINING (mineração de dados) $\Rightarrow$ é o processo de **DESCOBERTA** de **PADRÕES implícitos existentes** em **grandes massas** de **dados**.

Data Mining é **DESCOBRIR CONHECIMENTO novo escondido** em **grandes massas** de **dados** armazenadas em banco de dados.

FERRAMENTAS DE UM DATA WAREHOUSE

DATA MINING

- ✓ Resultados incluem $\Rightarrow$ associações, correlações, sequências, classificações, clustering (grupamento) e previsões.
- ✓ Não tem automatização simples e precisa ser conduzido por uma pessoa, preferencialmente com formação em Estatística ou áreas afins.

FERRAMENTAS DE UM DATA WAREHOUSE

DATA MINING - EXEMPLOS

- ✓ SITE DE VENDAS (armazenar acessos - cookies)
- ✓ WAL-MART (fraldas e cerveja)

DATA **WAREHOUSE** $\Rightarrow$ A **MEMÓRIA** DA EMPRESA !

DATA **MINING** $\Rightarrow$ A **INTELIGÊNCIA** DA EMPRESA !

QUESTÕES DE PROVAS

01- Um Técnico executou uma operação de visualização OLAP que rotacionou os eixos de um determinado cubo, provendo uma visão alternativa dos dados, ou seja, permitindo a visão de suas várias faces. Ele executou a operação

- (A) *slice*.
- (B) *dice*.
- (C) *drill across*.
- (D) *pivot*.
- (E) *roll up*.

02- O uso prático de data mining envolve o emprego de processos, ferramentas, técnicas e métodos oriundos da matemática, da estatística e da computação, inclusive de inteligência artificial.

02- CERTO

03- Quem utiliza o data mining tem como objetivo descobrir, explorar ou minerar relacionamentos, padrões e vínculos significativos presentes em grandes massas documentais registradas em arquivos físicos (analógicos) e arquivos lógicos (digitais).

03- ERRADO

04- A finalidade do uso do data mining em uma organização é subsidiar a produção de afirmações conclusivas acerca do padrão de comportamento exibido por agentes de interesse dessa organização.

02- CERTO

05- No ambiente organizacional, devido à grande quantidade de dados, não é recomendado o emprego de data mining para atividades ligadas a marketing.

05- ERRADO

06- No contexto de data warehouse e ferramentas OLAP, associe os seguintes conceitos às respectivas afirmativas:

Coluna 1 Conceitos

1. Slice 2. Dice 3. Roll up 4. Pivot

Coluna 2 Descrição

() Permite **rotacionar** um cubo de modo a visualizar os dados sob diferentes faces, modificando a orientação dimensional do cubo.

() Produz um subcubo através da seleção de **um subconjunto** de atributos de múltiplas dimensões.

() Produz um subcubo através da seleção de **uma dimensão** em detrimento de outras.

() Realiza operações de **agregação** em um cubo de dados.

Assinale a alternativa que indica a sequência correta, de cima para baixo.

a. () 1 • 4 • 2 • 3

b. () 1 • 4 • 3 • 2

c. () 3 • 2 • 1 • 4

d. () 4 • 1 • 2 • 3

e. () 4 • 2 • 1 • 3

07- A tecnologia OLAP organiza os dados de um banco de dados em estruturas multidimensionais. Essas estruturas são conhecidas como:

- a- redes
- b- matrizes
- c- cubos
- d- árvores
- e- listas

08- É utilizado em empresas para descobrir padrões de comportamento de determinados clientes visando a oferta de produtos ou serviços específicos. Essa estrutura é utilizada no(s):

- a- Sistemas Especialistas.
- b- Business Intelligence.
- c- Sistemas Distribuídos.
- d- OLAP.
- e- Data Marts.

09-

I. Mudança de perspectiva da visão - extração de um subcubo.

II. Corta o cubo mas mantém a mesma perspectiva de visualização dos dados.

I e II correspondem, respectivamente, às operações básicas OLAP

(A) *pivot* e *drill-throught*.

(B) *slice* e *dice*.

(C) *slice* e *pivot*.

(D) *dice* e *slice*.

(E) *dice* e *drill-across*.

10- Pivotamento está relacionado, na representação dimensional, com o conceito de

(A) Rotação de Planos.

PIVOTEAMENTO = ROTAÇÃO

(B) Valores em Cubos.

(C) Valor Pontual.

(D) Valores em Planos.

(E) *Drill-Down* e *Drill-Up*.

10- A

