

# Data Warehousing na prática

DATA WAREHOUSE NA PRÁTICA PARA AMBIENTE VIRTUAL DE APRENDIZAGEM - MOODLE



Sergio Fred Andrade

## DATA WAREHOUSE NA PRÁTICA PARA AMBIENTE VIRTUAL DE APRENDIZAGEM - MOODLE

### SUMÁRIO

#### RESUMO 4

#### 1. APRESENTAÇÃO 5

##### 1.1. Por que data warehouse para esse AVA? 6

##### 1.3. Objetivos 8

#### 2. CARACTERIZAÇÃO SOBRE DATA WAREHOUSE 9

##### 2.1. A Educação a Distância 9

##### 2.2. Ambiente Virtual de Aprendizagem 9

##### 2.4. Sistemas de Apoio à Tomada de Decisão 11

##### 2.5. Ferramentas para Visualização de Informações 11

##### 2.6. O Processo de Data Warehousing 12

##### 2.6.1. Tipo de Implementação de DW 13

##### 2.6.2. Granularidade 14

##### 2.6.3. Componentes de Data Warehousing 15

##### 2.6.4. Composição da Matriz de Barramento 16

2.6.5. Entidades de dimensões e fato 16

2.7. Modelo Multidimensional 17

3. OS PROCEDIMENTOS E MÉTODOS 19

3.2. Formulação dos questionamentos do domínio específico AVA 20

3.3. Modelagem Conceitual e Matriz de Barramento 20

3.4. Consultas OLAP 21

4. DATA WAREHOUSING NA PRÁTICA 22

4.1.1. Banco de Dados de produção do Moodle 22

4.2. Matriz de Barramento como Modelo Conceitual 23

4.3. Modelo Multidimensional para o DW Moodle 26

4.4. O Modelo de Data Warehouse 27

4.5. Visualização por OLAP 28

4.6. Visualização com MS Power BI 29

4.7. Visualização Toad for Mysql 31

5. CONSIDERAÇÕES FINAIS 33

REFERÊNCIAS 34

### RESUMO

O Ambiente Virtual de Aprendizagem (AVA) é um sistema de software que auxilia na disponibilização e acompanhamento de conteúdos para a educação à distância que são informações úteis ao processo de ensino-aprendizagem. Porém seus atores, sejam docentes, gestores e tutores, em geral, estão mais preocupados com as tarefas rotineiras de atribuição de atividades escolares, forum de discussões, aplicação de questionários e outros. Muitas vezes essas informações não são avaliadas, não são recuperadas adequadamente ou não são objetos de análises para extração de conhecimento na busca de melhorias no processo de decisão educacional.

Um exemplo são as avaliações realizadas, que são informações ricas em detalhes e podem indicar assuntos pouco assimilados do programa, comparação de notas por turma ou polo, desvio padrão das notas avaliativas, relação entre nível avaliativo entre alunos de classes ou regiões diferentes, entre outros.

Desta forma, este trabalho auxilia com uma modelagem para análise de dados através do data warehousing ou business intelligence. Para tanto, é desenvolvido um modelo baseado num AVA, o Moodle 4.0, o qual adota o módulo de questionário que representa uma das avaliações mais aplicadas nos ambientes de aprendizagem virtual e fornece condições sobre as questões, respostas do questionário, o processo de avaliação e a pontuação alcançada pelos alunos para extrair dos dados armazenados, itens significativos para apoiar políticas educacionais, planejamento, programas pedagógicos, atividades didáticas e outras práticas avaliativas, em favor da melhor qualidade no aprendizado.

O objetivo é oferecer um processo de desenvolvimento de ferramenta analítica para os técnicos, gestores e professores da EAD no sentido de se conhecer melhor como está o processo de assimilação do conteúdo ministrado, o aprendizado do aluno e a qualidade dos modelos de questionários aplicados como atividade avaliativa.

Como resultado foram demonstradas as etapas para a modelagem num Ambiente Virtual de Aprendizagem com o propósito de ofertar condições melhores de indicadores de desempenho no processo ensino-aprendizagem à distância.

# APRESENTAÇÃO

Com as inovações tecnológicas aplicadas na educação à distância, vários benefícios foram surgindo principalmente nos aspectos à prática pedagógica e da percepção e aprendizado do aluno. Isso reflete no planejamento pedagógico onde o professor vislumbra as condições de aprendizado e organiza os recursos e práticas didáticas para as contribuições cognitivas e para a boa formação do aluno. Nesse sentido, a avaliação é fundamental para conhecer como estão as etapas realizadas do conteúdo programático e o que deve ser replanejado na prática educacional.

A etapa da avaliação deve ser constante no processo de ensino, pois tem uma significativa função de diagnosticar o desempenho estudantil diante do que foi alvo da prática didática-pedagógica, o nível de aprendizado e o que deve ser reconsiderado para o próximo período. Na avaliação observa-se a evolução cognitiva do alunado e a promoção da articulação necessária entre o ensino e o emprego de testes, provas e relatos que os alunos fazem durante o processo de aprendizagem.

Considerando essas premissas, o presente trabalho contempla a técnica da análise de dados, da ciência de dados, para trazer suporte ao processo de decisão com base em informações armazenadas no sistema Moodle (Rice e William, 2006), onde os dados são inseridos, atualizados e removidos, porém muito pouco consultados para provimento em recursos gráficos, dashboards, matrizes e tabelas, com métodos estatísticos e computacionais para o processo decisório.

Sendo assim, foi desenvolvido uma modelagem para análise de dados baseado num AVA, o Moodle – uma plataforma livre de software de apoio às atividades do ensino e aprendizagem à distância. A abordagem adota o módulo de questionário que representa uma das avaliações mais aplicadas nos ambientes de aprendizagem virtual e fornece condições sobre as questões, respostas do questionário, o processo de avaliação e a pontuação alcançada pelos alunos para extrair dos dados armazenados descrições significativas para apoiar o políticas educacionais, planejamento, programas pedagógicos, atividades didáticas e outras páticas avaliativas em favor da melhor qualidade no aprendizado.

Desde os últimos 20 anos, a inteligência de negócios ou Business Intelligence (BI), vem se apresentando como condição fundamental para análise de dados nas tomadas de decisões das organizações, sejam Privadas, Ong's ou Públicas. As principais técnicas que fazem parte do processo de modelagem e aplicação são o Data Warehouse (DW) - modelagem multidimensional, e, o On-line Analytical Processing (OLAP) - ferramentas de consultas analíticas.

Data warehousing demanda incursão de várias etapas complementares, que vai da abstração do domínio de interesse e das especificações dos sistemas produtivos transacionais, passando pela extração de dados, até os relatórios e gráficos informacionais, que são primordiais para auxiliar os administradores e, em especial os gestores educacionais, por exemplo. (FAYYAD, 1996). O OLAP são ferramentas para visualização das informações estatísticas com gráficos, tabelas e dashboard's.

Nesse aspecto, o domínio de um Ambiente Virtual de Aprendizagem (AVA), como o Moodle, tem condições ideais para fornecer dados e informações para o processo decisório, pela análise de seus módulos formadores, como cadastro de alunos, professores e tutores, cadastro de componentes curriculares, fórum de debates, questionários nas diversas modalidades, midiatecas, disponibilização de materiais instrucionais, entre outros. Que é recurso fundamental para este trabalho.

### **Por que data warehouse para esse AVA?**

Considera-se como premissa as ideias de Hoffmann (2003) e Luckesi (2003) de que os educadores brasileiros têm uma concepção da avaliação da aprendizagem relacionada à aplicação de provas e atribuição de notas, atrelada fortemente a uma concepção de sentenças como julgamento de resultados, e por isso, exerce uma avaliação de forma arbitrária, com uma função classificatória e burocrática, em que se dissocia o ato de educar do ato de avaliar.

Outro enfoque mencionado por Laguardia (2007), é de que o uso de questionários é o método mais amplamente utilizado nos diversos tipos de avaliação de cursos online, pela vantagem da praticidade. Porém, existem desvantagens pelas precárias consultas ou interpretações dos dados postados, que podem distorcer as informações importantes.

Nesse sentido, observa-se no AVA que os docentes e gestores do processo de ensino estão mais preocupados com as tarefas rotineiras de atribuição de atividades escolares, forum de discussões, aplicação de questionários e outros, os quais estão revestidos de dados na inserção, atualização e acompanhamento nos sistemas onde os dados ficam armazenados, quando muitas vezes essas informações são deixadas para um segundo momento, não recuperadas adequadamente e não avaliadas, ou mesmo não serem alvos de análises para extração de conhecimento na busca de melhorias no processo decisório para desenvolvimento de melhores programas e recursos didático-pedagógicos.

Diante da metodologia questiona-se: é possível responder com modelo de data warehousing para resultar consultas analíticas, a partir de formulações de perguntas baseadas no domínio do Moodle - um Ambiente Virtual de Aprendizagem?

## Justificativa

Na avaliação observa-se a evolução cognitiva do alunado e a promoção da articulação necessária entre o ensino e o emprego de testes, provas e relatos que os alunos fazem durante o processo de aprendizagem.

As organizações eficientes estão preocupadas com a qualidade e eficiência de suas informações e como estas são usadas nos processos decisórios. Para as escolas, também é assim, quanto mais próximas da especificação dos requisitos pedagógicos as ferramentas do AVA estiverem, melhores serão empregadas no cotidiano dos atores envolvidos e na prática educacional.



Isso justifica a aplicação das técnicas de análise de dados com modelagem de DW e ferramentas de visualização de informações com OLAP, que podem ser úteis para melhorias nas práticas didático-pedagógicas para Educação à Distância (EAD). Especificamente, análise das informações do módulo de questionário que representa uma das avaliações mais aplicadas nos ambientes de aprendizagem virtual e fornece condições sobre as questões, respostas do questionário, o processo de avaliação e a pontuação alcançada pelos alunos para extrair dos dados armazenados descrições significativas, para apoiar as políticas educacionais, planejamento, programas pedagógicos, atividades didáticas.

Desta forma, é apresentada uma análise e modelagem multidimensional, com exemplo de uma base de dados simulada do Moodle, para mostrar estruturas de consultas OLAP contemplando questionários aplicados em uma disciplina com três turmas, num curso técnico de ensino médio em informática.

## Objetivos

### Objetivo Geral

Propor uma abordagem para análise de dados dos questionário de um AVA da modalidade EAD e modelagem pela técnica de data warehousing para facilitar o processo de consultas e visualização sobre informações das avaliações de questionários armazenados.

### Objetivos Específicos

1. Prover abordagem para melhorar o nível de avaliação pedagógica pela análise dos dados de questionários armazenados no software Moodle;
2. Definir a modelagem multidimensional necessária para constituir a base de um DW para análises educacionais a partir de um AVA;
3. Promover construção de cubos multidimensionais para consultas OLAP para indicar possíveis ferramentas de visualização das informações sobre avaliações pedagógicas;



# CARACTERIZAÇÃO SOBRE DATA WAREHOUSE

## A Educação a Distância

A EAD é uma modalidade de ensino e aprendizagem que usa tecnologias da informação online e possibilita novas formas de compartilhamento remoto entre os atores envolvidos na prática pedagógica. Por isso, permite vários aspectos significativos ao ambiente educacional com possibilidade da boa formação do desenvolvimento pessoal e habilidades profissionais importantes para o contexto individual e social.

A conceituação do MEC (2022), diz que a EAD é a modalidade educacional na qual alunos e professores estão separados, física ou temporalmente e, por isso, faz-se necessária a utilização de meios e tecnologias de informação e comunicação. Essa modalidade é regulada por uma legislação específica e pode ser implantada na educação básica (educação de jovens e adultos, educação profissional técnica de nível médio) e na educação superior.

De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira/Legislação e Documentos - INEP (2018), nos últimos anos, a EAD tem apresentado um crescimento contínuo e, pode ser considerada como uma alternativa à inserção educacional de pessoas de diversos contextos e classes sociais distintas.

Segundo Martins e From (2016), a educação, de forma geral, propicia a produção de conhecimento individual e coletivo e, no contexto da EAD, isto é favorecido pelos ambientes digitais e interativos de aprendizagem.

De acordo com Perreneoud (1999), na prática pedagógica é importante redefinir os critérios e procedimentos de avaliação da aprendizagem mediante novas correntes filosóficas e métodos, o que tem sido um desafio para estudiosos e pesquisadores da educação, pois a avaliação está no centro do sistema didático e do sistema de ensino.

A avaliação trata-se de um processo de análise comparativa com pontuações previamente definidas e que são coletadas no sistema de monitoramento do AVA.

## Ambiente Virtual de Aprendizagem

O AVA refere-se ao principal sistema utilizado pela instituição de ensino para disponibilizar aos discentes recursos pedagógicos e tecnológicos envolvidos na relação ensino e aprendizagem e na cooperação entre professor conteudista, professor mediador e discente referente ao conteúdo dos componentes curriculares.

Esse sistema apresenta materiais, recursos e tecnologias apropriadas, que permitem desenvolver a cooperação entre tutores, discentes e docentes, a reflexão sobre o conteúdo das disciplinas e a acessibilidade metodológica, instrumental e comunicacional. Permite a utilização de diversos objetos de aprendizagens, dentre eles pode-se destacar as categorias: textos, ferramentas de orientação, conteúdos, atividades avaliativas e interação.

Com o objetivo de atender ao modelo pedagógico de educação a distância, é utilizada uma plataforma de acesso e funcionamento integral, geralmente via web, a qual garante ao aluno flexibilidade de acesso, considerando-se a esfera temporal (qualquer dia e hora) e a esfera espacial/geográfica (de qualquer local), além de flexibilidade na organização dos estudos. (MEC, 2022)

O AVA é uma plataforma que permite a utilização de diversos objetos de aprendizagens, dentre eles pode-se destacar as categorias: textos, ferramentas de orientação, conteúdos, atividades avaliativas e interação. Segundo a Resolução CNE 01/2016, o AVA apresenta materiais, recursos e tecnologias apropriadas, que permitem desenvolver a cooperação entre tutores, discentes e docentes, a reflexão sobre o conteúdo das disciplinas e a acessibilidade metodológica, instrumental e comunicacional.

# Sistemas de Apoio à Tomada de Decisão

Segundo Sprague (1997), os Sistemas de Apoio à Decisão (SAD) podem ser determinado como sistemas consultivos por computador, mediante dados e modelos, que auxiliam nas tomadas de decisões, relativas aos problemas não estruturados ou semiestruturados. Os SAD's concedem total acesso à base de dados corporativa, estruturação da modelagem de problemas, simulações e possuem interface amigável. Nesse sentido, Singh (2001), menciona dois grandes grupos de SAD: as ferramentas de visualização e Data Warehouse.

## Ferramentas para Visualização de Informações

O conjunto de processos para a criação, gerência e manipulação de dados multidimensionais para análise e visualização pelo usuário é chamado de ferramentas Online Analytical Processing (OLAP). Um conjunto de dados multidimensionais formado por uma ou mais tuplas de dimensões e fato, abstraída por um "cubo". Os cubos, que são representações abstratas de um modelo multidimensional específico, são fundamentais para geração das consultas e relatórios OLAP.

Portanto, cubos são estruturas dinâmicas para consultas online e para gerar gráficos e relatórios não triviais. Podem ter várias dimensões: exemplo de tempo, região, produto, dentre outros. Podem ser manipulados e visualizados por diversos ângulos e diferentes níveis de agregação, onde é possível, através de suas dimensões, proporcionar análises de determinadas situações.

Um cubo é criado conforme o tipo de consulta que se deseja fazer. O projeto de um cubo deve priorizar a escolha e a organização das dimensões que podem influenciar nas medidas relacionais para a tomada de decisão. Ele é implementado separadamente, onde cada cubo é uma visão multidimensional de dados, por uma tabela fato e n tabelas dimensões. O DW é a origem para formação de cubos para visualizações por OLAP.

# O Processo de Data Warehousing

Na teoria de banco de dados as informações são armazenadas com propósito maior de operações de inserção e atualização, com a necessidade de garantir a integridade destes (MACHADO, 2008). Neste contexto, o DW que faz parte dessa teoria serve para gerar consultas OLAP em dados não triviais.

Com a crescente necessidade de desenvolver sistemas de suporte à decisão caracterizado pelo novo ambiente de negócios, surgiu a necessidade de separar o ambiente de processamento operacional do ambiente analítico. Desta evolução natural surgiu em 1992 o conceito de data warehouse como requisito aos emergentes sistemas informacionais. Posteriormente surgiu o conceito de data warehousing abrangendo o conjunto de tecnologias empregadas nestes ambientes (INMON, 1997).

Para Singh e Malhotra (2011) a grande vantagem da escolha do DW está no fato de que o armazenamento de dados se apresenta bastante útil para a prestação de informações coletivas aos seus usuários. Eles também destacam que nos DW's a informação está sempre disponível com muita facilidade por um longo período de tempo.

O data warehouse pega os dados do sistema e melhora a relação entre as variáveis no banco de dados de origem. Em geral, a modelagem de dados tem sido um assunto associado diretamente aos procedimentos dos modelos lógico e físico, aplicando-se de imediato diagramas entidade-relacionamento (COUGO, 1997).

Segundo Kimball e Ross (2002), os elementos do DW estão no esquema estrela, formado pelo modelo lógico entre as tabelas de dimensões e fato, em banco de dados relacional. No domínio de TI, Sistemas de Data Warehousing (SDWs) são utilizados como meio de armazenamento de grande volume de dados para a aplicação de ferramentas de BI, ressaltando Campos (2012, p.1).

Inmon (2008) definiu DW como uma coleção de dados orientada por assuntos, variante no tempo e não volátil. Ele é orientado a assunto, pois é específico em torno dos grandes assuntos gerais da organização, é integrado, pois a integração pode ocorrer ao consistir nomes descritivos, medidas de variáveis, estruturas de dados, atributos físicos de dados entre outros; e não volátil, pois as modificações dos dados já existentes não ocorrem no DW. Existem, também, dois tipos de carga de dados, uma inicial e outra periódica estabelecida, variante no tempo, já que os dados podem ser apresentados em tempos definidos, demonstrando posições históricas das atividades no tempo.

O DW é considerado único para um conjunto de departamentos organizacionais ou para a organização toda. Cada departamento por ter seu DW por assunto setorizado, que é conhecido como Data Mart (DM). O DM é menor e específico e tem os mesmos elementos e recursos que um DW.

### **Tipo de Implementação de DW**

Data Mart (DM) é um subconjunto lógico e físico da área de apresentação do data warehouse (DW). São DW's departamentais, pois têm as mesmas características de um data warehouse comum, e representam determinados processos de negócios. Tornam-se consideravelmente menos onerosos e sua construção é muito mais rápida em relação ao projeto do DW corporativo (BONIFATI, 2001).

Segundo Kimball e Ross (2002, p. 455), DM é um conjunto flexível baseado nos dados com alto nível de granularidade, extraído de fontes operacionais de informação e apresentado em um modelo dimensional. Já na visão de Inmon (2008), DM ou base multidimensional por assunto, oferece flexibilidade aos usuários finais e são contidos no DW central. Deste DW, originam os dados e proporcionam a fonte robusta e útil para os data mart's.

### **Granularidade**

O conceito mais claro sobre granularidade menciona que: "diz respeito ao nível de detalhe ou de resumo contido nas unidades de dados existentes do data warehouse. Quanto mais detalhe, mais baixo o nível de granularidade. Quanto menos detalhe, mais alto o nível de granularidade." (INMON, 2008).

Para Kimball e Ross (2002), declarar a granularidade significa especificar com clareza o que uma linha do relacionamento das tabelas da base de dados do DW representa. A granularidade exprime o nível de detalhes associados às medidas dos fatos, como os dados descritivos e suas mensurações ou quantificações.

São exemplos de declarações de granularidade baixa de muito detalhe: uma nota avaliativa de um aluno de uma determinada turma, curso e determinada carga horária; ou em maior granularidade, uma questão de um questionário de determinada avaliação de unidade de um disciplina, num curso, numa turma e num determinado período avaliativo.

Percebe-se que para cada nível de detalhe um horizonte de tempo é estabelecido. Portanto, há uma relação inversa proporcional ao nível de granularidade e a possibilidade de utilização dos dados para consulta. Ou seja, definindo-se um nível muito baixo de granularidade é possível responder a quaisquer consultas, pois o nível de detalhe na base é maior. Situação que pode ser notado na Figura 1, onde a técnica OLAP de drill-down representa essa orientação com o exemplo de um item de informação com um baixo nível de granularidade e um alto nível de detalhes, adequado aos gráficos com dados individuais, como exemplo de uma nota de aluno ou uma aula ministrada. Na técnica de drill-up que permite agregar dados tornando-se baixo nível de detalhamento com alta granularidade é ideal para totalizações ou elevações em níveis hierárquicos, como média ou desvio padrão da turma em uma avaliação.

Figura 1 - Níveis de dados por granularidade.





## Componentes de Data Warehousing

Para Kimball e Ross (2002) faz-se necessário compreender quais são os componentes que formam um ambiente de data warehousing antes de começar a combiná-los para criar o DW. Os quatro componentes distintos são: sistemas operacionais de origem, data staging area, área de apresentação dedados e ferramentas de acesso a dados. A Figura 2, a seguir, ilustra esses elementos de forma integrada.

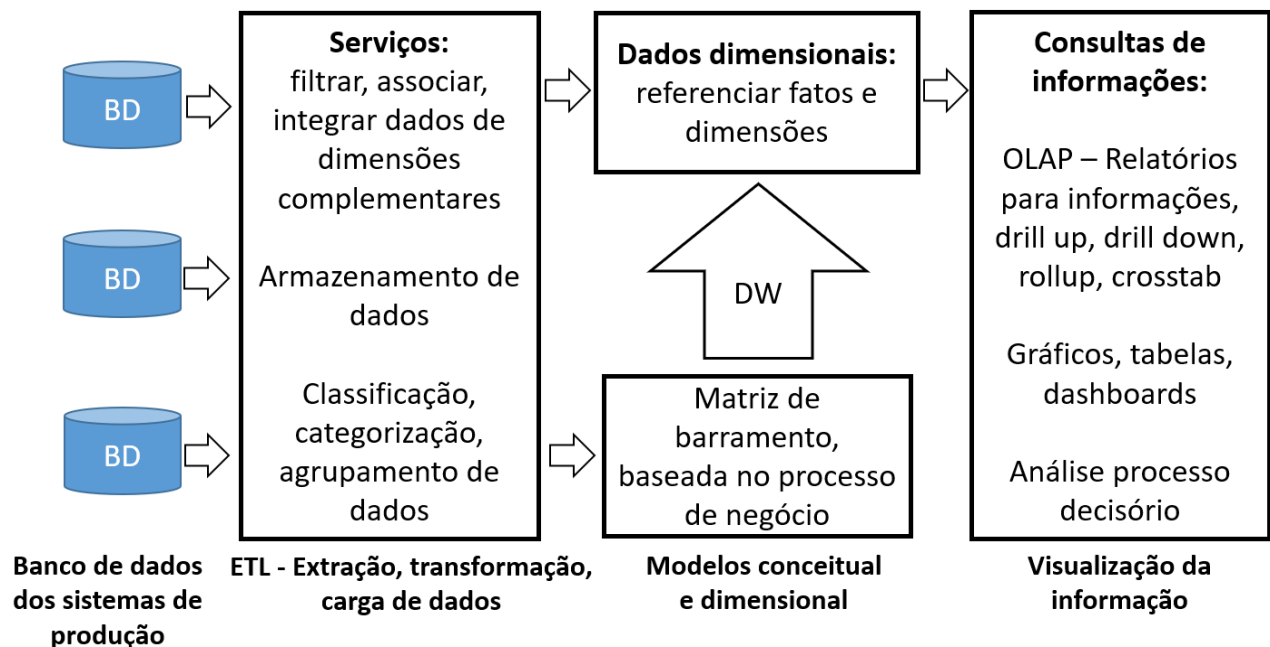


Figura 2 - Componentes do data warehouse.



Os sistemas operacionais de origem são os sistemas de registros, ou transacionais que capturam as transações da organização. Eles devem ser considerados como externos ao DW porque presume-se que se tenha pouco ou nenhum controle sobre o conteúdo e formato desses dados nesses sistemas. Os sistemas de origem também são chamados Sistemas Legados ou Online Transaction Processing (OLTP).

Data staging area é um conjunto de processos que normalmente é denominado de Extract, Transformation and Load (ETL), ou seja, extração, transformação e carga, podendo ser também uma área de armazenamento temporária.

O primeiro passo a ser tomado é a definição das fontes de dados e a extração desses dados. As origens podem ser várias e também em diferentes formatos, onde pode-se encontrar desde dados estruturados nos sistemas transacionais ou Sistema Gerenciador de Banco de dados (SGBD), planilhas, arquivos de multimídia e, dados não estruturados como textos. A transformação dos dados compreende a limpeza, a preparação, correções de erros e integração dos dados para a próxima etapa que é a de carga no DW.

As Ferramentas de acesso a dados na área de apresentação, formam os recursos necessários para os usuários acessarem a base do DW através do conceito da matriz de barramento, usando processo de negócio específico. São vários os recursos existentes nessas ferramentas, que usam a forma dimensional específica e geralmente são baseados em cubos multidimensionais de dados, como exemplo de aplicações de visualização que são ferramentas de acesso às informações pelas técnicas OLAP de drill-up, drill-down, pivot, roll-up, grouping, entre outras.

### **Composição da Matriz de Barramento**

A estrutura da matriz de barramento é binária onde as linhas correspondem aos processos de negócios e as colunas correspondem às dimensões que estão vinculadas diretamente aos mesmos processos. Conhecendo as entidades de dimensões e os questionamentos sobre os processos de negócios, suas estatísticas, avaliações, dados numéricos, volumes e totalizações dos valores inerentes aos processos, pode-se inferir sobre a entidade de fato, os cubos multidimensionais para responder ao questionamento organizacional.

### Entidades de dimensões e fato

As tabelas de dimensões são descritivas da modelagem dimensional e dizem respeito às entidades que servem de perspectiva de análise para o domínio de negócio da organização e devem estar em conformidade com os assuntos escolhidos. As dimensões são os pontos de entrada da estrutura do DW e têm como principal função reunir os atributos que serão utilizados para qualificar e descrever as métricas. São compostas de atributos textuais, que servem também como filtros e como parâmetros dos relatórios. Atributos de dimensões produzem recursos de separação e combinação analíticos eficazes. As dimensões descrevem a interface de usuário nas visualizações OLAP, para DW (KIMBALL e ROSS, 2002).

A tabela de fato é a tabela principal em um modelo DW podendo até existir mais de uma. Essa tabela é responsável pelo armazenamento de valores das medidas, métricas, sumarizações e estatísticas, contendo um grande volume de dados históricos. Em uma tabela de fato uma linha corresponde a uma medição e deve estar direcionada às dimensões relacionadas e todas devem estar alinhadas na mesma granularidade. Essa tabela de fato é relacionada basicamente pelas chaves estrangeiras para atributos qualitativos das tabelas de dimensão.

Faz-se necessário manter a conformidade e coerência entre fatos e as medidas, com o mesmo sentido. Outro cuidado que se deve tomar é na escolha de seus campos mantendo compromisso com a performance.

### Modelo Esquema Estrela

O modelo esquema estrela é a estrutura básica relacional de um modelo de dados multidimensional. Sua estrutura típica possui uma grande entidade central (fato), e um conjunto de entidades menores (dimensões) arranjadas ao redor dessa entidade central, formando a abstração de uma estrela. Apresenta uma estrutura física simples, poucas tabelas (recomendado até 12) e relacionamentos, que promove aos usuários facilidade de entendimento e compreensão dos dados.

O relacionamento entre a entidade fato e cada dimensão é uma simples ligação entre duas entidades com um relacionamento de um-para-muitos no sentido da dimensão para o fato.

## Modelo Multidimensional

A modelagem multidimensional é a abordagem de concepção, estruturação e visualização de um domínio de dados em mais duas dimensões de assuntos que descrevem processos de negócios.

Conceitualmente, o modelo multidimensional está representado pela matriz de barramento que estrutura os processos de negócios e as dimensões do modelo (MACHADO, 2004). Elmasri (2011) comenta que modelos multidimensionais tiram proveito dos relacionamentos inerente nos dados para preencher os dados em matrizes multidimensionais representando cubos de dados. Para os dados que se prestam à formatação dimensional, o desempenho da consulta nas matrizes multidimensionais pode ser muito melhor do que no modelo de dados relacional.

## OS PROCEDIMENTOS E MÉTODOS

Para alcance dos objetivos deste trabalho foram aplicadas as tarefas metodológicas a seguir baseadas na tecnologia de data warehousing. Foi considerado o banco de dados MySQL versão 5.7, sistema Moodle versão 4.0.

O Moodle apresenta diversos módulos e respectivas tabelas no banco de dados para atividades como fórum, questionários, questões, chat, glossário, lições, estatísticas, administração, usuários e outros.

Para demonstrar a abordagem para análise de dados foram considerados os módulos de questionário (quiz), questões (question), notas (gradebook), cursos (courses) e usuários (users).

## Tarefas metodológicas

As seguintes tarefas metodológicas foram adotadas para a modelagem e desenvolvimento do ambiente de data warehouse e apresentação de visualizações para análise de dados.

1. Formulação dos questionamentos do domínio AVA para composição das consultas e análise de dados;
2. Levantamento dos requisitos de modelos multidimensionais, processo ETL e formação de cubos dimensionais de DW;
3. Modelagem conceitual para o processo de dimensionamento do DW através da elaboração da Matriz de Barramento;
4. Determinação dos processos organizacionais, nos módulos do Moodle: quiz, question, grade e user, e operações do processo de aplicação dos questionários e avaliação pedagógica para consultas OLAP;
5. Processo de extração, transformação e carga de dados no DW, para preparação e aplicação com as ferramentas para visualizações OLAP, com as ferramentas Toad for MySQL 7.0 e Ms Power BI;
6. Aplicação das ferramentas OLAP para gráficos, dashboards e tabelas para drill-up, drill-down, pivot, rollap, grouping, crosstab, slice and dice

## Formulação dos questionamentos do domínio específico AVA

Os questionamentos propostos para serem investigados que podem influenciar na qualidade de avaliações num curso EAD são os seguintes:

1. Qual a média de aproveitamento geral do curso?
2. Qual a média de aproveitamento de cada disciplina?
3. Qual disciplina possui a menor média de aproveitamento?
4. Como está o desvio padrão entre as avaliações por questionário?
5. Como está o desvio padrão entre as avaliações por turma?
6. Qual a média de aproveitamento ou desempenho geral de cada polo de ensino?
7. Qual a média de aproveitamento por assunto do conteúdo programático?
8. Qual a média de aproveitamento por professor/formador ou tutor?

9. Qual a média de aproveitamento dos questionários avaliativos?
10. Qual o nível de desempenho entre questionários avaliativos e avaliações presenciais?
11. Qual patamar de evasão escolar de cada polo de ensino?
12. Qual o patamar de evasão escolar em cada semestre ou disciplina?

## Modelagem Conceitual e Matriz de Barramento

O dicionário de dados do banco de dados do Moodle, a descrição do domínio Moodle, as práticas pedagógicas e as operações escolares formam a base de requisitos para a modelagem conceitual multidimensional, representada pela matriz de barramento.

Conhecendo os processos pedagógicos e as dimensões, pode-se chegar aos questionamentos sobre seus dados numéricos, estatísticos, mensurações, volumes e totalizações, possibilitando definir as entidades fatos. As entidades fatos constituem as principais tabelas do modelo multidimensional, pois armazenam dados numéricos com relacionamento direto para todas as dimensões, definindo a granularidade dos dados.

Portanto, para responder aos questionamentos de cada processo pedagógico são necessárias informações provenientes das dimensões e fatos vinculados a ele, assim como, as orientações ao processo decisório das técnicas OLAP que serão utilizadas.

## Consultas OLAP

As operações básicas para visualização por OLAP utilizadas por um modelo multidimensional são: drill-up, drill down, rollup, pivot e slice and dice, grouping e crosstab, conforme Machado (2008). O drill-up e drill down ocorrem quando o usuário aumenta ou diminui o nível de detalhe da informação, alterando o nível de granularidade. O rollup, grouping e crosstab fazem o movimento com agrupamentos das informações, modificando o nível de granularidade das informações, com totalizações, médias, desvio-padrão, contagem e outros.

Slice (ou fatia) é a seleção ou projeção com variação de valores de uma ou mais dimensões. Exemplo: a média das avaliações, da turma 1, no semestre 2, de todas as disciplinas. Dice (dados) é a seleção ou projeção de valores fixos em todas as dimensões. Exemplo: média das notas, no questionário 2, da turma 3, do semestre 2, por disciplina, no curso 1.

Para visualização das informações resultantes das consultas no DW, de rollup, pivot, grouping e crosstab, sendo utilizado o software MS Power BI, Microsoft (2022), que é apresentado com diversos módulos e serviços, conectores com bases de dados e recursos que trabalham para extrair, transformar e carregar dados de um repositório. Os dados deste trabalho são retirados do DW originado das tabelas do Moodle, descritas anteriormente. A escolha do MS Power BI recaiu sobre a facilidade de operação, a diversidade de consultas OLAP possíveis, acesso facilitado para a aplicação, serviços de infra-estrutura em nuvem, painéis, gráficos e dashboards personalizáveis e análise em tempo real para os dados.

Outra ferramenta utilizada para os resultados das consultas OLAP é o software TOAD for Mysql 7.0, Quest (2022), que possibilita que sejam geradas pelos usuários gráficos e grid's de drill-up e drill-down para cruzar informações com agrupamentos, classificações e associações, através de recursos de slice and dice com serviços de alteração da perspectiva da informação de forma visual e serve para alterar a transversalidade da informação por substituição de linhas e colunas nos grid's e matrizes.

## DATA WAREHOUSING NA PRÁTICA

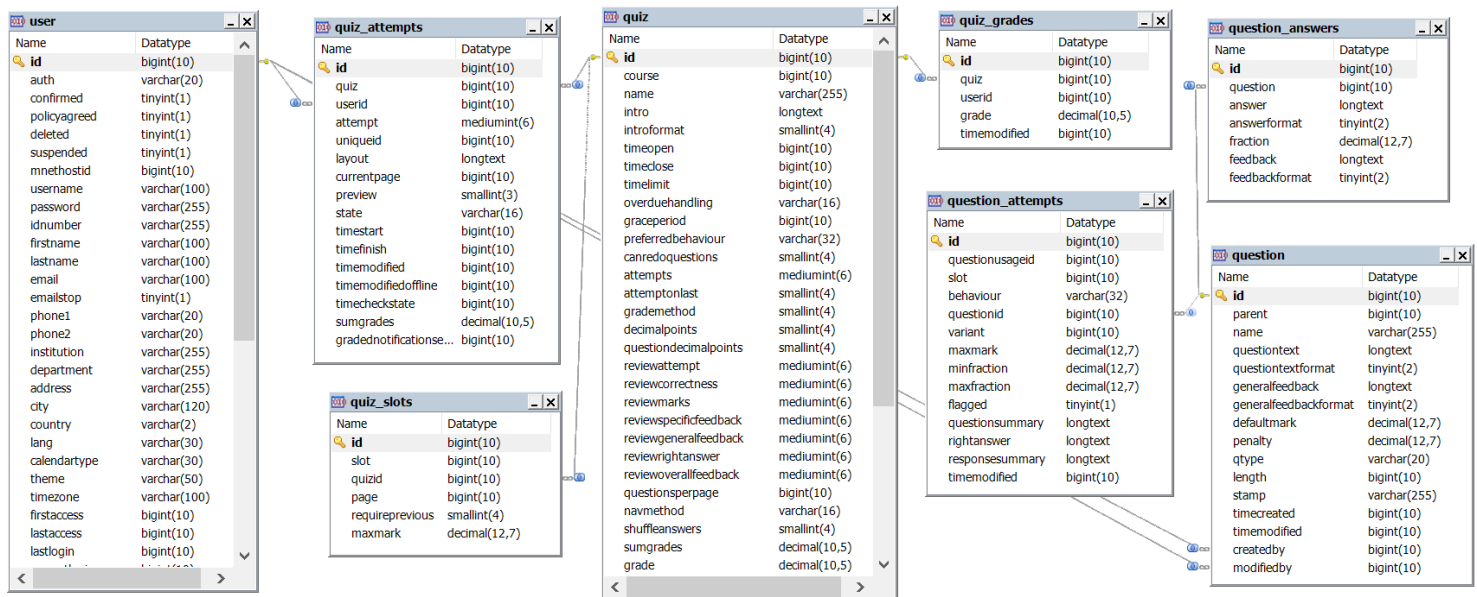
### Banco de Dados de Produção

Trata-se da base de dados do sistema Moodle para cadastros, atualizações e exclusões de participantes – docentes, tutores, alunos e técnicos e dos dados necessários para produção das atividades, interação, disponibilização de recursos didáticos, midiatecas e outros. Como também, das avaliações dos componentes, através de fórum de discussão, questionários, chat's, apresentações e avaliações gerais para o processo de ensino-aprendizagem. Que serviu de origem dos dados para o DW.



Na Figura 3, é mostrado o esquema lógico do banco de dados do sistema Moodle 4.0, que registra todas as operações das atividades de ensino e compartilhamento entre os atores envolvidos.

Figura 3 - Banco de dados Moodle, parte das tabelas dos módulos de avaliação.



### **Banco de Dados de produção do Moodle**

Diante das fontes de dados e da caracterização dos processos escolares, a modelagem conceitual é apresentada através da matriz de barramento. A abordagem pela matriz de barramento é adaptada do modelo tradicional de Kimball et al. (2002), com incrementos direcionados para as operações de OLAP.

### **Matriz de Barramento como Modelo Conceitual**

A matriz de barramento na Tabela 2, é considerada um elemento dinâmico pois serve em todo ciclo de vida do DW até as aplicações das consultas, e deve ser empregada na modelagem conceitual, na multidimensional, na arquitetura do DW, na geração dos cubos dimensionais e na orientação das técnicas OLAP. A granularidade dos dados usada corresponde a uma questão respondida do questionário e sua respectiva pontuação, que é a unidade mínima que pode ser objeto de um processo avaliativo. A matriz de barramento é formada pelos seguintes elementos:

1. Processos de negócios e consultas OLAP: é um grupo de colunas que representa as atividades formadas por operações organizacionais, consultas decisórias, fontes dimensionais e técnicas OLAP.
2. Operações organizacionais: são operações diretas, inerentes às atividades escolares como participantes (professor, tutor, aluno), fórum de discussão, questionários, midiateca, avaliações e outras, cujas atividades de ensino que compõem a estrutura educacional têm suas atribuições definidas internamente. Em geral, as operações organizacionais ou processos escolares dão origem aos data mart's, que são orientados por assunto. As operações organizacionais da matriz de barramento, indicadas a seguir, foram originadas pelo BD do Moodle, que representam atividades escolares: atores participantes; curso, disciplinas, turmas; questionários avaliativos; banco de questões; respostas das questões avaliativas; e, notas avaliativas.
3. Consultas decisórias: são orientações relativas aos processos escolares, que auxiliam nas consultas para tomadas de decisões através das dimensões, fatos e atributos específicos e relacionadas, são descritas com aspectos qualitativos ou quantitativos. Servem para auxiliar o tomador de decisão nas formulações de perguntas que serão respondidas pelas técnicas OLAP em conformidade com os cubos dimensionais.

Exemplo: médias das avaliações, por turma, em determinado período ou unidade, em determinado polo. Neste caso, são indicadas consultas que empregam dimensões e fatos e visualizadas por técnicas em "cubo" como dice, crosstab, drill thought, roll up, pivot e grouping;

1. Fontes: a matriz de barramento traz orientações para utilização das dimensões ou fatos. A depender da consulta requerida, poderão ser empregadas somente dimensões para respostas qualitativas ou, dimensões e fatos para respostas com mensurações e totalizações. Essa orientação ajuda na construção do cubo dimensional com a indicação de quais dimensões, fatos e atributos farão parte do cubo dimensional para emprego da adequada técnica OLAP.
2. OLAP: é a aplicação da técnica OLAP (slice, dice, drill-down, drill-up, rollup, pivot, grouping e crosstab) sugerida para obtenção da resposta requerida de acordo com o processo de negócio e a operação organizacional.
3. Dimensões e fatos: representam todas as dimensões e fatos do modelo estrela. A escolha das dimensões deve refletir as operações organizacionais representadas no Banco de Dado (BD) do Moodle. Recomenda-se os seguintes passos:
4. Identificar a integração entre dados existentes nas tabelas do BD. Como trata-se de um único sistema AVA para uma unidade escolar, representa dados integrados que formam um processo escolar único e cíclico. Esse passo verifica e compartilha informações similares, complementares e/ou relacionadas.
5. Identificar os processos escolares. Para que haja atividade de ensino completa, há um fluxo de informações entre o processo ensino-aprendizagem, os quais estão representados nas tabelas do BD do Moodle. Esses processos são iniciados com aulas expositivas ou síncronas, a disponibilização do material de aula (apostilas, slides, artigos, vídeos e outros), depois a disponibilização de avaliações (fórum de debate, questionários, atividades postadas - timeline, palavra cruzada, podcast e outros), retroalimentação com revisões, pesquisas e replanejamentos, e por fim, as correções e resultados das avaliações qualitativas e quantitativas.

Determinar as dimensões que devem ser originadas das operações escolares. Este modelo implementado pode ser visto na Figura 4. Por convenção, as tabelas são nomeadas usando-se o prefixo “Dim\_” para melhor identificação no esquema estrela e implementação física. Estas tabelas dimensionais são: Dim\_course, Dim\_quiz, Dim\_question, Dim\_user, Dim\_date.

A tabela de dimensão Dim\_date, é uma tabela específica que armazena a série de tempo do histórico de registros contidos no Moodle. Seus atributos são especiais e servem também para auxiliar na granularidade do DW. Já que a granularidade definida corresponde a uma avaliação de questionário individual do aluno em determinado tempo.

1. Determinar a tabela fato: constitui na principal tabela do modelo dimensional, pois representa a definição de granularidade do DW e tem como objetivo o armazenamento dos dados de medição, dados quantitativos, valorativos e aditivos. A tabela de fatos é nomeada pelo prefixo “fat\_”, para facilitar a identificação no esquema estrela. Sua denominação é Fat\_dw\_moodle.

| Processos escolares e consultas OLAP |                                                                                                                                                |                            |                                                                     | Dimensões e fato |            |          |              |          |               |
|--------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|---------------------------------------------------------------------|------------------|------------|----------|--------------|----------|---------------|
| Operações avaliativas                | Consultas decisórias                                                                                                                           | Fonte                      | OLAP                                                                | Dim_user         | Dim course | Dim_quiz | Dim_question | Dim_date | Fat Dw_Moodle |
| Atores participantes                 | Docentes, tutores, alunos que são participantes do processo de ensino-aprendizagem                                                             | Busca em dimensões         | Dice, Crosstab, Pivot, Grouping                                     | X                | X          |          |              | X        | X             |
|                                      | Alunos por turma, alunos por disciplina, alunos por curso, alunos por polo, frequência por turma e disciplina                                  |                            |                                                                     | X                | X          | X        |              | X        | X             |
|                                      | Docentes e tutores por turma, por disciplina, por curso                                                                                        |                            |                                                                     | X                | X          | X        |              | X        | X             |
| Curso, disciplinas, turmas           | Curso e componentes curriculares e alocação em respectivas turmas                                                                              |                            |                                                                     | X                | X          |          |              | X        | X             |
|                                      | Turmas por disciplinas, disciplina e conteúdo programático e carga horária                                                                     |                            |                                                                     |                  | X          |          |              | X        | X             |
| Questionários                        | Modalidade de questionário, modalidade de avaliação                                                                                            |                            |                                                                     |                  |            | X        | X            | X        | X             |
|                                      | Questionário e sua aplicação, questões por questionário                                                                                        |                            |                                                                     |                  |            | X        | X            | X        | X             |
|                                      | Quantidade de questões, tempo de aplicação do questionário                                                                                     |                            |                                                                     |                  |            | X        | X            | X        | X             |
| Banco de questões                    | Questões disponíveis, modalidade de questões, pontuação de questões                                                                            |                            |                                                                     |                  |            | X        | X            | X        | X             |
| Respostas das questões               | Valor da questão, questões acertadas, total de questões acertadas, total da avaliação                                                          | Busca em fatos e dimensões | Dice, Crosstab, Pivot, Grouping, Roll-up, Drill-down, Drill-Through | X                | X          | X        | X            | X        | X             |
|                                      | Média avaliativa questionário, média avaliativa por assunto, média avaliativa por conteúdo programático                                        |                            |                                                                     | X                | X          | X        |              | X        | X             |
| Notas avaliativas                    | Pontuação de corte em notas, Avaliação qualitativa/quantitativa, Faixa de valores avaliativos                                                  |                            |                                                                     |                  |            | X        | X            | X        | X             |
|                                      | Notas por aluno, notas por turma, notas por disciplina, notas por curso                                                                        |                            |                                                                     |                  | X          | X        |              | X        | X             |
|                                      | Média avaliativa questionário, média avaliativa por turma, desvio padrão por questionário, desvio padrão por turma, evasão escolar, reprovação |                            |                                                                     | X                | X          | X        | X            | X        | X             |

Tabela 2 - Matriz de barramento para origem da ontologia de Empresa Comercial.

## Modelo Multidimensional para o DW Moodle

A matriz de barramento específico do DW tem cinco dimensões, um fato e seis processos escolares. A célula de intercepção entre determinada dimensão e o processo escolar vinculado é marcada com "X" quando há uma interdependência entre esses fatores e, portanto, serve para responder questionamentos sobre a operação e alcance de objetivos para as atividades do ensino-aprendizagem.

O resultado encontrado na modelagem do DW do Moodle, representado aqui, é um modelo de cubo dimensões e fato. As dimensões e fato com os seus atributos estão ilustrados na Figura 4. Tendo modelado esse cubo o usuário é capaz de responder às várias perguntas que podem ser feitas para tomada de decisões, como as Consultas Decisórias básicas na matriz de barramento na Tabela 2.

Como exemplo demonstrativo, considera-se um cubo multidimensional de alunos, por polo, período, para determinado questionário por assunto e sua avaliação nas turmas onde foi aplicado. Imagina-se que o professor esteja interessado em saber qual o desvio padrão (dispersão nas notas quantitativas por assunto ou verificar qual turma/polo houve melhor regularidade nas notas), das questões respondidas do questionário em determinado período. Assim, para suporte à decisão pretende-se medir os valores de grade (notas) e fraction (valor da questão), conforme atributos no modelo da Figura 4. Assim, na entidade de fatos os atributos (M) corresponde a:

$$M = \{\text{grade, fraction, rate\_grade, rate\_question}\}$$

Ou seja, imagina-se que um professor quer analisar as informações sobre as notas do questionário considerando quatro dimensões: DATA, ALUNOS, QUESTÃO, NOTAS.

A dimensão DATA (Dim\_date) é descrita pelas propriedades dia, mês e ano. A dimensão do ALUNOS (Dim\_user) tem propriedades nome, sobrenome, cidade polo. A dimensão QUESTÃO (Dim\_question) tem: assunto da questão, questão, resposta da questão, e o fato (Fat\_dw\_moodle) tem propriedade de nota, valor da questão, media das notas, desvio padrão das notas. Assim, obteve-se os atributos (A) das dimensões:

$$A = \{\text{dia, mês, ano, nome, sobrenome, cidade polo, assunto da questão, questão, resposta da questão}\}$$

Cada uma das dimensões é descrita por propriedades específicas e são disjuntas mutuamente e corresponde aos quantitativos no fato respeitando a notação de DW. O cubo multidimensional  $f()$  ficou assim definido:

$$f(\text{Dim\_date}) = \{\text{day, month, year}\}$$



$f(\text{Dim\_user}) = \{\text{firstname, lastname, city}\}$

$f(\text{Dim\_question}) = \{\text{question\_name, questionsummary, responsesummary}\}$

$f(\text{Fat\_dw\_moodle}) = \{\text{grade, fraction, rate\_course, rate\_question}\}$

Nota-se que as características desse cubo do DW Moodle entre os relacionamentos de  $f()$  satisfazem à definição para responder ao questionamento do tomador de decisão.

## O Modelo de Data Warehouse

A Figura 4 representa o modelo físico do DW. Através desse esquema estrela observa-se a entidade fato como elemento central e as cinco dimensões criadas. As respostas para as possíveis perguntas consideram o referenciamento entre as tabelas de dimensões-fato e poderão ser acessadas através de junções conforme o que se deseja ilustrar nas visualizações de OLAP, ou seja, os relatórios e gráficos serão extraídos das cinco dimensões integradas ao fato.

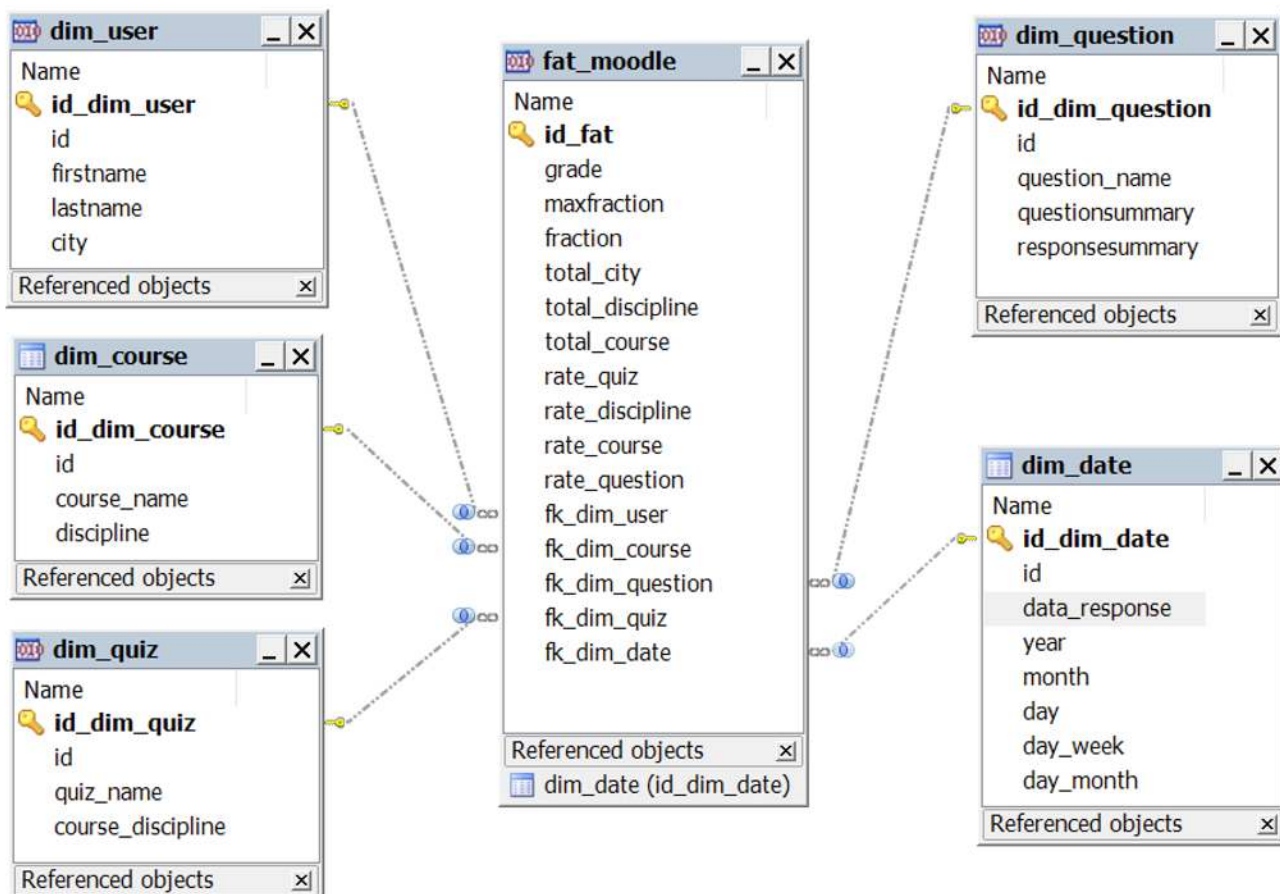


Figura 4 - Esquema de DW

Esse modelo representado de DW servirá de suporte geral onde pode ser fonte maior das informações das áreas envolvidas. Esse modelo deve ser capaz de responder a questões gerais de todos os setores do domínio – avaliações no Moodle, fazendo as combinações necessárias entre o fato e as suas dimensões. A construção do modelo DW pode ser unificada por diversos DM modelados conforme o assunto que se quer abordar, conforme os processos organizacionais na matriz de barramento.

## Visualização por OLAP

A visualização OLAP foi representada pelas ferramentas MS Power BI e Toad for MySQL. Com MS Power BI foram utilizados os serviços internos de Key Performance Indicators (KPI)'s e dashboard's para recursos de slice, dice, crosstab e grouping (Figura 5) que são recursos para gráficos com dados que medem a acurácia e desempenho dos processos escolares. São indicadores que usam técnicas estatísticas para totalizações, médias, medianas, desvio padrão, resumo de dados comparativos, correlações, máximo e mínimo quantitativos, entre outros.

A ferramenta Toad for MySQL disponibiliza um módulo para OLAP - o Pivot Grid, que é específico para consultas e análises de dados dinâmicos diretamente das tabelas de dimensões e fato do DW. Permite recursos de drill-down, drill-up, pivot e rollup (Figura 5) através de um grid que possibilita crosstab ou conversão entre linhas e colunas do grid, oferecendo informações resumidas e detalhadas entre dados articulados pelo usuário com construção de gráficos e dashboard's.

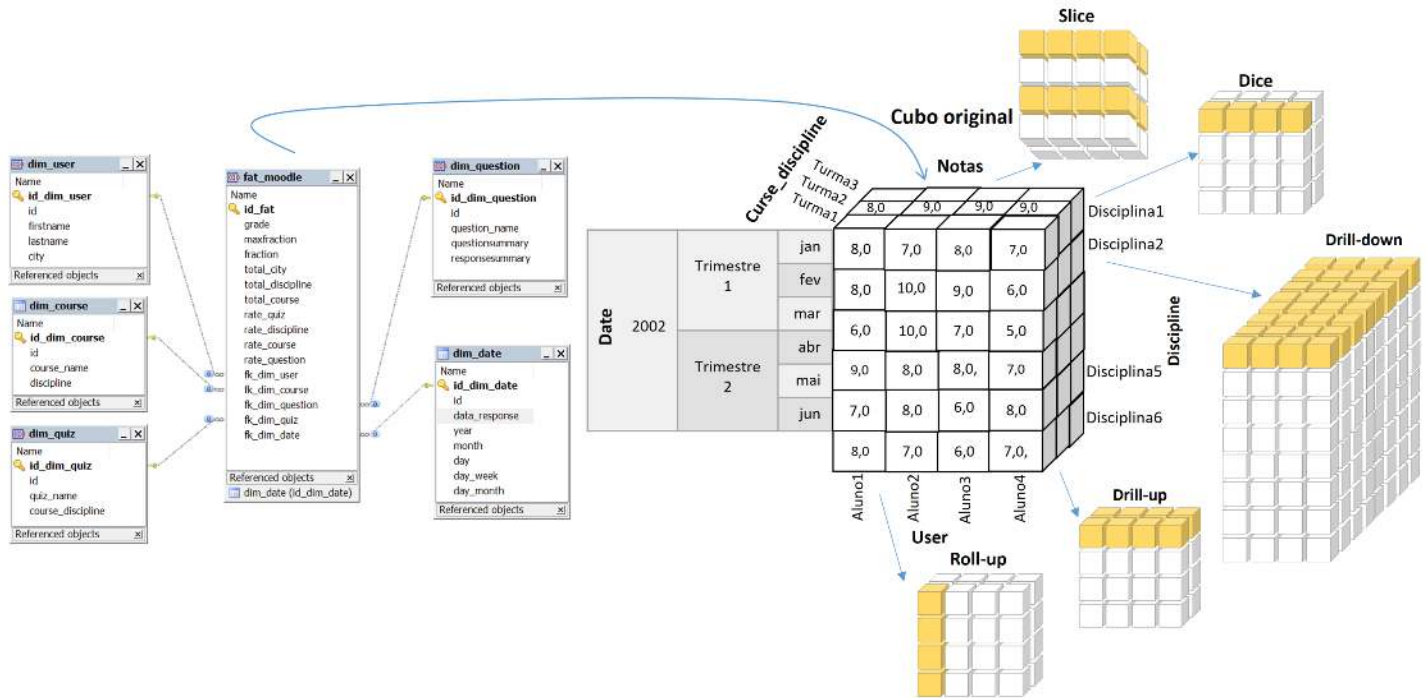


Figura 5 – Origem dos dados do DW e formação de cubos para OLAP.

## Visualização com MS Power BI

Na Figura 6 foram utilizadas técnicas OLAP para construção dos gráficos dashboard's "Média de grade por prova" que mostra turmas diferentes da mesma disciplina e as médias por notas, ilustrando o nível quantitativo das notas por turma e possibilitando a comparação do entendimento do conteúdo programático ensinado. No gráfico de barras "Desvio padrão de fraction por assunto e prova", mostra o desvio padrão com a dispersão nas notas quantitativas por assunto, na mesma prova e em turmas diferentes, permitindo uma análise mais acurada do processo ensino-aprendizado considerando os assuntos ministrados e aprendidos.

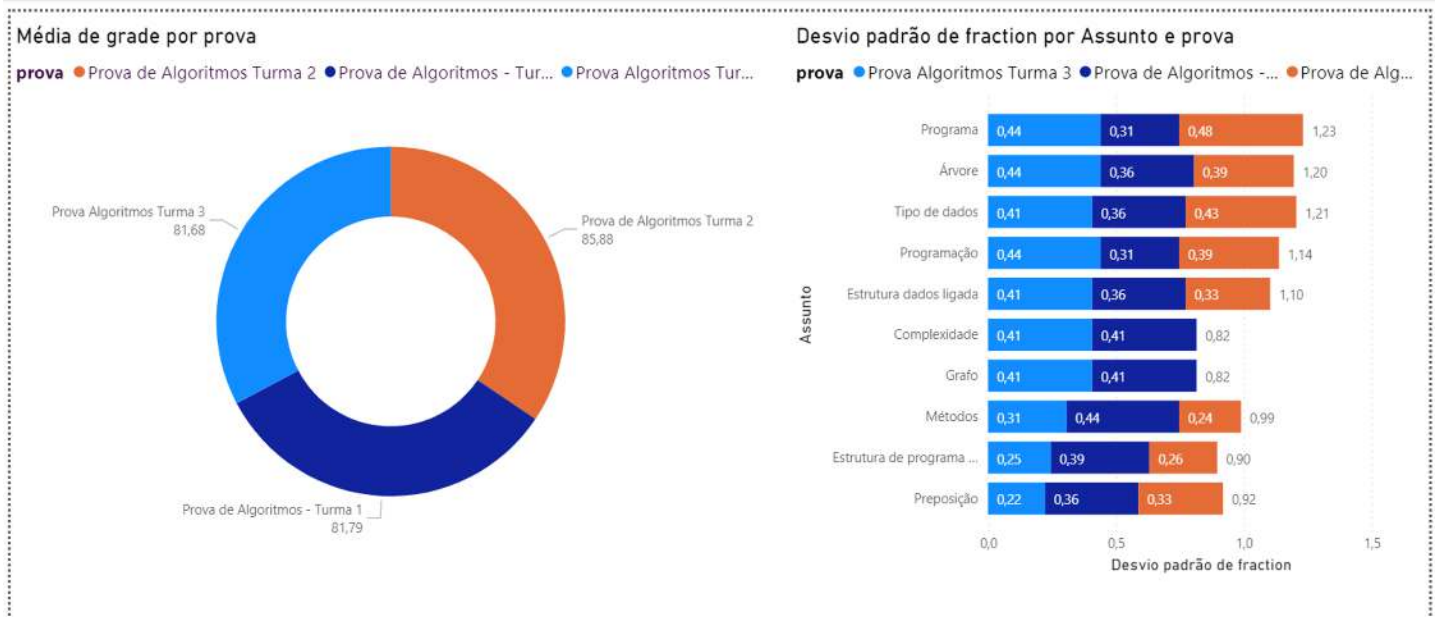


Figura 6 – Gráficos para média de avaliação por prova.

Consultas OLAP foram empregadas para as técnicas de slice, dice, crosstab e grouping com dashboard gráficos. Esses gráficos permitem análise dados originados das dimensões ou indicadores de desempenho para monitoração de processos escolares. O propósito é permitir aos docentes e gestores da educação ter uma visão mais ampla dos processos para auxiliar em replanejamentos, alterações nos conteúdos programáticos e aplicação de novas modalidades avaliativas.

Na Figura 7 é mostrado um dashboard da “Média de notas por polo da EAD” através de um gráfico “tunel” apresenta um comparativo entre as médias avaliativas por turma e disciplina em determinada avaliação. Ao lado a tabela “Desvio padrão de notas”, que trata de um recorte gerado da técnica slice com a dispersão em desvio padrão por prova em determinada turma para um assunto do conteúdo programático. Desta forma, o docente pode ter ideia onde é necessário replanejar ou mudar a estratégia educacional, para rever o processo ensino-aprendizado, com noção clara qual o Polo da EAD precisa de maior reforço educacional.



Figura 7 – Gráfico em tunel para avaliações do aprendizado.

Outros dashboard's de OLAP estão ilustrados na Figura 8 com os gráficos “Média de notas por assunto” e “Média de notas por polo EAD”. Observa-se que são gráficos que oferecem ampla visão do desempenho da turma por assunto ensinado e das turmas no Polo de EAD. Como as técnicas slice, dice e grouping são recortes de parte do cubo dimensional, pode-se fazer outras modalidades de consultas com diversas granularidades e combinações de dimensões. Como notas por questões, notas por avaliação, média geral ou média por Polo da EAD, desvio padrão das médias ou das notas por turma, comparações e correlações nos resultados avaliativos.





Figura 8 – Visualização OLAP por médias avaliativas.

## Visualização Toad for Mysql

Na Figura 9 é exemplificado um grid com drill-down de dados com turma da disciplina, em Polos de EAD e média avaliativa das notas. Podendo ser aumentada a granularidade por região ou por determinado período avaliativo. Ao lado um dashboard de radar linear com a dispersão da média das notas de turma e polo possibilitando diversas análises e tomadas de decisão.

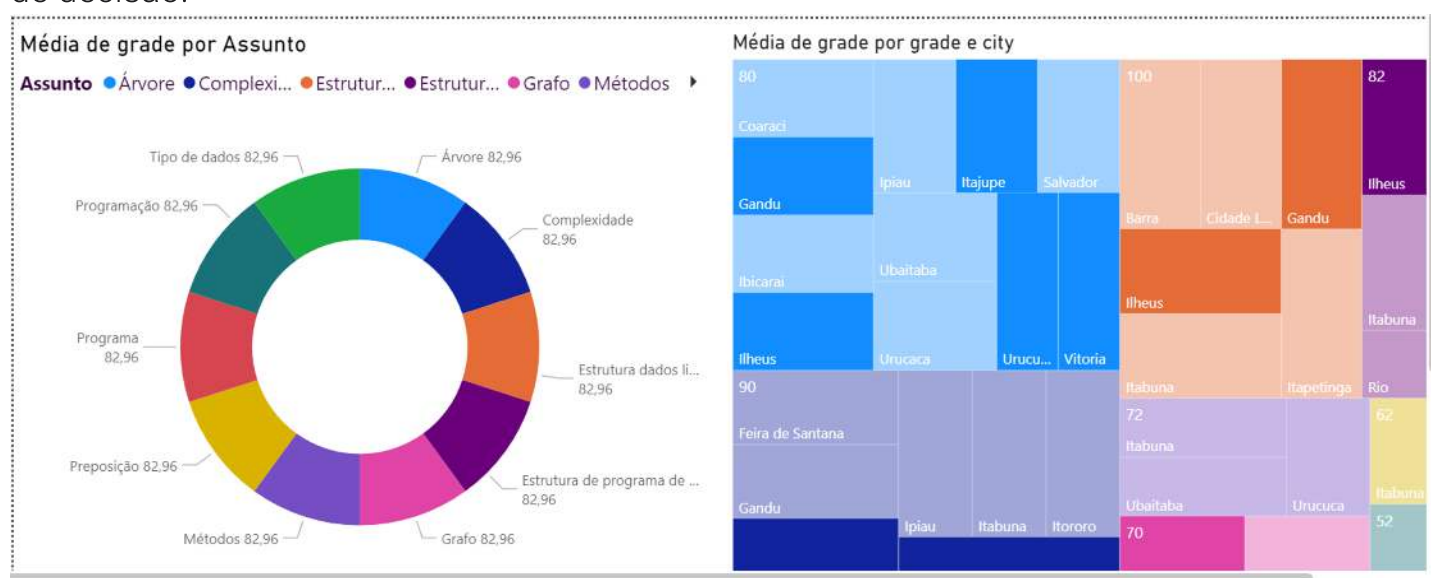


Figura 9 – Visualização OLAP do Toad for Mysql.

Na Figura 10 é mostrado um grid com drill-up por disciplina, turma, assunto do conteúdo programático ensinado e a respectiva nota avaliativa por questão. O que permite análises do assunto lecionado e do aprendizado numa comparação entre as turmas e o Polo da EAD sob outro aspecto visual.

Esses dashboard's oferecem um sistema de retroalimentação para acompanhar o progresso e a obtenção dos resultados, com informação sobre o desempenho das notas avaliativas para revisão da estratégia educacional. Esses recursos trazem as vantagens de: acompanhar a performance de indicadores; oferecer a visualização inteligente dos processos; transparência na análise dos resultados; e, auxiliar o processo gestão.



Portanto, é mostrado grid como origem dos dados, por disciplina, assunto do conteúdo programático ensinado e a respectiva nota avaliativa por questão e possibilita interpretação em conjunto com o gráfico apresentado em drill-up, ao lado direito da tela sobre o assunto lecionado, a disciplina, a turma, o questionário aplicado e as questões avaliativas, por Polo da EAD.

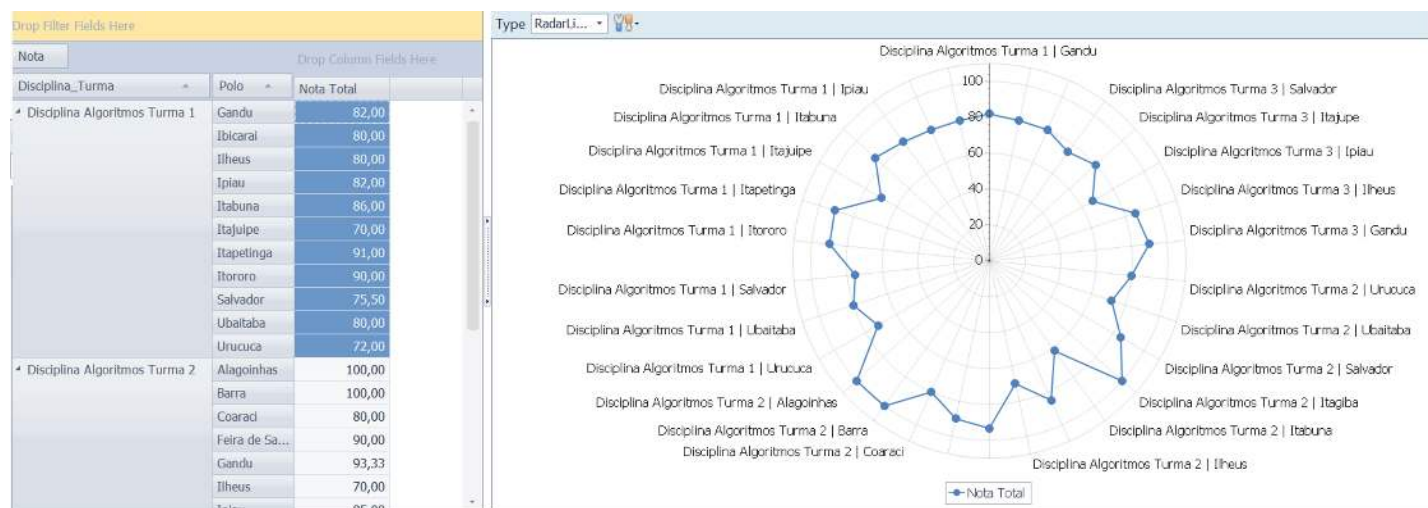


Figura 10 – OLAP drill-up por disciplina e assunto ministrado.

## CONSIDERAÇÕES FINAIS

O desenvolvimento desse trabalho busca contribuir para a aplicação da modelagem multidimensional nos processos escolares num AVA, com o Moodle 4.0 e respectivo banco de dados, para construções de modelos para data warehousing e análise de dados OLAP.

A modelagem teve o propósito de mostrar o entendimento da formação das entidades de dimensões e fato, em razão da vinculação das atividades do ensino com os processos de ensino, em especial, os resultados avaliativos originados da aplicação dos questionários.

A metodologia indica a aplicação de uma matriz de barramento adaptada, para definição dos cubos dimensionais e considera os processos das atividades de ensino-aprendizagem na EAD. Para isso, foram adotados os passos e roteiros do processo de modelagem de DW para elaboração do modelo lógico multidimensional, implementação física e visualização por OLAP.

Em suma, espera-se que este trabalho contribua com um modelo multidimensional onde se privilegia os processos do ensino-aprendizagem e análise de dados da EAD, com indicação de construção de consultas mais próximas aos requisitos dos gestores educacionais e professores, para auxiliar em planejamentos e na condução da formação educacional esperada.

## REFERÊNCIAS

BONIFATI, Angela et al. Warehousing workflow data: Challenges and opportunities. In: VLDB. 2001. p. 649-652.

CAMPOS, Sandir Rodrigues. Validação de dados em sistemas de data warehouse através de índice de similaridade no processo de etl e mapeamento de trilhas de auditoria utilizando indexação ontológica - 2013. XVII, 156 p: il.; 30 cm. Dissertação (Mestrado) - Universidade de Brasília, Faculdade de Tecnologia. Departamento de Engenharia Elétrica, 2012.

CNE - Resolução CNE/CEB nº 1, de 2 de fevereiro de 2016 - Define Diretrizes Operacionais Nacionais para o credenciamento institucional e a oferta de cursos e programas de Ensino Médio, de Educação Profissional Técnica de Nível Médio e de Educação de Jovens e Adultos, nas etapas do Ensino Fundamental e do Ensino Médio, na modalidade Educação a Distância, em regime de colaboração entre os sistemas de ensino.

COUGO, Paulo Sérgio (1997). Modelagem conceitual e projeto de bancos de dados. Campus.

ELMASRI, R.; NAVATHE, S. R. Sistemas de Banco de Dados, 4ª ed., Pearson Addison Wesley, São Paulo-SP, 2011.

FAYYAD, U.; Piatetsky-Shapiro, Gregory; Padhraic, Smith (1996). From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence. Menlo Park, CA, USA, novembro, p. 1 – 34.

HOFFMANN, Jussara. Avaliação: mito e desafio: uma perspectiva construtivista. 2003.

INMON William H. DW 2.0 - Architecture for the Next Generation of Data Warehousing. With Derek Strauss and Genia Neushloss, Elsevier Press, 2008.

KAMPFF, Adriana Justin Cerveira. Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente. 2009.

KIMBALL, Ralph; ROSS, Margy. The Data Warehouse Toolkit. Guia completo para modelagem. Ed. Wiley, 2002.

LAGUARDIA, Josué; PORTELA, Margareth Crisóstomo; VASCONCELLOS, Miguel Murat. Avaliação em ambientes virtuais de aprendizagem. Educação e pesquisa, v. 33, p. 513-530, 2007.

LUCKESI, Cipriano Carlos. Formação do educador sob uma ótica transdisciplinar. Revista ABC EDUCATIO, São Paulo: Editora CRIAP, v. 4, n. 29, p. 1-17, 2003.

MACHADO, Felipe Nery Rodrigues. Banco de Dados: Projeto e Implementação. São Paulo: 2004.

MACHADO, Rosa Teresa Moreira. Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. Rev. Adm. Pública [online]. 2008, vol.42, n.3, pp. 495-528. ISSN 0034-7612.

MARTINS, Karine; FROM, Danieli Aparecida. A importância da educação a distância na sociedade atual. 2016.

MACHADO, Felipe Nery Rodrigues. Tecnologia e projeto de Data Warehouse: uma visão multidimensional. 4. ed. São Paulo: Érica, 2008.

MORENO, Jomara Ramos et al. Ambiente de apoio à tomada de decisão no âmbito da gestão de educação a distância. 2022.

PERRENOUD, Philippe. Formar professores em contextos sociais em mudança: prática reflexiva e participação crítica. Revista brasileira de educação, v. 12, n. 5-21, 1999.

RICE, William; WILLIAM, H. Moodle. Birmingham: Packt publishing, 2006.

SPRAGUE, R. H. Decision Support Systems. North Holland, n. 3, p. 197-200, 1997.

SINGH, Harry S. Data Warehouse: Conceitos, Tecnologias, Implementação e Gerenciamento. 1ª ed. São Paulo: Makron Books, 2001.

SINGH, S.; MALHOTRA, S. Data Warehouse and its methods. Journal of Global Research in Computer Science, India, v. 2, n. 5, p. 113-115, maio 2011.

## Sergio Fred Andrade

PhD in Science, Master in Systems and Computing, specialist in Advanced Computing and Distance Education, Bachelor in Information Systems and Mathematics. He has experience in computing, analysis and information systems, database, data science, and software engineering.

