

Instituto INFNET

**Pós-graduação - MIT em Engenharia de Bancos de dados
com Oracle, SQL Server e Cloud DB**

Modelagem de bases de dados

Projeto

Rossana Santos

Introdução

Este projeto entrega o processo de modelagem de um banco de dados relacional. Foi escolhida uma situação verdadeira, onde nomes, números e demais referências são mascaradas (anonimizadas).

Resumo

A empresa Educredita provisiona crédito exclusivamente para área educacional. Já está no mercado há mais de 25 anos, sendo uma financeira desenvolvida e bem sucedida:

- sua carteira tem cerca de 75 mil alunos (clientes)
- sua principal parceira, Business To Business (B2B), é Instituição de Ensino (IE) privada; total de 120 e estão distribuídas pelo Brasil.
- possui próprio fundo de investimento como originador de crédito, assim como outros originadores parceiros, sejam bancos tradicionais ou outros fundos de investimento (Fundo)s.
- vários produtos especializados, dependendo do aluno, IE e do Fundo.

Até o momento a Educredita tem um banco de dados (PostgreSQL) que é usado para análises da área de negócio (dados analíticos - DW).

Os dados operacionais estão misturados neste mesmo banco de dados e a visibilidade para tratamento destes dados está extremamente complexa, além de muitas das operações serem realizadas de forma manual.

A Educredita pretende escalar seus produtos de forma a atingir 2 milhões de alunos em 300 IEs, nos próximos 4 anos e tornou-se imprescindível criar um banco de dados operacional que continue abastecendo o banco de dados analítico (DW), e que também permita o escalonamento para atingir a meta.

Requisitos

Devem ser cadastrados:

- IEs
- Fundos e Bancos
- Alunos e Fiadores
- Produtos e suas regras de distribuição

Haverá registro e estado de operações:

- propostas originadas pelos alunos (simulações)
- contratos (proposta aceita) e
- títulos (parcela) gerados a partir de contrato.

Questionamentos

Para construção do novo banco de dados, incluindo os processos de construção do Modelo Conceitual, do Modelo Lógico e então do Modelo Físico foram deslocados seis profissionais, a saber:

Rossana

Fabício

Elenice

Inácio

Telma

Adriano

Formando assim a equipe **RFEITA**.

Nas várias interações com a área de negócio foram levantadas algumas questões:

1.) Quais são os dados coletados para uma pessoa ser um cliente?

R. Ser Pessoa Física (PF), CPF, nome, data de nascimento, fone, e-mail, endereço, renda mensal

2.) Se o cliente-Aluno é menor de idade ou não tem renda mensal suficiente, há alguma outra forma dele vir a ser um cliente da Educredita?

R. Sim. Se for menor de idade, deve apresentar um Responsável (cliente-R). Se não possuir renda suficiente, deve apresentar um Garantidor (cliente-G) que componha a renda. Se o cliente-R não possuir renda suficiente, ainda poderá incluir um cliente-G. Obs. A renda composta pode ter no máximo 2 clientes.

3.) Quais são os dados coletados de uma IE?

R. Ser Pessoa Jurídica (PJ), CNPJ, nome, cursos, modalidades de ensino (Fundamental, Médio, Superior, Especialização, etc), disponibilidade de vagas, formato (presencial, EAD, híbrido).

4.) Quais são os dados coletados para um fornecedor de crédito?

R. Ser Pessoa Jurídica (PJ), CNPJ, identificação como banco ou como fundo, dados para a geração de contrato e geração de títulos.

5.) Quais são os períodos mínimos e máximos de um contrato?

R. Todo contrato tem apenas 6 (seis) meses de duração e 12 títulos a serem pagos. Caso o cliente seja um bom pagador, haverá renovação de negócio e o próximo contrato terá as respectivas parcelas cobradas (títulos) com datas posteriores ao término das parcelas do contrato anterior. Por exemplo, um primeiro contrato financia um semestre e é quitado em um ano. Caso renove o negócio com a Educredita, pode financiar o próximo semestre, cuja primeira parcela será posterior à última do contrato anterior.

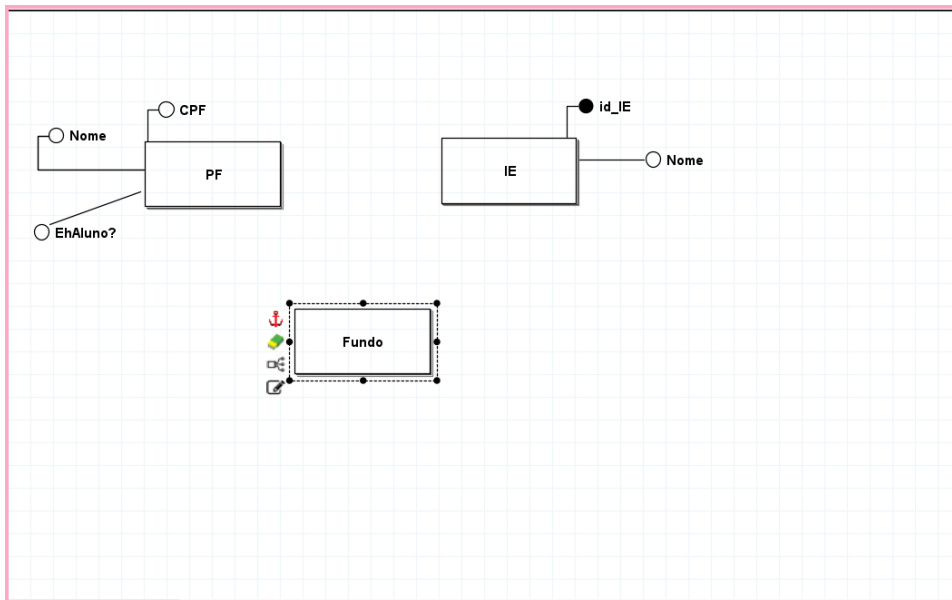
Num exemplo simples de produto Educredita, um curso de 5 anos (10 semestres) terá 10 contratos, 120 títulos e levará 10 anos para ser quitado. O aluno termina o curso e continua pagando as parcelas faltantes por um igual período no qual estudou.

6.) Há algum tipo de controle sobre alunos em atraso com suas parcelas?

R. Sim. De 20 a 30 dias de atraso, as pessoas envolvidas no contrato (cliente-Aluno, cliente-R, cliente-G) passam a ser negativadas nos organismos responsáveis por isto (tipo Serasa e Boa Vista). Antes disto acontecer são disparadas comunicações (SMS, e-mail) alertando-os sobre o atraso e suas implicações. A quantidade de dias de atraso também entra no cálculo da provisão de risco no crédito (PDD, Provisão para Devedores Duvidosos).

Obs. O cálculo da PDD não ficará neste banco operacional, mas fornecerá os dados para tal.

Modelo Conceitual



Com base nas respostas reais/hipotéticas fornecidas na etapa anterior (levantamento de requisitos), deve-se elaborar um Modelo de Entidades e Relacionamentos (M.E.R.) possuindo no mínimo:

- 10 entidades (todas com chave primária)
- 40 atributos
- Relacionamentos das cardinalidades 1xN e MxN

3) Modelo Lógico

Com base no M.E.R. construído na etapa anterior (modelo conceitual), deve-se gerar um **modelo lógico** no qual conste a criação de todas as tabelas e relacionamentos lá apresentadas.

A partir do modelo desenvolvido:

- 3.1) Apresente exemplos que utilizem conceitos de seleção e de projeção;
- 3.2) Apresente exemplos que utilizem conceitos de união e diferença;
- 3.3) Apresente exemplos que utilizem conceitos de produto cartesiano e interseção;
- 3.4) Apresente exemplos que utilizem conceitos de divisão e agregação.

4) Modelo Físico

Com base no **modelo lógico** construído na etapa anterior deve-se gerar um **modelo físico** e um conjunto de instruções SQL compatíveis com o SGBD escolhido (Oracle ou SQL Server) para criação dos objetos no banco. Não esquecer de fazer menção às FKs e PKs.

Esquema: Deve-se criar um **schema Oracle** ou uma **base SQL Server** e executar as instruções SQL geradas nesta etapa.

Implementação: As instruções geradas devem gerar uma estrutura de dados física, na qual os dados devem ser armazenados (especificação na próxima etapa) para continuar o desenvolvimento do projeto.

- 4.1) Devem ser criadas tabelas de acordo com as entidades apresentadas no MER;
- 4.2) Deve ser criada ao menos uma tabela que possua um trigger;
- 4.3) Deve ser criada ao menos uma view para visualização de dados;

- 4.4) Não esqueça de criar as constraints PK e FK;
- 4.5) Deve ser utilizado ao menos 4 atributos que possuam a constraint NOT NULL;
- 4.6) Deve ser utilizado ao menos 2 atributos que possuam a constraint UNIQUE;
- 4.7) Deve ser utilizado ao menos 2 atributos que possuam a constraint DEFAULT.

5) Carga de Dados

Rotinas: Com base na estrutura de dados gerada na etapa anterior (modelo físico), devem ser criadas rotinas/scripts em PL/SQL ou TSQL para alimentar tabelas.

Não deve existir preocupação em produzir nomes “reais”. Por exemplo, uma pessoa poderia se chamar P001 e residir no logradouro L001.

Volumetria: Devem existir tabelas atendendo às condições:

- Ao menos uma tabela deve possuir mais que um milhão de linhas (repositórios gerados por transações costumam atingir grandes volumes)
- No mínimo três (3) tabelas de cadastro. Exemplos: Cliente, Aluno, Livro, Filme etc.

Não se esqueça de utilizar os conceitos de concorrência, transações (BEGIN TRANSACTION, ROLLBACK, COMMIT) e bloqueios em suas cargas de dados. Explique por que utilizar a transação neste ponto da carga.

Nas cargas devem ser utilizadas operações de INCLUSÃO, ALTERAÇÃO E EXCLUSÃO de dados.

6) Consultas de dados

Devem ser construídas consultas que “respondam” às questões levantadas na etapa “1) Definição de mini mundo”.

Por exemplo, no tema “Conserto de Roupas”, poderia existir esta questão:

- Onde moram e como poderia contatar os cinco (5) clientes cujos consertos levaram mais tempo nos últimos doze meses?

E a resposta seria uma relação com cinco linhas contendo nome, endereço e telefone dos clientes teoricamente mais insatisfeitos.

Devem ser construídas ao menos uma (1) consulta para cada questão levantada na fase “1) Definição de mini mundo”.

Além das consultas que respondam às questões apresentadas anteriormente, a partir do modelo desenvolvido:

6.1) Crie consultas que utilizem conceitos de agregação de dados utilizando as cláusulas **Count** e **Sum**;

6.2) Crie consultas que utilizem conceitos de agregação de dados utilizando as cláusulas **Min** e **Max**;

6.3) Crie consultas que utilizem conceitos de agregação de dados utilizando a cláusula **AVG**;

6.4) Crie consultas que utilizem conceitos de agregação utilizando as cláusulas **GROUP BY** e **HAVING**.

Como anunciado inicialmente este projeto será apresentado em duas partes, uma em cada disciplina do bloco.

A primeira parte, que deve ser apresentada como **projeto desta disciplina**, é o desenvolvimento deste projeto **até a fase 3 (Modelo Lógico)**.

O apresentável do projeto deve ser um arquivo PDF com a especificação dos itens das fases 1, 2 e 3 do projeto. Assim que terminar, salve o seu arquivo PDF e poste no Moodle. Utilize o nome do aluno para nomear o arquivo, identificando também a disciplina no seguinte formato:

“nomedoaluno_nomedadisciplina_pd.PDF”.