



Projeto da Disciplina MATA60 (Bancos de Dados)

Elaboração, implantação, governança e uso de banco de dados em estudo de caso

(individual ou grupo de até 4 pessoas)

1. Introdução

Esta atividade tem a intencionalidade de provocar discentes a explorar os aspectos de projeto, implantação, otimização, suporte e análise de bancos de dados. No percurso da avaliação se observarão a criatividade, descoberta, rigor, eficiência e capacidade de planejamento na resolução de um problema que mimetiza a complexidade de cenários reais na indústria, mercado e pesquisa.

O ponto de partida para a condução do projeto é o refinamento de um desafio fictício atribuído pelo professor. O texto inicial oferecido pelo professor busca emular uma complexidade suficiente para uma exploração completa dos itens avaliados em cada entrega. A partir disso, cada discente, individualmente, deve refletir sobre os requisitos, elaborar novos requisitos e especificar o minimundo de banco de dados. Os produtos entregues devem refletir o planejamento, implantação, operação, otimização e administração do banco de dados, observando os requisitos. Todas as demais entregas serão avaliadas a partir de sua coerência com esta delimitação do minimundo, sua modelagem inicial e os documentos de governança. Adicionalmente, é importante que estes documentos também materializem as demandas do projeto lógico/conceitual.

Este projeto é organizado em torno de duas avaliações. Cada avaliação se refere a um marco de entrega. A primeira entrega abriga os desenvolvimentos, descobertas e definições da concepção e implantação do projeto de sistema de informação. Neste processo, cada discente deve incorporar as melhores práticas de modelagem, cumprindo com as determinações dos artefatos de Governança de Dados. Estes artefatos descrevem, principalmente, a Metodologia de Administração de Dados (MAD).

Seguindo a definição do DATASUS¹, o MAD contém a regulamentação e descrição que estabelecem padrões e procedimentos a serem adotados quando necessária a realização de tarefas relativas à área de Administração de Dados, considerando o contexto da Governança de Dados. Parte substantiva desta metodologia é suportada pela ISO/IEC 11179-5:2015², principalmente no seu capítulo orientando estratégias de nomenclatura³. Desta forma, todas as abstrações, relações, restrições e implementações desta entrega devem estar associadas aos

¹ <https://datasus.saude.gov.br/metodologia-de-administracao-de-dados-mad/>

² https://ibamanet.ibama.gov.br/phocadownload/2023/TI/MAD_modelo_de_administracao_de_dados.pdf

³ <https://www.iso.org/standard/62453.html>



requisitos básicos e estendidos do caso atribuído ao discente. O sucesso nesta entrega está diretamente relacionado à capacidade de produzir uma modelagem adequada, uma estratégia de tradução entre modelos (Entidade-Relacionamento para o Relacional) amparados nos fundamentos teóricos oferecidos nos livros-texto da disciplina e, também, no uso efetivo dos mecanismos de otimização do modelo físico.

As entregas do Marco 2 são compostas pelos avanços na elaboração de rotinas para otimização dos processos de administração do banco de dados pelo sistema de informação e pelos seus atores (Administradores de Bancos de Dados, Analistas de Dados, Gestores, etc.). Estão compreendidos nesta entrega um conjunto de *queries* de manipulação de dados, as visões materializadas, *stored procedures*, transações e otimização de consultas. Um aprofundamento teórico-prático dos conceitos de bancos de dados e exploração de seus subsistemas para cumprimento dos requisitos previamente estabelecidos é a chave para uma entrega de excelência.

No momento da segunda entrega, será possível revisitar a modelagem feita nas entregas do Marco 1. Este exercício de melhoria é de suma importância para garantir que os ajustes provocados pelo feedback do Marco 1 permitam o cumprimento pleno de requisitos e acesso a escalas mais altas de avaliação. Em seguida, este documento descreverá os marcos, suas entregas, avaliações e materiais de apoio.

2. Marco 1: Elaboração, modelagem e implantação ótima do banco de dados

(MIBD)

2.1. Descritivo das entregas

A partir do estudo de caso atribuído pelo professor e dos documentos de governança oferecidos, cada grupo deve propor uma extensão do estudo de caso, modelar o banco de dados, implantar o banco de dados usando o modelo relacional no SGBD PostgreSQL, popular as relações no SGBD e otimizar o modelo físico relacional.

Propor uma extensão ao estudo de caso: Cada discente deverá analisar criticamente o estudo de caso atribuído e propor uma extensão de sua descrição e requisitos que devem ser cumpridos. Adicionalmente, de maneira justificada, o discente deve especificar um minimundo que inclua as tabelas, relacionamentos e restrições suficientes para cumprir com todos os requisitos (originais e estendidos).

Modelar o banco de dados: Cada discente deve, a partir da descrição do projeto, seu minimundo e requisitos, elaborar uma modelagem conceitual e lógica usando o Modelo de Entidade-Relacionamento (MER). A modelagem deve seguir a notação de Peter Chen, descrita no principal livro-texto da disciplina.

Implantar o banco de dados: Utilizando o PostgreSQL, discentes devem escrever scripts SQL-DDL (*data definition language*) para materializar a modelagem MER num banco de dados físico.



Popular o banco de dados: Para garantir uma avaliação adequada das restrições e do impacto de eventuais otimizações, cada discente deve popular seu banco com um volume de dados sintéticos.

Otimizar o modelo físico relacional: Cada discente deverá elaborar planos de indexação, justificados de acordo com os requisitos e a natureza das tabelas propostas. Para cumprir com a avaliação de desempenho, o discente deve criar 10 (dez) consultas de complexidade intermediária e outras 10 (dez) avançadas.

Relatório de entrega:

O relatório da Entrega 1 deve seguir o modelo de Capítulo de Livros da SBC (Sociedade Brasileira de Computação). Seu conteúdo deve estar organizado em, pelo menos cinco seções, como descrito no Quadro 1.

Quadro 1. Sugestão de organização da Entrega 1

Seção	Descrição
Descritivo do Projeto	Descrição final do projeto, original e estendido, contendo os requisitos e minimundo.
Modelagem	Apresentação e comentários da modelagem conceitual e lógica. Os artefatos de modelagem devem ser: <ul style="list-style-type: none">- Uma ou mais imagens produzidas a partir do BrModelo (ou outra ferramenta que adote plenamente o Modelo de Peter Chen).- Texto que explique o MER através da definição das entidades, suas relações, atributos, tipos e chaves.
Implantação	Apresentação do script DDL usado na implantação do banco de dados com foco na estratégia para “tradução” do MER para o Relacional.
População e consultas	Apresentação do código que popula o banco de dados e criação das 30 consultas (10 consultas intermediárias e 20 consultas avançadas). Deve-se inserir cinco mil registros em pelo menos duas das relações criadas. Todas as consultas devem estar relacionadas a um requisito. Sugere-se criar uma tabela descrevendo a consulta e justificando como ela cumpre ou se relaciona com um determinado requisito.
Plano de indexação	Apresentação dos planos de indexação e do resultado da avaliação de desempenho para decidir qual melhor otimiza as queries elaboradas.
Anexos	Os anexos devem incluir os códigos e tutorial necessários para a reprodução de seu projeto.



INSTITUTO DE COMPUTAÇÃO

www.computacao.ufba.br

MATA6o - Banco de Dados

Prof. Robespierre Pita

robespierre.pita@ufba.br



Regras:

- Discentes podem usar apenas SQL ANSI para operar no banco de dados, sendo proibido o uso de conexão remota ao SGBD através de outras linguagens de propósito geral (Python, R, Java, Julia, C, etc).
- Os anexos devem estar organizados para permitir que a correção aconteça sem erros. Desta forma, discentes podem optar por incluir links externos ou documentos auxiliares. É de responsabilidade do aluno que o acesso a qualquer material externo ao relatório deve estar disponível publicamente sem a necessidade de autenticação.
- A avaliação de desempenho deve considerar o tempo de execução médio de, pelo menos, 20 execuções das queries e o desvio padrão. O resultado deve avaliar o speedup entre o baseline (sem qualquer indexação) e os planos de indexação propostos.
- Consultas intermediárias envolvem no mínimo 3 tabelas E devem utilizar pelo menos 2 das seguintes funções: JOIN, GROUP BY, WINDOW e COUNT.
- Consultas avançadas envolvem no mínimo 3 tabelas E devem utilizar pelo menos 3 das seguintes funções: SUB-CONSULTAS, JOIN, GROUP BY, WINDOW e COUNT.
- Todas as consultas devem estar associadas a pelo menos um requisito, isso deve estar explícito no relatório.



2.2. Sugestão de cronograma para cumprimento do prazo

A seguir, se oferece uma sugestão de planejamento para cumprimento pleno do prazo. Recomenda-se que esta proposta seja adaptada à realidade de cada grupo e que utilizem os espaços de orientação e monitoria para garantir boas entregas e avaliações.

MARCO 1											
Metas	Descrição	Atividades	Item(ns) no Barema	Semana de início	Semana final	S1	S2	S3	S4	S5	S6
M1. Definição e extensão do Estudo de Caso	Nesta meta estão concentradas a apropriação dos estudos de caso atribuídos para a equipe, a definição do estudo de caso que será conduzido e a proposta de extensão dos requisitos e minimundo.	M1. A1. Definição do estudo de caso	P1.Q1 P1.Q2	1	3						
		M1. A2. Detalhamento do texto que descreve o sistema (a partir da seção "Introdução")	P1.Q1 P1.Q2								
		M1. A3. Avaliação crítica e extensão dos requisitos oferecidos inicialmente.	P1.Q1 P1.Q2 P1.Q3								
		M1. A4. Detalhamento e Delimitação do mini-mundo (a partir do texto inicial).	P1.Q1 P1.Q2 P1.Q3								
M2. Modelagem e Tradução entre modelos	Nesta meta estão concentradas as atividades de modelagem, conforme o mini-mundo e os requisitos produzidos na meta anterior, e os esforços preparatórios para a implantação.	M2.A1. Definição do MER	P2.Q1 P2.Q2	2	4						
		M2.A2. Definição do modelo lógico	P2.Q1 P2.Q2								
		M2.A3. Ajuste do modelo lógico com base na estratégia de tradução entre o MER e o MR.	P2.Q1 P2.Q2								
M3. Implantação e Desempenho	Esta meta concentra as atividades de implantação e otimização do modelo físico do banco de dados produzido a partir da modelagem.	M3.A1. Criação de DDL (inclusive de constraints)	P2.Q1 P2.Q2	3	5						
		M3.A2. População do Banco de Dados	P2.Q1								
		M3.A3. Produção das queries	P3.Q1								
		M3.A4. Produção do plano de indexação	P3.Q1 P3.Q2								
		M3.A5. Condução de experimentos e produção do relatório	P3.Q2								
M4. Relatório e entrega	Aqui se concentra a confecção e entrega do relatório.	M4.A1. Escrita do relatório	NA	4	6						
		M4.A2. Entrega	NA								

➤ SUGESTÃO DE CRONOGRAMA



2.3. Barema

Quadro 2. Barema de avaliação da Entrega 1.

Item	Peso	Avaliação											
P1. Estudo de caso	3	[Q1] = Entrega apresenta requisitos estendidos adequados e bem justificados.				[Q2] = Q1 + Elaboração do minimundo é adequada, bem justificada e cumpre com os requisitos dos documentos de governança.				[Q3] = Q2 + Cada aspecto do minimundo está devidamente associado a um requisito. E Relatório apresentado segue o modelo de Capítulo de Livro da SBC.			
P2. Modelagem e implantação	5	[Q1] = Modelos conceitual e lógico são adequados e cumprem com o minimundo. E DDL é funcional e cumpre adequadamente o minimundo (materializando a estratégia de tradução entre modelos). E Base de dados é populada (pelo menos duas tabelas devem ter, pelo menos, cinco mil registros). E Cumprir com as demandas de padrão para nomenclaturas do MAD					[Q2] = Q1 + Modelagem é adequadamente comentada no relatório, principalmente relacionando as estratégias com o minimundo. E O relatório contém uma exploração dos metadados conforme apresentado em uma das aulas expositivas. E Relatório apresentado segue o modelo de Capítulo de Livro da SBC.						
P3. Desempenho	2	[Q1] = Elaborou 10 consultas intermediárias e 20 consultas avançadas, funcionais, com retorno não-vazio e adequadamente associadas a pelo menos um requisito.					[Q2] = Q1 + Avaliação de desempenho coletou tempo de execução de cada consulta pelo menos 10 vezes, e utilizou a média e o desvio padrão dos tempos de execução para calcular o speedup. E Relatório apresentado segue o modelo de Capítulo de Livro da SBC.						
nota		0	1	2	3	4	5	6	7	8	9	10	



3. Entrega 2: Rotinas avançadas (OCRA)

3.1. Descritivo das entregas

Uma vez implantado o banco de dados para o sistema de informação, deve-se aprimorar sua implantação com o suporte a tarefas analíticas, privacidade e preservação de dados. Desta forma, cada discente deverá elaborar um relatório evidenciando como endereçar estas demandas usando *materialized views*, *stored procedures* e transações.

Para possibilitar esta avaliação, cada discente deve elaborar um conjunto de cláusulas SQL para servir de back-end nos seguintes artefatos:

Quadro 3. Artefatos do SI, comandos necessários e sugestão para aplicação de rotinas

Artefato do sistema de informação	Descrição dos comandos	Sugestão de Rotina
Tela 1: cadastro com validação	1 comando SELECT 2 comandos INSERT 2 comandos UPDATE	Materialized View Stored Procedure
Tela 2: cadastro ou validação	2 comandos SELECT 1 comando INSERT 2 comandos UPDATE 1 comandos DELETE	Materialized View Stored Procedure Transação
Dashboard 1: 4 gráficos analíticos estratégicos	Cada gráfico corresponde ao resultado de uma consulta avançada (diferentes da Entrega 1)	Materialized View Stored Procedure
Dashboard 1: 6 gráficos analíticos operacionais	2 gráficos devem ser geradas por uma consulta avançada e 4 por consultas intermediárias (diferentes da Entrega 1)	Materialized View Stored Procedure

3.2. Relatório de entrega

O relatório da Entrega 2 deve seguir o modelo de Capítulo de Livros da SBC (Sociedade Brasileira de Computação). Seu conteúdo deve estar organizado em, pelo menos cinco seções, como descrito no Quadro 4.



Quadro 4. Sugestão de organização da Entrega 2

Seção	Descrição
Descritivo do Projeto	Comentários gerais sobre eventuais mudanças necessárias na Entrega 1, com base nas observações do professor ou amadurecimento do processo.
Artefatos	Apresentação dos artefatos, contextualizados no SI, e dos comandos necessários para cumprir com as demandas de cada artefato no quadro 3
Rotinas	Scripts SQL e demonstração de uso no SGBD PostgreSQL
Políticas	Apresentação e demonstração das políticas de backup e efeito das configurações de privacidade e segurança.
Anexos	Os anexos devem incluir os códigos e tutorial necessários para a reprodução de seu projeto. Discentes devem considerar o envio de um vídeo de demonstração para as demonstrações.

Regras:

- Discentes podem usar apenas SQL ANSI para operar no banco de dados, sendo proibido o uso de conexão remota ao SGBD através de outras linguagens de propósito geral (Python, R, Java, Julia, C, etc).
- Os anexos devem estar organizados para permitir que a correção aconteça sem erros. Desta forma, discentes podem optar por incluir links externos ou documentos auxiliares. Nesta etapa, o envio de um vídeo de demonstração da entrega é encorajado. É de responsabilidade do aluno que o acesso a qualquer material externo ao relatório deve estar disponível publicamente sem a necessidade de autenticação.
- Consultas intermediárias envolvem no mínimo 3 tabelas E devem utilizar pelo menos 2 das seguintes funções: JOIN, GROUP BY, WINDOW e COUNT.
- Consultas avançadas envolvem no mínimo 3 tabelas E devem utilizar pelo menos 3 das seguintes funções: SUB-CONSULTAS, JOIN, GROUP BY, WINDOW e COUNT.
- Todas as consultas e rotinas devem estar associadas a pelo menos um requisito, isso deve estar explícito no relatório.



3.4. Barema

Quadro 5. Barema de avaliação da Entrega 2.

Item	Peso	Avaliação											
P1. Consultas e comandos	4	[Q1] = Consultas e comandos atendem adequadamente as demandas do Quadro 3.				[Q2] = Q1 + Consultas e comandos, bem como as telas e dashboards estão devidamente previstos em requisitos do sistema de informação.							
P2. Modelagem e implantação	3	[Q1] = Elaborou 2 Stored Procedures, 2 materialized Views e 2 Transações adequadamente.						[Q2] = Q1 + Rotinas elaboradas estão devidamente previstos em requisitos do sistema de informação.					
P3. Políticas	3	[Q1] = Cumpriu adequadamente com as políticas de backup e privacidade.											
nota		0	1	2	3	4	5	6	7	8	9	10	

4. Arguição oral para o Marco 1 e Marco 2.

A arguição oral de cada Marco será conduzida através de quatro questões cobrindo aspectos diferentes do projeto entregue por cada equipe. As questões serão direcionadas a diferentes componentes de forma alternada.

4.1. Barema para avaliação do Marco 1 (PO_{MIDB})

Item	Peso	Avaliação									
P1. Questão oral 1: Mini-mundo e modelagem	4	Q1 = Resposta foi clara e objetiva. E A fala ocupou entre 1 e 5 minutos.				[Q2] = Q1 + Resposta foi correta. E Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.					
P2. Questão oral 1: Implantação	3	Q1 = Resposta foi clara e objetiva. E				[Q2] = Q1 + Resposta foi correta. E Demonstrou domínio sobre a entrega					



INSTITUTO DE COMPUTAÇÃO

www.computacao.ufba.br

MATA60 - Banco de Dados

Prof. Robespierre Pita

robespierre.pita@ufba.br




		A fala ocupou entre 1 e 5 minutos.	e os conceitos fundamentais da disciplina.									
P3. Questão oral 3: Consultas		Q1 = Resposta foi clara e objetiva. <div>E</div> A fala ocupou entre 1 e 5 minutos.	[Q2] = Q1 + Resposta foi correta. <div>E</div> Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.									
P4. Questão oral 4: Desempenho		Q1 = Resposta foi clara e objetiva. <div>E</div> A fala ocupou entre 1 e 5 minutos.	[Q2] = Q1 + Resposta foi correta. <div>E</div> Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.									
nota		0	1	2	3	4	5	6	7	8	9	10

$$PO_{MIDB}: (P1 + P2 + P3 + P4) / 40$$

4.2. Barema para avaliação do Marco 2 (PO_{OCRA})

Item	Peso	Avaliação	
P1. Questão oral 1: Mini-mundo e modelagem	4	Q1 = Resposta foi clara e objetiva. E A fala ocupou entre 1 e 5 minutos.	[Q2] = Q1 + Resposta foi correta. E Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.
P2. Questão oral 1: Implantação	3	Q1 = Resposta foi clara e objetiva. E A fala ocupou entre 1 e 5 minutos.	[Q2] = Q1 + Resposta foi correta. E Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.
P3. Questão oral 3: Consultas		Q1 = Resposta foi clara e objetiva. E A fala ocupou entre 1 e 5 minutos.	[Q2] = Q1 + Resposta foi correta. E Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.
P4. Questão oral 4: Desempenho		Q1 = Resposta foi clara e objetiva. E	[Q2] = Q1 + Resposta foi correta. E



		 A fala ocupou entre 1 e 5 minutos.	Demonstrou domínio sobre a entrega e os conceitos fundamentais da disciplina.									
nota		0	1	2	3	4	5	6	7	8	9	10

$$PO_{OCRA}: (P1 + P2 + P3 + P4) / 40$$

5. Cálculo de nota final

A nota para a avaliação será composta pela seguinte equação:

$$MÉDIA FINAL = ((MIBD * PO_{MIBD}) + (OCRA * PO_{OCRA})) / 20$$