
CURSO: Ciência de Dados e Inteligência Artificial – 2º semestre de 2023
DISCIPLINA: **Banco de Dados**
PROFESSOR(ES): Júlio César Chaves
CARGA HORÁRIA: 60h
PRÉ-REQUISITO: Introdução à Computação
HORÁRIO E SALA DE ATENDIMENTO: Todas as sextas-feiras de 13h às 15h.
SALA: 502

COMPLEMENTAÇÃO DE CARGA HORÁRIA: 2 aulas de 1h40min

PLANO DE ENSINO

1. Ementa

Modelo Entidade-Relacionamento. Tipos de entidades, conjuntos de entidades e atributos-chaves. Tipos de relacionamentos, papéis e restrições estruturais. Diagrama Entidade-Relacionamento. Modelo de Dados Relacional. Atributos-chaves de uma relação. Esquemas de Bases de Dados Relacionais e Restrições de integridade. Linguagens Formais de Consulta: Álgebra Relacional. A linguagem SQL. Bancos de Dados NoSQL. Manipulação de bases de dados publicamente disponíveis.

2. Objetivos da disciplina

Bancos de dados estão presentes no cotidiano, seja fazendo uma compra no supermercado, jogando on-line, ou visualizando um conteúdo na internet. Esta disciplina tem como objetivo geral capacitar o aluno para compreender, criar e manipular as organizações lógicas (e até físicas) em torno de bancos de dados.

3. Procedimentos de ensino (metodologia)

A exposição dos conteúdos ocorrerá em módulos, seguindo a estrutura didática do livro do Prof Jukic, da Loyola University Chicago.

Apesar da existência de diversas tecnologias, o foco maior será na capacidade do aluno pensar sobre como criar, manipular, gerir, e administrar dados em um SGBD. Algo que o aluno poderá tirar proveito em qualquer tecnologia.

Plataforma de Laboratório:

<https://pesquisador.fgv.br>

Além de servidor de contingência na nuvem: RDPAZUCPR0003.acad.fgv.br:53287 (IP 191.235.80.65)

Ferramentas principais da disciplina:

O servidor de laboratório da EMAP.

ERDPLUS (<https://erdplus.com/>) – Modelagem em diagramas ER (Conceitual) e Relacional.

Visual Paradigm (online.visual-paradigm.com) – Modelagem de diversos tipos de diagramas.

MySQL – Banco de dados para modelo relacional.

Gremlin (Gremlify.com) – Plataforma para manipular bases de dados de grafos, como exemplo de dados não-estruturados.

4. Conteúdo programático detalhado

Datas	Tópico	Atividades
8/8/23	Apresentação, acordos, eclass, introdução histórica	Fórum e Questionário
10/8/23	Introdução – Terminologia Inicial e Etapas	Lab 1 - Acessando o seu Lab
15/8/23	Requisitos de BD e modelagem ER (conceitual)	Dúvidas da Introdução, Questionário
17/8/23	Requisitos de BD e modelagem ER (conceitual)	Lab 2 – Acessando o seu servidor
22/8/23	Requisitos de BD e modelagem ER (conceitual)	Questionário
24/8/23	Requisitos de BD (apêndice I)	Lab 3 – Instalando e autenticando-se no seu SGBD
29/8/23	Modelo lógico	Formação dos grupos, definição de tema por grupo.
31/8/23	Modelo lógico	Lab 4 – Criando a minha primeira tabela.
5/9/23	Estabelecimento dos trios da A1, escopo do trabalho: portais de preferência.	Entrega da Fase 1 do projeto em grupo, lista de no mín. 10 requisitos de BD.
12/9/23	Anomalias e normalização	Questionário
14/9/23	Consulta médica	
15/9/23	Compensação, verificar sala e horário Anomalias e normalização	Lab 5 – Usando o workbench.
19/9/23	SQL, álgebra relacional	Entrega da Fase 2 do projeto em grupo: mín. 20 req BD + ER
21/9/23	Apresentação externa: Gerente de Banco de Dados da FGV, Otavio Freitas Teixeira.	

26/9/23	Espaço de revisão e dúvidas sobre a entrega da Fase 3 do projeto.	Entrega de atividade: Resumo da apresentação externa (mín. três parágrafos)
28/9/23	Apresentação A1 - Modelo ER e relacional, implantando banco de dados vazio	Entrega + apresentação da Fase 3 do projeto em grupo: req BD + ER + esquema relacional + excel com dados de exemplo
3/10/23 (A1)	Apresentação A1 para quem não conseguiu	
5/10/23 (A1)		
10/10/23	SQL	
17/10/23	SQL analítico	Questionário
19/10/23	Implantação e uso de BD	Lab 6 – Implantando as Lojas ZAGI
20/10/23	(Complemento de horário) Implantação e uso de BD	Dúvidas do lab 6
24/10/23	Implantação e uso de BD	Dúvidas do lab 6
26/10/23	Implantação e uso de BD (apêndice)	Entrega da Fase 4 do projeto em grupo: Implantação do banco de dados no SGBD, carga inicial (mín 5 linhas p/tabela) e no mínimo 3 instruções SQL úteis e explicadas. Lab 7 - Implantando as Lojas ZAGI - continuação
31/10/23	NoSQL, Gremlin	Lab 8 – Consultas em bases de dados em grafo
7/11/23	ACID, bloqueio impeditivo e deadlock. Indexação, controle de acesso, plano de execução	
9/11/23	SGBD em cluster, carga de dados, particionamento de tabelas	Lab 9 – Simulando bloqueios impeditivos
14/11/23	Instalação, backup, monitoração	
16/11/23	Dados georreferenciados. Carga de shapefile, funções auxiliares de mapa em SGBDR.	Lab 10 – Consultas geoespaciais
17/11/23	(Complemento de horário) Otimização de consultas em grandes volumes de dados	Lab 11 – Consultas em bases de dados colunares.
21/11/23	Espaço de revisão e dúvidas sobre a entrega da Fase final do projeto.	

23/11/23	Apresentação de trabalho A2 – CRUD	Entrega + apresentação da Fase final do projeto em grupo
28/11/23 (A2)	Apresentação A2 para quem não conseguiu	
30/11/23 (A2)		
14-20/12/23 (AS)	Modelar, implantar e simular operações sistêmicas	

Adicionar:

Lab – instalando o próprio SGBD

Lab – criar lab para um CRUD com python + mysql

5. Procedimentos de avaliação

Frequência

Espera-se que todos os alunos compareçam a todas as aulas. Os alunos serão, em todos os momentos, responsáveis por todos os materiais e informações discutidos e fornecidos em sala de aula. Isso será aplicável independentemente de o aluno assistir ou não a uma aula específica. Faltar às aulas afetará a sua avaliação no quesito de participação nas aulas.

Participação

Espera-se que todos os membros da classe participem de todas as classes. A participação é uma das principais chaves para a aprendizagem. Cada aluno será julgado pela qualidade e quantidade de participação nas discussões em classe. Esperamos que os alunos sejam interativos na aula. Contribuições que agreguem novos insights ou avancem nas discussões, inclusive esclarecendo dúvidas, serão premiadas.

Atividades

Os alunos trabalharão em várias tarefas de casa ao longo do curso. Todas as tarefas devem ser feitas individualmente, salvo indicação em contrário.

Projeto

Você também trabalhará em um projeto em equipe. Os projetos serão baseados em minimundos fictícios ou reais. Você precisará modelar os dados usando as metodologias apropriadas e apresentar suas descobertas para a turma no final da A1 e da A2. A participação é crítica em qualquer projeto de grupo - não deixe apenas seu parceiro resolver os problemas para você. Embora você possa dividir certos aspectos do projeto, espera-se que você seja capaz de apresentar 100% do trabalho sozinho. Iremos ajustar a pontuação de um aluno com base no resultado da avaliação dos colegas ou na decorrência da não participação no trabalho.

Entregas em atraso

Não haverá reposição de quaisquer tarefas ou laboratórios em sala de aula devido à ausência do aluno. Haverá uma multa na nota de 10% ao dia para as tarefas recebidas após a data de vencimento. As entregas que ultrapassarem 8h a partir da hora final da entrega contarão como 1 dia de atraso.

Trabalho em equipe

Trabalho em equipe e em grupos de estudo são incentivados e podem ser necessários para algumas tarefas. No entanto, copiar, plagiar, etc. não será tolerado.

Composição da nota

A avaliação será feita através da participação na entrega dos exercícios, questionários, fóruns e trabalhos, da seguinte forma:

(4 pontos) Entrega em fases e apresentação de trabalho em grupo nas avaliações finais A1 e A2.

Trabalho A1 - Engenharia reversa de portal da internet.

Trabalho A2 – CRUD do mesmo projeto da A1.

(4 pontos) Atividades individuais e em grupo fora da avaliação final (Trabalho)

Fóruns, Questionários, Entregas individuais, Entregas em grupo fora da final.

(2 pontos) Participação individual: (-) negativa ou (+) positiva

(-) dormiu na aula,

(-) está alheio à aula,

(-) chegou com atraso ou saiu antes do fim da aula,

(-) não seguiu as instruções do lab

(+) respondeu corretamente quando perguntado pelo professor,

(+) ajudou colegas no lab quando requisitado pelo professor

AS = Prova (10 pontos).

6. Bibliografia Obrigatória

Jukic, N., Vrbsky, S., Nestorov, S., & Sharma, A. (2021). Database Systems: An Introduction to Databases and Data Warehouses. 005.74 D232da 2nd.

KORTH, H. F.; SILBERSCHATZ, A.; SUDARSHAN, S.. Sistema de Banco de Dados. 5a ed., Campus, 2006. 005.74 S582d 6. Ed (tem on-line na minhabiblioteca).

HEUSER, Carlos Alberto. Projeto de Banco de Dados. 6a Ed., Bookman, 2008. 005.74H595p 6. ed (tem on-line na minhabiblioteca).

GARCIA-MOLINA Hector, ULLMAN, Jeffrey D., WIDOM, Jennifer. Database systems, 2014. 005.74 G216d 2.ed.

7. Bibliografia Complementar

Korth, Henri F. Data base systems concepts. Sixth edition. Mc Graw-Hill.

DATE, C. J. Introdução a Sistemas de Banco de Dados. Elsevier Editora, 2004.

Marz, Nathan; W, James. Big data: Principles and best practices of scalable realtime data systems. 1st. Manning Publication

RAMAKRISHNAN, R.; GEHRKE, J. Sistemas de Gerenciamentos de Bancos de Dados. 3a ed., McGraw Hill Brasil, 2008.

Database Systems: the complete book. 2a ed., Prentice Hall, 2008.

(<https://github.com/hexanome/Denmark/blob/master/Advanced%20Databases/Prentice.Hall.Database.Systems.The.Complete.Book.2nd.Edition.Jun.2008.pdf>)

Semantic web for the working ontologist: modeling in RDF, RDFS and OWL. Dean Allemang, James Hendler. 025.04 A424s.

8. Minicurrículo do(s) Professor(s)

Concluiu a graduação em tecnologia em informática na UNIVERCIDADE/RJ (2006). Possui mestrado (2011) e doutorado (2018) pela COPPE/UFRJ sob o programa de engenharia civil, tendo como foco de pesquisa a mobilidade urbana a partir de dados georreferenciados, tais como dados providos por operadoras de telefonia móvel, do IBGE e do setor de transportes. As pesquisas buscam enriquecer, a partir da análise abrangente dos dados, as metodologias tradicionais de inferência de mobilidade. Atualmente coordena a recém-criada área de governança de infraestrutura para pesquisas na Fundação Getúlio Vargas.

Formação acadêmica/titulação

2012 - 2018

Doutorado em Engenharia Civil.

Universidade Federal do Rio de Janeiro, UFRJ, Brasil.

Título: ESTIMATIVA DE MATRIZES DE ORIGEM-DESTINO A PARTIR DE REGISTROS DE CHAMADAS DE TELEFONIA CELULAR, Ano de obtenção: 2018.

Orientador: Alexandre Gonçalves Evsukoff.

Coorientador: Moacyr Alvim Horta Barbosa da Silva.

Palavras-chave: mobilidade; urbana; cdr; celular; matriz; origem.

Grande área: Ciências Exatas e da Terra

Setores de atividade: Transporte terrestre; Transporte aéreo; Obras de infra-estrutura.

2009 - 2011

Mestrado em Engenharia Civil (Conceito CAPES 6).

Universidade Federal do Rio de Janeiro, UFRJ, Brasil.

Título: AVALIAÇÃO DE SEGMENTAÇÕES URBANAS ATRAVÉS DE INTELIGÊNCIA COMPUTACIONAL, Ano de Obtenção: 2011.

Orientador: Alexandre Gonçalves Evsukoff.

Grande área: Ciências Exatas e da Terra.

Atuação Profissional

FGV - Fundação Getúlio Vargas

2019 - Atual

Vínculo: Colaborador, Enquadramento Funcional: Governança de Infraestrutura para Pesquisas, Carga horária: 44, Regime: Dedicção exclusiva.

2010 - 2018

Vínculo: Colaborador, Enquadramento Funcional: DBA líder, Carga horária: 44, Regime: Dedicção exclusiva.

2007 - 2010

Vínculo: Outro, Enquadramento Funcional: Administrador de Banco de Dados. Carga horária: 44, Regime: Dedicção exclusiva.

VARIGLOG

2005 – 2007

Vínculo: Outro, Enquadramento Funcional: Administrador de Banco de Dados. Carga horária: 44, Regime: Dedicção exclusiva.

CLARO

2003 – 2005

Vínculo: Outro, Enquadramento Funcional: Administrador de Banco de Dados. Carga horária: 44, Regime: Dedicção exclusiva.

DANA/ATH - Albarus/GKN

1994 – 2002

Vínculo: Outro, Enquadramento Funcional: Aprendiz SENAI em Mecânica Geral, Atendente de Service Desk, Aprendiz de Administrador de Banco de Dados. Carga horária: 44, Regime: Dedicção exclusiva.

Áreas de atuação

1. Grande área: Ciências Exatas e da Terra / Área: Ciência da Computação / Subárea: Sistemas de Computação.

Idiomas

Espanhol

Compreende Razoavelmente, Fala Pouco, Lê Razoavelmente, Escreve Pouco.

Inglês

Compreende Bem, Fala Razoavelmente, Lê Bem, Escreve Razoavelmente.

Português

Compreende Bem, Fala Bem, Lê Bem, Escreve Bem.

Produções

Chaves, J. C., A.H.B. da Silva, M., & Evsukoff, A. G. (2021). Gravity Model Parameter Variation during a Long-Term Study using Mobile Phone Data in the Rio de Janeiro Metropolitan Area. Case Studies on Transport Policy. <https://doi.org/10.1016/J.CSTP.2021.12.017>

BARBOZA, MATHEUS H. C. ; ALENCAR, RICARDO DE S. ; CHAVES, JULIO C. ; SILVA, MOACYR A. H. B. ; ORRICO, ROMULO D. ; EVSUKOFF, ALEXANDRE G. . Identifying Human Mobility Patterns in the Rio de Janeiro Metropolitan Area using Call Detail Records. TRANSPORTATION RESEARCH RECORD, v. 1, p. 036119812097765, 2020.

Citações:1

LENORMAND, MAXIME ; SAMANIEGO, HORACIO ; Chaves, Júlio César ; DA FONSECA VIEIRA, VINÍCIUS ; DA SILVA, MOACYR ALVIM HORTA BARBOSA ; EVSUKOFF, ALEXANDRE GONÇALVES . Entropy as a Measure of Attractiveness and Socioeconomic Complexity in Rio de Janeiro Metropolitan Area. Entropy, v. 22, p. 368, 2020.

Citações:3

Apresentações de Trabalho

1.

Chaves, Júlio César; ALMEIDA, M. S. G. ; BACHA, M. N. ; AYALA, K. ; FERREIRA, I. ; VINHAS, F. . I ENCONTRO DE GESTÃO DE DADOS DE PESQUISA EM HUMANIDADES E CIÊNCIAS SOCIAIS. 2019. (Apresentação de Trabalho/Outra).

2.

CHAVES, J. C.. Arquiteturas de TI como suporte de processamento para pesquisas. 2019. (Apresentação de Trabalho/Seminário).

Eventos

1.
OPEN SCIENCE/OPEN DATA. 2018. (Outra).
2.
Workshop on Urban Computing and Society. 2018. (Outra).

9. Link para o Currículo Lattes

<http://lattes.cnpq.br/6558177209348272>