

TRABALHO PRÁTICO

FUNDAMENTOS DE BANCO DE DADOS 2019/1

Sergio Mergen

Objetivo Geral

- O trabalho é dividido em 5 etapas
 - **(1 ponto)** Criação de um banco de dados desnormalizado
 - **(2 pontos)** Normalização do esquema desnormalizado
 - **(1 ponto)** Modelagem conceitual do esquema normalizado
 - **(4 pontos)** Carga dos dados para o esquema normalizado
 - **(2 pontos)** Consultas sobre o esquema normalizado
- O SGBD a ser utilizado é o MySQL

Etapa 1

- **Criação de um banco de dados desnormalizado**
- **Peso: 1 ponto**
- **Passos**
 - Escolher uma fonte de dados com **dados abertos** disponibilizados na Web
 - Ex. arquivos csv, planilhas excel, ...
 - Gerar o script SQL para a criação de um banco de dados com uma tabela que contenha todas as colunas existentes.
 - Executar o script.

Dados Abertos

- Dados que devem ter livre acesso por qualquer pessoa
- Normalmente publicados na Web
- O uso é irrestrito
 - Em alguns casos se exige apenas o reconhecimento de quem disponibilizou o dado
 - Ainda se exige que os dados gerados a partir de dados abertos também sejam abertos
- Algumas fontes de dados abertos
 - Governo (esfera municipal, estadual, federal)
 - Empresas privadas
 - ...

Dados Abertos

The screenshot shows the homepage of dados.gov.br, the Brazilian Open Data Portal. The browser address bar shows the URL 'dados.gov.br'. The page features a green header with the portal's name and logo, and a navigation menu with links to 'Participe', 'Serviços', 'Legislação', and 'Canais'. A search bar is located below the header, displaying 'Pesquisa ...' and 'PESQUISAR'. Below the search bar, it states 'em 1122 conjuntos de dados com 8993 recursos (o que é isto?)'. The main content area is divided into two columns. The left column contains a section titled 'Dados em destaque' (Data in focus) with a link to 'Reclamações do Consumidor.gov.br' and a link to 'Lista de Eleitores Filiados aos Partidos Políticos'. The right column contains a section titled 'Publicações mais recentes' (Most recent publications) with a table of recent data releases.

Conjunto de dados	Autor	Quando
Áreas Embargadas	Edson Eyji Sano - CSR ...	17 Abr
REGIÃO DE INFORMAÇÃO DE VOO - FIR	Divisão de Operações	03 Abr
AEROVIAS	Divisão de Operações	03 Abr
AERÓDROMOS DO BRASIL	Divisão de	03 Abr

The Windows taskbar at the bottom shows the system clock as 17:18 on 04/05/2016, and the battery level is at 56%.

Dados Abertos

The screenshot shows a web browser window with the URL `dados.gov.br/dataset/obras-do-pac-programa-de-aceleracao-do-crescimento`. The page features a sidebar on the left with social media links (Google+, Twitter, Facebook) and a 'Licença' section titled 'Licença Aberta para Bases de Dados (ODbL) do Open Data Commons'. The main content area displays a list of datasets under the heading 'Dados e recursos'. Each dataset entry includes a file icon, a title, and an 'Explorar' button. The datasets listed are:

- DICIONÁRIO DE DADOS – Empreendimentos do PAC
- Obras do PAC em ODS referentes ao 2º Balanço de ...
- Obras do PAC em CSV referentes ao 2º balanço de ...
- Obras do PAC em XLSX referentes ao 2º balanço de ...
- Obras do PAC em XML referentes ao 2º balanço de ...
- Empreendimentos regionais do PAC referentes ao 2º ...
- Obras do PAC em ODS referentes ao 1º Balanço de ...
- Obras do PAC em XLSX referentes ao 1º Balanço de ...
- Empreendimentos regionais do PAC referentes ao 1º ...

At the top of the main content area, there is a paragraph of text: 'Os dados mais atuais são os do 2º balanço de 2013-2016, referente ao ano de 2013, e fornecem informações acerca de todos os empreendimentos presentes na carteira de projetos do programa.'

Dados Abertos

- **IMPORTANTE:** As fontes de dados a serem usadas **devem** ser nacionais
- Existem diversas fontes brasileiras de dados abertos
- Boa procura. 😊

Etapa 2

- **Normalização do esquema desnormalizado**
- **Peso: 2 pontos**
- **Passos**
 - Usar as formas normais para decompor a tabela desnormalizada em várias tabelas
 - Gerar o script SQL para a criação das tabelas normalizadas dentro do mesmo banco de dados criado na etapa 1.
 - Executar o script.

Etapa 3

- **Modelagem conceitual do esquema normalizado**
- **Peso: 1 ponto**
- **Passos**
 - Gerar uma representação gráfica do modelo relacional.
 - Realizar engenharia reversa do esquema de banco de dados normalizado para que seja gerado o modelo conceitual correspondente (modelo ER)
 - Gerar uma representação gráfica do modelo ER.
- Obs. O modelo dito 'ER' gerado pelo editor do MySQL é na verdade o modelo relacional, e não o conceitual

Etapa 4

- **Carga dos dados para o esquema normalizado**
- **Peso máximo: 4 pontos**
- A nota pode variar da seguinte forma:
 - 4 pontos (100%) – Carga 1
 - 2,8 pontos (70%) – Carga 2
 - 1,6 pontos (40%) – Carga 3

Etapa 4

- **Carga dos dados para o esquema normalizado**
- 4 pontos (100%) – Carga 1
 - Usar um extrator que carregue os dados da fonte para a tabela desnormalizada
 - Usar um script SQL para que os dados da tabela desnormalizada sejam copiados para as respectivas tabelas normalizadas

Etapa 4

- **Carga dos dados para o esquema normalizado**
- 2,8 pontos (70%) – Carga 2
 - Carregar manualmente os dados da fonte para a tabela desnormalizada
 - Usar um script SQL para que os dados da tabela desnormalizada sejam copiados para as respectivas tabelas normalizadas

Etapa 4

- **Carga dos dados para o esquema normalizado**
- 1,6 pontos (40%) – Carga 3
 - Carregar manualmente os dados da fonte diretamente para as tabelas normalizadas
 - Sem passar pela tabela desnormalizada

Etapa 5

- **Consultas sobre o esquema normalizado**
- Passos
 - Criar consultas SQL sobre as tabelas normalizadas
 - Executar as consultas para obter as respostas.
- **Peso: 2 pontos**

Etapa 5

- **Consultas sobre o esquema normalizado**
- Criar pelo menos quatro consultas que envolvam alguns desses construtores:
 - agregação
 - operador de conjunto
 - junção externa
 - junções entre três ou mais tabelas
 - subconsulta
- Não é necessário usar todos os construtores
 - Pode-se usar o mesmo para todas as consultas
- Só serão aceitas consultas cujo resultado seja significativo

Formação dos grupos

- Os alunos podem formar grupos de até três integrantes.
- Cada grupo poderá escolher uma fonte a partir de dados abertos na Web
 - Desde que a fonte seja nacional
- Terá direito a usar uma fonte o grupo que primeiro fizer a indicação ao professor
- Ou seja, o grupo deve verificar com o professor se a fonte escolhida já está sendo usada por outro grupo

O que entregar

- Fonte escolhida
- Criação do banco
 - Script de criação da tabela desnormalizada
 - Script de criação das tabelas normalizadas
- Alimentação do banco
 - Scripts sql de alimentação, códigos de extração, ...
- Esquemas gráficos
 - Modelo ER
 - Modelo relacional
- Consultas elaboradas
 - Scripts SQL
- Apresentação final
 - slides onde a solução adotada é relatada.

O que entregar

- A entrega será pelo moodle.
- Apenas um membro do grupo precisa entregar
- Apenas um arquivo deve ser entregue
 - Todos os documentos devem ser comprimidos em um só.

Apresentação

- O que apresentar:
 - A descrição da solução adotada, em slides.
 - O esquema do banco de dados criado
 - Esquemas gráficos para ER e relacional
 - A execução dos scripts SQL
 - Script SQL de criação do banco
 - Script SQL de alimentação da tabela desnormalizada
 - Script SQL de alimentação das tabelas normalizadas
 - Scripts SQLs de consultas
- Salvo decisão posterior, a execução ocorrerá a partir do computador do professor.

Apresentação

- Tempo de apresentação: até 20 minutos por grupo.
- Escolha do apresentador
 - A apresentação deverá ser feita por apenas uma pessoa.
 - A escolha será feita por sorteio, no dia da apresentação.

Descontos da nota

- O grupo perde dois pontos da nota final para cada aluno sorteado que
 - não quiser realizar a apresentação ou
 - Não demonstrar conhecimento à respeito do trabalho
- Alunos que se ausentarem da apresentação sem justificativa receberão zero na nota final.
- Anulação da nota do grupo: O grupo receberá nota zero caso o prazo de entrega não seja respeitado.

Descontos da nota

- A comprovação de realização das etapas se dá mediante
 - a entrega da documentação necessária
 - E a apresentação para o grupo
- Ou seja, a apresentação deve fornecer indícios que comprovem que as etapas foram realizadas
- Não entregar ou não apresentar esses indícios
 - Levam a anulação daquela parcela da nota

Datas

- Data da entrega
 - Todos os grupos enviarão o trabalho na mesma data
 - A data da entrega será acertada com o professor.
- Data da apresentação
 - As apresentações ocorrerão durante três aulas.
 - As datas das apresentações serão acertadas com o professor.
 - Assim como os grupos que farão a apresentação em cada aula.