

TRABALHO PRÁTICO

FUNDAMENTOS DE BANCO DE DADOS 2022/1

Sergio Mergen

Objetivo Geral

- O objetivo principal do trabalho é a criação de um dashboard para visualização de dados abertos.
- O trabalho é dividido em 5 etapas
 - **(1 ponto)** Definição de um tema
 - **(2 pontos)** Criação de um esquema relacional normalizado
 - **(1 ponto)** Representação gráfica do modelo
 - **(3 pontos)** Carga dos dados
 - **(3 pontos)** Visualização dos dados

Etapa 1

- **Definição de um tema**
- **Peso: 1 ponto**
- **Passos**
 - Escolher uma fonte de dados com **dados abertos** disponibilizados na Web
 - Ex. arquivos csv, planilhas excel, ...

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`. A search bar at the top contains the text "odont" and shows "Nenhum resultado" (No results). The main header features the portal's logo and navigation links: "Acesso à informação", "Participe", "Serviços", "Legislação", and "Canais". Below the header, a green banner displays the portal's name and tagline: "PORTAL BRASILEIRO DE DADOS ABERTOS".

The main content area is divided into several sections:

- Pesquisa ...**: A search bar with the text "Pesquisa ..." and a "PESQUISAR" button. Below it, it states "em 1122 conjuntos de dados com 8993 recursos (o que é isto?)".
- Dados em destaque**: A section highlighting featured data sets, including "Reclamações do Consumidor.gov.br" and "Lista de Eleitores Filiados aos Partidos Políticos".
- Publicações mais recentes**: A table listing recent publications.

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

The Windows taskbar at the bottom shows the system clock as 17:18 on 04/05/2016, along with various application icons and a battery level indicator 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 (PDF icon)
- Obras do PAC em ODS referentes ao 2º Balanço de ... (DATA icon)
- Obras do PAC em CSV referentes ao 2º balanço de ... (CSV icon)
- Obras do PAC em XLSX referentes ao 2º balanço de ... (XLS icon)
- Obras do PAC em XML referentes ao 2º balanço de ... (XML icon)
- Empreendimentos regionais do PAC referentes ao 2º ... (CSV icon)
- Obras do PAC em ODS referentes ao 1º Balanço de ... (DATA icon)
- Obras do PAC em XLSX referentes ao 1º Balanço de ... (XLS icon)
- Empreendimentos regionais do PAC referentes ao 1º ... (DATA icon)

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.'

The Windows taskbar at the bottom shows the search bar 'Pesquisar na Web e no Windows', several application icons, and the system clock indicating 17:18 on 04/05/2016.

Etapa 2

- **Criação de um esquema normalizado**
- **Peso: 2 pontos**
- **Passos**
 - Usar as formas normais para decompor os dados da(s) fonte(s) escolhida(s) em tabelas normalizadas
 - Gerar o script SQL para a criação das tabelas normalizadas.
 - Executar o script.

Etapa 3

- **Representação gráfica do modelo**
- **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 ER (conceitual) correspondente
 - Gerar uma representação gráfica do modelo relacional (lógico).

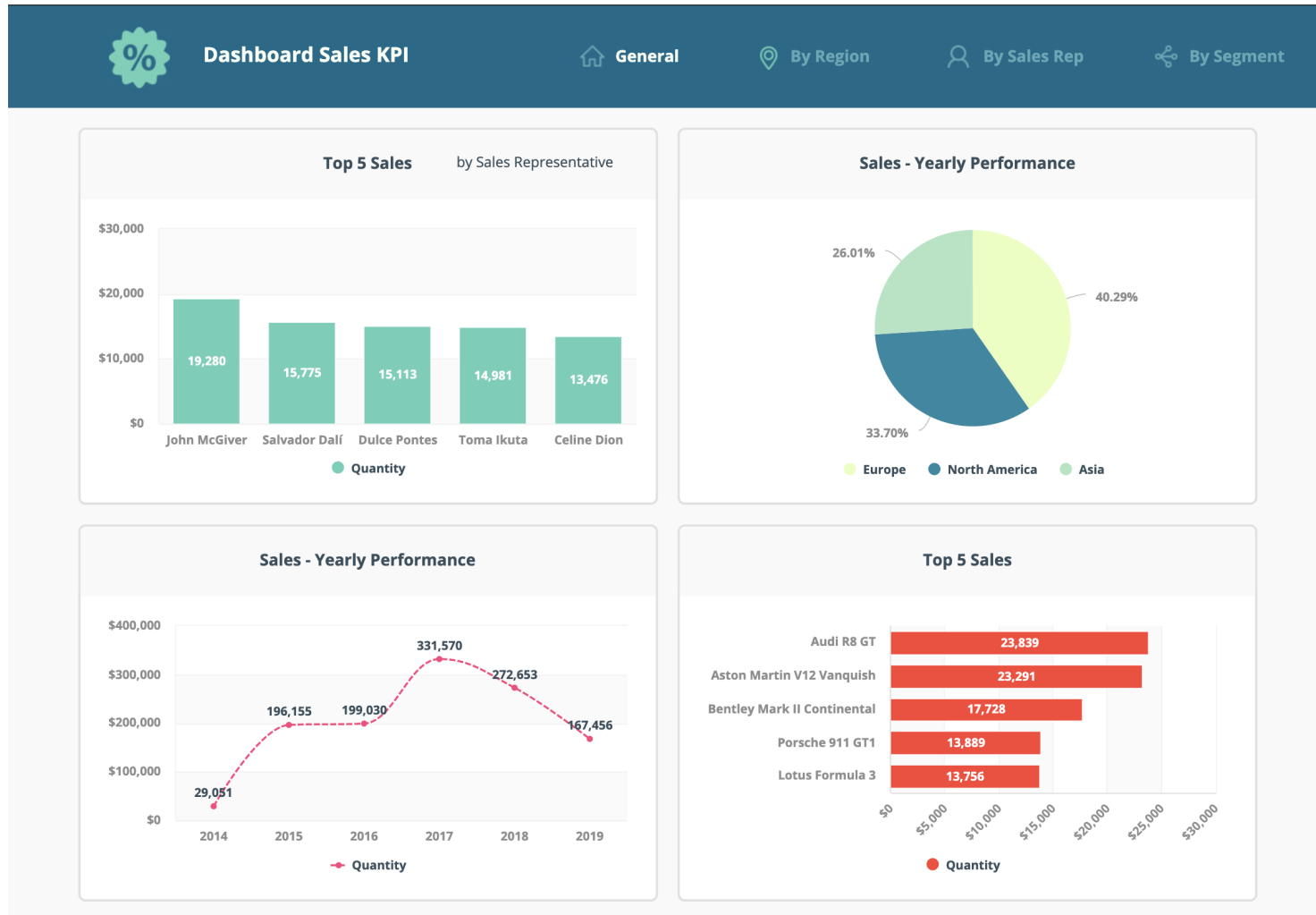
Etapa 4

- **Carga dos dados**
- **Peso: 3 pontos**
- **Passos**
 - Importar os dados da(s) fonte(s) escolhida(s) para o banco de dados relacional
- **Obs.**
 - A carga pode ser feita através de scripts de código fonte ou com o auxílio de ferramentas
 - Caso a carga seja feita de forma manual, a nota máxima equivale a 1,5 pontos

Etapa 5

- **Visualização dos dados**
- Peso: 3 pontos
- Passos
 - Escolher uma ferramenta de visualização de dados
 - Criar um dashboard a partir do banco de dados criado
 - Demonstrar o dashboard criado, interagindo com os parâmetros

Exemplo de Dashboard



Formação dos grupos

- Os alunos podem formar grupos de até três integrantes.
- Cada grupo poderá escolher uma fonte (ou mais) a partir de dados abertos na Web
- 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 (caso não seja muito grande)
- Criação do banco
 - Script de criação das tabelas normalizadas
- Alimentação do banco
 - Scripts sql de alimentação, códigos de extração, ...
- Esquemas gráficos
 - Modelo ER
 - Modelo relacional
- Dashboard criado
- 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 os modelos ER e relacional
 - A execução do processo de criação e alimentação do banco
 - A execução do processo de criação do dashboard

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

- A nota em cada etapa está associada a
 - o grau de dificuldade para a execução dos passos
 - relevância do tema escolhido
- Por exemplo
 - A escolha de uma fonte de dados com poucos registros
 - simplifica a etapa de alimentação do banco
 - Não permite que sejam feitas análises significativas sobre os dados
 - Isso será considerado na atribuição da nota

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 for 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 postarão 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.