

Data Warehouse e Mineração de Dados – Trabalho Prático Data Warehouse e OLAP

Data de Apresentação: 21/10/2016

Prof. Ismael Santana Silva (ismaelsantana@decom.cefetmg.br)

Valor 26 pontos

Trabalho em trio.

Este trabalho tem como objetivo exercitar os conceitos de Data Warehouse, ETL, OLAP.

Para sua execução, cada grupo deverá escolher pelo menos dois conjuntos de dados disponíveis no:

- Portal Brasileiro de Dados Abertos (<http://dados.gov.br>) ou;
- Portal da Transparência do Estado de Minas Gerais (<http://www.transparencia.mg.gov.br/dados-abertos>) ou;
- Dados Abertos da Assembleia de Minas (https://www.almg.gov.br/sobre/dados_abertos).

Devem ser escolhidos conjuntos de dados interessantes, por exemplo:

- Compras públicas (<http://dados.gov.br/dataset/compras-publicas-do-governo-federal>)
- Despesas Diretas do Poder Executivo Federal;
- Prestação de Contas das Campanhas Eleitorais;
- Orçamento Federal.

O grupo deve informar no questionário <http://goo.gl/forms/5VVwqbgQ6M> os conjuntos de dados escolhidos até o dia **07/10/2016**, para que não haja conflitos com o mesmo conjunto de dados. Os temas escolhidos serão publicados no endereço <https://goo.gl/MTsla8>.

Definido o conjunto de dados o trio deverá:

- Realizar uma caracterização dos dados, apresentar informações como: número de registros, número de atributos, categorias presentes nos dados;
- Fazer uma análise crítica sobre a qualidade dos dados;
- Modelar um *Data Warehouse* multidimensional;
- Realizar a Extração, Transformação e Carga (ETL);
- Criar visualizações para análise dos dados e padrões utilizando uma ferramenta OLAP;
- Criar uma apresentação explicando todo o processo executado e a análise dos dados a partir da caracterização e visualizações geradas;
 - Você pode buscar informações em notícias para embasar suas análises.
- Criar um repositório no **GitHub** e inserir todos os artefatos criados e utilizados no trabalho.
- Enviar a apresentação com o link para o repositório para o e-mail ismaelsantana@decom.cefetmg.br.

O trabalho será avaliado por uma banca externa que votará na melhor apresentação. A apresentação mais votada será premiada. **Será avaliado na apresentação:** postura, oralidade, **empenho**, clareza, duração, **conteúdo** e **aprendizado**. O grupo deve entregar os slides da apresentação via e-mail e todos os artefatos gerados por e-mail.

A aparência dos slides também será avaliada apresentação deve ser observada, serão levados em consideração aspectos como:

- Tamanho da fonte;
- Quantidade de texto;

Bom trabalho!