



Unidade Curricular: Armazenamento e Processamento Analítico de Dados

MiniProjecto II (Final)

Concepção, Desenvolvimento e Utilização de um DataMart

No âmbito do segundo trabalho teórico-prático de avaliação da unidade curricular Armazenamento e Processamento Analítico de Dados dos Mestrados de Sistemas e Tecnologias de Informação para as Organizações e Engenharia Informática e Sistemas de Informação, pretende-se que os mestrandos elaborem um projeto de criação e exploração de um DataMart.

Descrição do trabalho

Um dos elementos de avaliação desta unidade curricular consiste na elaboração de um projeto relativo à conceção e implementação de um DW/DM. O projeto tem como finalidade aprofundar os conhecimentos teóricos relativos à conceção do DW/DM, no seu modelo dimensional, e a perspetiva de integração, mostrando como se efetuará a extração dos dados e a transformação para o modelo dimensional. Importa também investigar e apresentar a utilização de sistemas analíticos que permitam dar resposta às necessidades de quem utiliza os sistemas de DW/DM.

Em resumo, pretende-se que cada grupo de mestrandos realize um trabalho de análise, planeamento, arquitetura e implementação de um sistema de Data Warehouse ou DataMart para uma dada instituição/tema.

Escolha do tema

A seleção do tema, que servirá como caso de estudo, será da responsabilidade de cada grupo de trabalho, que deverá assim indicar qual o tipo de instituição/tema que vai servir como modelo para a implementação do sistema requerido e apresentar as suas características básicas.

A escolha do tema deve ser realizada até dia **14 Junho de 2021**.

Constituição do Grupo

Os grupos serão constituídos por um máximo de 3 elementos. O número de elementos do grupo será considerado na avaliação do MiniProjecto. Sempre que se justificar, poderão ser desenvolvidos MiniProjectos individuais.

Todos os elementos do grupo devem participar na realização do MiniProjecto e estarem preparados para responder às perguntas que surgirem na sua avaliação.

Componentes do trabalho:

1. Relatório do MiniProjecto desenvolvido em forma de artigo: Ponderação de 35%.

Avaliação: Média (Resumo; Introdução; Contextualização e Caso de Estudo; Conceitos Teóricos/ Tecnologias; Conclusões e Trabalho Futuro; Referências; Nível Global do Artigo);

2. Elaboração do MiniProjeto: Ponderação de 50%.

Avaliação: Média (Sistema Concebido/Implementação: Sistema Fonte e ETL; Sistema Concebido/Implementação: DM (Modelo Dimensional); Resultados Experimentais; Nível Global de Concepção / Implementação);

3. Apresentação 20 minutos (max.): Ponderação de 15% - Componente obrigatória. Todos os membros do grupo necessitam de apresentar o trabalho.

Avaliação: Qualidade dos Diapositivos (10%); Qualidade da Exposição dos Conteúdos (15%); Modelo Relacional (10%); Modelo Dimensional (10%); ETL (15%); Consultas (15%); Interesse e Qualidade das Respostas às Questões (25%).

Componentes obrigatórias do MiniProjeto

1. Base de dados relacional onde devem constar os dados operacionais. Devem definir e construir os esquemas conceptuais. Podem escolher qualquer SGBD.
2. Desenvolvimento de um Sistema ETL, nas suas vertentes de extração, limpeza e transformação. Em resumo devem desenvolver um programa que deve aceder às fontes de dados originais e transpor o seu conteúdo para a base de dados multidimensional. Atenção: o programa deve padronizar e sumarizar os dados (ex. normalização de unidades de medida). Podem escolher qualquer linguagem de programação.
3. Desenvolvimento de uma base de dados multidimensional de acordo com a metodologia de Kimball. Podem escolher qualquer SGBD.
4. Utilização ou criação de uma ferramenta OLAP para consultas ou desenvolvimento de um programa para exploração dos dados que permita efetuar análises sobre os mesmos, devendo mostrar a importância de cada consulta e a sua efetivação. Podem escolher qualquer linguagem de programação e/ou ferramenta.

Valorizações do MiniProjeto

1. Integração de múltiplas fontes de dados, tais como outras BD's, ficheiros csv, txt, pdf, excel, etc.
2. Utilização de dois motores de base de dados diferentes (exemplo: utilização para a BD relacional MySql e para a BD Multidimensional SQL Server).
3. Desenvolvimento de um *frontend* para exploração de dados, em detrimento de uma aplicação OLAP.

Notas

1. As linguagens de programação, motores de base de dados e ferramentas que optem por utilizar ficam ao critério dos alunos.
2. Além do relatório, cada grupo deve entregar todo o código e scripts SQL.

Documentos a entregar:

1. No dia 1 ou 13 de Julho:
 - Um ficheiro com o artigo correspondente à descrição do MiniProjecto, concepção da solução, sua implementação e resultados obtidos (gráficos

e/ou consultas). Este deverá ser escrito num formato de artigo de 2 colunas. O modelo é disponibilizado no Moodle.

- Um ficheiro ZIP com o projeto, scripts e códigos SQL.
2. No dia da apresentação, um ficheiro pdf com os diapositivos.
 3. A entrega dos artefactos correspondentes à resolução do MiniProjecto II (componentes dos itens 1 e 2) é feita exclusivamente pela plataforma Moodle.

Data de entrega:

Época Normal

1 de Julho de 2021 (véspera da data do Exame).

Época Recurso

13 de Julho (véspera da data do Exame de Recurso).

Apresentação/defesa:

Época Normal

2 de Julho de 2021 (após exame) e, se necessário, no dia 06 de Julho, às 19H00.

Época Recurso

Após o exame de recurso (14 de Julho) e, se necessário, no dia 15 de Julho, às 19H00.