



Folha de Exercícios n.º PL04	
Unidade Curricular:	Análise de Dados
Ano Letivo:	2019/2020 - 1.º Semestre

1 Modelação Dimensional em Constelação de Factos

O principal objetivo da resolução da primeira parte deste exercício é analisar a base de dados *sakila*, bem como o ficheiro *calendario.xlsx*, disponibilizados durante as aulas práticas laboratoriais desta unidade curricular, escolher a informação de interesse para futura análise de dados e, consequentemente, definir um modelo dimensional no formato de constelação de factos.

É de relembrar que um modelo dimensional baseado no formato em constelação de factos é constituído por duas ou mais tabelas de factos que podem partilhar entre elas tabelas de dimensão.

Numa segunda parte, procederá ao povoamento do *data warehouse* definido e implementado, bem como à gestão dos seus processos.

É de notar que pode consultar mais informação de apoio sobre a base de dados *sakila* disponibilizada na seguinte referência: <https://dev.mysql.com/doc/sakila/en/>.

Com base no caso apresentado, pretende-se que:

1. Implemente a base de dados *sakila* no MySQL Workbench com o ficheiro *sakila-schema.sql*.
2. Povoie as tabelas da base de dados criada no passo anterior com o ficheiro *sakila-data.sql*.
3. Defina um modelo dimensional em constelação de factos a partir da base de dados *sakila* (ver o ficheiro *sakila.mwb*) – *EER Diagram*. No entanto, deverá ter em consideração os seguintes três pontos:
 - (a) Deverá definir duas tabelas de factos: *FACTS_PAYMENT* (para os pagamentos) e *FACTS_RENTAL* (para os alugueres);
 - (b) Deverá definir obrigatoriamente uma tabela de dimensão para o tempo denominada “*DIM_TIME*”. Tenha em atenção à granularidade que definirá para esta tabela de dimensão, uma vez que terá de guardar na mesma pelo menos um identificador único para a data (*id*), valor, dia, mês, ano, dia da semana, semana do ano, se é dia útil ou não (ver ficheiro *calendario.xlsx*), se é feriado ou não (ver ficheiro *calendario.xlsx*), entre outros;
 - (c) Adicione uma coluna denominada “*etiqueta_DTA*” em cada tabela do modelo dimensional de modo a guardar o *datetime* de povoamento de cada linha.
4. Converta o modelo dimensional definido para o respetivo modelo físico numa base de dados denominada *data_warehouse* (Database > Forward Engineer).
5. Guarde num ficheiro *.sql* a *script* de criação das tabelas do modelo dimensional definido.