

Administración de Bases de Datos – Traballos tutelados

TT1: Deseño, implementación e alimentación dunha base de datos.

Curso 2023/2024

Data límite para a entrega: 8 de marzo de 2024

1. Contextualización

O obxectivo xeral dos traballos tutelados en ABD é deseñar unha pequena base de datos, que será posteriormente implementada utilizando Oracle, alimentada con datos, e afinada para que se comporte de xeito eficiente.

Para iso realizaranse 2 traballos:

TT1: (1,2 puntos) Deseño, implementación e alimentación dunha base de datos.

Realizarase o deseño conceptual e lóxico da base de datos. Posteriormente a base de datos será implementada en Oracle, e alimentada con datos a través de sentencias SQL e con ferramentas externas como SQL*Loader.

TT2: (1,8 puntos) Axuste da base de datos.

Debe describirse unha carga de traballo habitual para a base de datos, indicando unha serie de consultas frecuentes que queremos optimizar.

Utilizando esa información, a base de datos debe ser afinada para que responda coa maior eficiencia posible a esa carga de traballo.

Como tamaño orientativo para a base de datos, esta debería conter ó redor de 5 táboas unha vez transformada ó modelo relacional.

2. TT1: Traballo a realizar

O traballo a realizar para esta primeira entrega consiste nos seguintes pasos:

1. Realizar o deseño conceptual da base de datos:

Deseña o diagrama E-R. Este debe incluír os tipos de entidade, atributos e tipos de relación, especificando correctamente a participación e cardinalidade dos tipos de entidade nos tipos de relación.

Recorda que a base de datos, unha vez pasada a relacional, debe conter unhas 5 táboas. Non sobre-dimensiones o modelo con demasiados tipos de entidade ou tipos de relación M:N. Ten en conta tamén que debes ter un número razoablemente alto de atributos, polo que tampouco debes abusar de tipos de entidade que só teñan un identificador e un atributo, ou relacións M:N con poucos ou ningún atributo.

2. Transformar o deseño conceptual a un deseño lóxico.

Debe transformarse o deseño conceptual do punto anterior ó modelo relacional. No esquema relacional debe indicarse claramente cada táboa creada, identificando as claves primarias e foráneas (indicando a quen referencian utilizando frechas).

3. Crear un pequeno diccionario de datos para documentar o modelo.

Para cada táboa deberás dar a definición do seu nome e describir cada un dos seus atributos, indicando unha breve descrición se o nome non é significativo, o seu tipo de datos, e calquera outra característica que consideres relevante.

Describe todas as restricións que afecten á táboa ou a calquera atributo: clave primaria, outras claves candidatas se existen, restricións de unicidade, not null, checks e claves foráneas.

Para as claves foráneas debes especificar a acción referencial que leva asociada ó borrado (`on delete ...`), e xustificar por que escolles esa acción. Dado que se vai implementar en Oracle, que non admite a cláusula `on update ...`, non indiques a acción referencial para as actualizacións.

Na parte final deste enunciado tedes un exemplo de como documentar o dicionario de datos.

4. Implementar a BD en Oracle:

Crearase un script SQL coas sentencias de creación das táboas.

Este script deberá funcionar correctamente na instancia de Oracle da Facultade.

Todas as restricións, coa posible excepción das `NOT NULL`, deber ser nomeadas.

Crea todas as restricións nas sentencias `create table`. Non utilices sentencias `alter table` adicionais (a non ser que sexa absolutamente necesario, por exemplo se hai ciclos referenciais).

5. Alimentar a base de datos:

Deben engadirse filas ás táboas da base de datos. Para iso utilizarás un script con sentencias `insert` para unha das táboas, e utilizarás a ferramenta `SQL*Loader` para todas as demais. Para explorar as posibilidades que ofrece `SQL*Loader`, debes ter variacións nos ficheiros de control e/ou datos para realizar a carga. Por exemplo, usar datos separados por comas (CSV) ou de ancho fixo, cambiar separadores, poñer entre comiñas os valores etc.

Todas as táboas deben ter polo menos 100 filas, e unha delas debe ter polo menos 10.000.

Os datos en si non son relevantes, no sentido de que nos da o mesmo que vendamos produtos que se chamen *Produto1*, *Produto2*, ... , pero si deberían manter a distribución de valores que seguiría cada atributo. Por exemplo, se o noso modelo é dunha empresa de desenvolvemento de software, e na táboa de empregados temos un campo *salario*, non sería lóxico que todos os empregados cobrasen igual, nin que todos cobrasen diferente coa mesma probabilidade. É máis común que haxa uns poucos con salarios altos (iguais ou distintos entre si) e moitos máis empregados con salarios máis baixos. Da mesma forma, se tivesemos o campo *posto de traballo* non debería haber a mesma porcentaxe de xefes de proxecto que de programadores. Isto ten unha gran relevancia para a utilización de índices.

3. Normas de entrega e data límite

O traballo realizarase en grupos de 2 persoas (que non teñen por que estar no mesmo grupo das clases de prácticas).

A documentación a entregar será:

- Un documento en formato PDF coa memoria da práctica, que incluírá a parte de modelización conceptual e lóxica da base de datos e o dicionario de datos.

Na portada debe aparecer o nome, DNI e correo electrónico dos membros do grupo.

- O script SQL coas sentencias de creación das táboas.
- O script SQL para inserción de datos.
- Os ficheiros de control e de datos utilizados para a carga usando `SQL*Loader`.

A entrega realizarase usando o campus virtual da UDC, só por parte dunha das persoas que forman o grupo.

A data límite para a entrega é o día 8 de marzo de 2024.

Apéndice: exemplo de dicionario de datos

Este é un exemplo de como podedes documentar o dicionario de datos.

O formato específico das táboas deste exemplo non é obrigatorio, só orientativo, pero si que debes utilizar un formato tabular para o dicionario.

DEPARTAMENTO			
Unidade organizativa e administrativa da empresa			
ATRIBUTOS			
<i>Nome atributo</i>	<i>Tipo de datos</i>	<i>Nulos?</i>	<i>Descrición</i>
CODDEP	NUMERIC(3)	NON	Código identificador do departamento
NOMDEP	VARCHAR(20)	NON	Nome do departamento
LOCDEP	VARCHAR(50)		Localidade onde se ubica o departamento
CLAVE PRIMARIA:			
CODDEP			
RESTRICIÓN DE UNICIDADE OU CLAVES CANDIDATAS (1 por fila):			
NOMDEP			

EMPREGADO			
Persoa contratada pola empresa que realiza un traballo a cambio dun salario			
ATRIBUTOS			
<i>Nome atributo</i>	<i>Tipo de datos</i>	<i>Nulos?</i>	<i>Descrición</i>
CODEMP	NUMERIC(4)	NON	Código identificador do empregado na empresa
NOMEMP	VARCHAR(40)	NON	Nome de pila do empregado
APELEMP	VARCHAR(40)	NON	Apelidos do empregado
SALEMP	NUMERIC(7,2)	NON	Salario bruto anual do empregado
CODDEP	NUMERIC(3)		Código do departamento ó que está adscrito o empregado
CLAVE PRIMARIA:			
CODEMP			
RESTRICIÓN DE UNICIDADE OU CLAVES CANDIDATAS (1 por fila)			
CHECKS:			
SALEMP > 14000			
CLAVES FORÁNEAS:			
CODDEP	Referencia: DEPARTAMENTO (CODDEP) Acción referencial: ON DELETE SET NULL Xustificación: Ó eliminar un departamento non se eliminan (despiden) os seus empregados, quedan (temporalmente) sen adscribir a ningún departamento		