

Complementos de Bases de Dados

Filipe Fidalgo

ffidalgo@ipcb.pt

Informações

Suponha que foi contactado no sentido de desenvolver um sistema de suporte à decisão, para uma organização que realiza análise registos de chuva, em cada freguesia. Cada registo de chuva realizado, deve conter a data em que ocorreu, o tipo de chuva (ex: fraca, forte, moderada,...), a duração e ainda a quantidade de precipitação registada.

Depois de uma conversa com o decisor da organização foram definidas um conjunto de questões que irão servir de base para a construção do referido sistema.

Assim, para a realização deste trabalho deve cada grupo:

- 1) Definir o modelo multidimensional (tendo por base as questões identificadas);
- 2) Criar o código necessário à criação da estrutura física do datawarehouse (de acordo com o definido no pdf da aula de SSD);
- 3) Popular as respetivas tabelas do modelo criado, com base nos exemplos de procedimentos t-sql fornecidos)
- 4) Responder às questões:
 - 1) em SQL;
 - 2) em Excel (recorrendo a PowerPivot).

A entrega deve realizar-se até às **23:59 do dia 22/Dez.**

Informações

As questões a que devem dar respostas (e logo conduzir a modelação analítica, assumindo que os dados se referem ao ano 2023) são:

1) Qual o número de ocorrências de precipitação, por dia de semana, por tipo de chuva?

Exemplo de resposta:

	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a	Sab	Dom
Fraca	500	100	75	50	500	1600	2000
Moderada	30	30	32	20	150	200	500
Forte	1000	950	850	900	1500	3000	2200
...			

Deve-se ter sempre presente qual a informação que se retira dos dados, por exemplo neste caso observa-se que:

- No país o maior número de ocorrências é do tipo 'forte', com maior prevalência aos fins de semana;

Informações

2) Qual a quantidade total, por tipo de pluviosidade, por mês?

Exemplo de resposta:

	Fraca	Moderada	Forte	...
Janeiro	500	600	700	...
Fevereiro	120	70	44	...
Março	220	143	123	...
...				

Deve-se ter sempre presente qual a informação que se retira dos dados, por exemplo neste caso observa-se que:

- O mês de janeiro regista a maior quantidade de chuva em todos os tipos de pluviosidade...;

Informações

3) Qual o tempo médio de tempo duração e de quantidade, por tipo de pluviosidade?

Ex de resposta:

	Duração	Quantidade
Fraca	13	5
Moderada	4	6
Forte	7	19
...		

Deve-se ter sempre presente qual a informação que se retira dos dados, por exemplo neste caso observa-se que:

- Existe uma maior duração média tipo de pluviosidade 'Fraca' e uma maior quantidade média no tipo de pluviosidade 'Forte'.

Informações

4) Quais os valores médios, por tipo de precipitação, por distrito?

Ex de resposta:

		Fraca	Moderada	Forte	...
Angra do Heroísmo	do	500	600	700	...
Aveiro		120	70	44	...
Beja		220	143	123	...
...					

Deve-se ter sempre presente qual a informação que se retira dos dados, por exemplo neste caso observa-se que:

- Para cada distrito, conhece-se o valor médio anual, por tipo de precipitação...

NOTA:

Os 'Ex de resposta' são apenas ilustrativos, não tem de ser obrigatoriamente esse o formato a apresentar na resposta (só o conteúdo!).

Informações

-----ESTRUTURA FÍSICA PARA O SSD-----

1) O SGBD que vamos usar para esta fase vai ser o SqlServer (encontra-se na mesma máquina que temos usado: bdsO.est.ipcb.pt)

2) Para aceder devem usar SQL Management Studio (cliente da Microsoft)

3) Terão de criar um utilizador, para tal poderão logar-se com o utilizador:

User: ffidalgo_dba

Pass: fidalgo

E executar o seguinte código (alterando a parte sublinhada, para cada caso particular IM / INF)

```
USE [master]
GO
CREATE LOGIN [INF_N_ALUNO_2425] WITH PASSWORD=N'cbd', DEFAULT_DATABASE=[master], CHECK_EXPIRATION=OFF,
CHECK_POLICY=OFF
GO
EXEC master..sp_addsrvrolemember @loginame = N'INF_N_ALUNO_2425', @rolename=N'dbcreator'
GO
```

Informações

-----ESTRUTURA FÍSICA PARA O SSD-----

Depois terão de criar uma base de dados (IMPORTANTE: já com o vosso utilizador):

```
create database INF_N_ALUNO_2425
```

(NOTA: Depois de criarem a base de dados, devem “mudar” da base de dados “master” para a vossa, através do comando: use INF_N_ALUNO_2425)

4) Em seguida, estarão em condições de realizar os passos necessários para a criação da estrutura física do DW (que constam do pdf da aula de SSD):

Cada grupo deverá construir 4 filegroups diferentes:

- Um para as dimensões (aconselha-se +- 4 dimensões);
- Um para os índices (chaves primárias);
- Dois para as duas partições (assumamos que o modelo analítico vai conter informações do ano de atividade 2023, e que as partições serão, por exemplo, uma para cada semestre) da tabela de factos.

Os “files” dos filegroups (deverão usar nomes que invalidem sobreposições para os “Filegroups” e para os “Files” associados, como por exemplo o nº de aluno), deverão ser criado em:

- “f:\fgf\...”; (para as tabelas de dimensão)
- “g:\fgg\...”; (para os índices)
- “h:\fgh\...”; (para a 1ª partição da tabela de factos)
- “i:\fgi\...”; (para a 2ª partição da tabela de factos)