

Trabalho Prático 2 - BD.N. RELACIONAL – DSM 3

Nomes:

Alan Rodrigues Monteiro Matias

Camile Benucci Costa

Diego Pereira da Silva

Gabriel Rigoni Martins

Guilherme Pereira Cardoso de Andrade

Lucca Casado Santiago

Exercício: Modelagem Dimensional com ETL

* Em grupo

Tarefas

1. Desenhar o Modelo Dimensional - Made by: Guilherme Pereira



2. Definição de Grãos Made by: Camile

Definição de Grãos da Tabela Fato 'Fato_Vendas'

Cada linha representa: Uma venda específica.

Nível de detalhe:

Data da venda: quando ocorreu.(id_tempo)

Cliente: quem fez a compra.(id_cliente)

Produto: o item que foi comprado.(id_producto)

Loja: onde a venda foi realizada.(id_loja)

Quantidade: quantidade de produto vendido.(qt_vendas)

Valor Total: total monetário da transação.(valor_total)(geralmente a soma de quantidade * preço).

3. Definição das Métricas e Atributos – Made by: Diego

Dim_Produto:

(id_produto)

nm_produto

nm_categoria

nm_marca

vl_preco

Dim_Cliente:

(id_cliente)

nm_cliente

nm_cliente

dt_cadastro

Dim_Tempo

(id_tempo)

dt_data

dt_mes

dt_ano

dt_trimestre

dt_dia

Dim_Loja

(id_loja)

nm_loja

ds_endereco

cd_cep

4. Desenho das Transformações de ETL – Made by: Alan

	A	B	C	D	E
1	Origem	Campo de Origem	Transformação	Destino	Campo de Destino
2	Produto	produto_id	s/ transformação	Dim_Produto	id_produto
3	Produto	nome_produto	s/ transformação	Dim_Produto	nm_produto
4	Produto	categoria	s/ transformação	Dim_Produto	nm_categoria
5	Produto	marca	s/ transformação	Dim_Produto	nm_marca
6	Produto	preco	alteração para duas casas decimais depois da virgula	Dim_Produto	vl_preco
7	Cliente	cliente_id	s/ transformação	Dim_Cliente	id_cliente
8	Cliente	nome_cliente	s/ transformação	Dim_Cliente	nm_cliente
9	Cliente	email	s/ transformação	Dim_Cliente	nm_email
10	Cliente	data_cadastro	De MM/DD/YYYY para DD/MM/AA	Dim_Cliente	dt_cadastro
11	Loja	loja_id	s/ transformação	Dim_Loja	id_loja
12	Loja	nome_loja	s/ transformação	Dim_Loja	nm_loja
13	Loja	localizacao_loja	extrair somente CEP que é somente numero e transf para String "XX.XX	Dim_Loja	cd_cep
14	Loja	localizacao_loja	s/ transformação	Dim_Loja	ds_endereço
15	Loja	localizacao_loja	pegar a regiao coforme o Estado	Dim_Loja	nm_regiao
16	Venda	data_venda	data completa	Dim_Tempo	dt_data
17	Venda	data_venda	extrair somente o periodo (manha=05:00-11:59h/tarde=12:00-17:59h/noite=18:00-04:59h)	Dim_Tempo	dt_periodo
18	Venda	data_venda	extrair somente mes	Dim_Tempo	dt_mes
19	Venda	data_venda	extrair somente ano	Dim_Tempo	dt_ano
20	Venda	data_venda	extrair somente trimetre (3 meses)	Dim_Tempo	dt_trimestre
21	Venda	data_venda	extrair somente o dia	Dim_Tempo	dt_dia
22	Venda	quantidade	s/ transformação	Fato_Vendas	qt_vendas
23	Venda	valor_total	quantidade*preco	Fato_Vendas	vl_total_vendas
24	Dim_Produto	id_produto	s/ transformação	Fato_Vendas	id_produto
25	Dim_Cliente	id_cliente	s/ transformação	Fato_Vendas	id_cliente
26	Dim_Loja	id_loja	s/ transformação	Fato_Vendas	id_loja
27	Dim_tempo	id_tempo	s/ transformação	Fato_Vendas	id_tempo

5. Escrever o DDL para o Modelo Dimensional - Made by: Lucca

```
2_script_ddl_dimensional.sql X
MODELAGEM_DW > PROJETO_Scripts_ETL_DW > 2_script_ddl_dimensional.sql > {} CREATE TABLE
1 -- Tabela Dimensão Produto
2 CREATE TABLE Dim_Produto (
3     id_produto INT PRIMARY KEY,
4     nm_produto VARCHAR(100),
5     nm_categoria VARCHAR(50),
6     nm_marca VARCHAR(50),
7     vl_preco DECIMAL(10,2)
8 ) TABLESPACE usuarios;
9
10 -- Tabela Dimensão Cliente
11 CREATE TABLE Dim_Cliente (
12     id_cliente INT PRIMARY KEY,
13     nm_cliente VARCHAR(50),
14     nm_email VARCHAR(100),
15     dt_cadastro DATE --Todas as datas serão formatadas no Select para DD-MM-YY
16 ) TABLESPACE usuarios;
17
18 -- Tabela Dimensão Loja
19 CREATE TABLE Dim_Loja (
20     id_loja INT PRIMARY KEY,
21     nm_loja VARCHAR(100),
22     ds_endereco VARCHAR(100),
23     nm_regiao VARCHAR(50),
24     cd_cep VARCHAR(11) --Qtd de caracteres junto com os "." do CEP
25 ) TABLESPACE usuarios;
26
27 -- Tabela Dimensão Tempo
28 CREATE TABLE Dim_Tempo (
29     id_tempo INT PRIMARY KEY,
30     dt_periodo VARCHAR(10),
31     dt_data DATE,
32     dt_mes INT,
33     dt_ano INT,
34     dt_trimestre INT,
35     dt_dia INT
36 ) TABLESPACE usuarios;
37
38 -- Tabela Fato Vendas
39 CREATE TABLE Fato_Vendas (
40     id_fato INT PRIMARY KEY, --Criação do ID pra Fato
41     id_produto INT,
42     id_cliente INT,
43     id_loja INT,
44     id_tempo INT,
45     qt_vendas INT,
46     vl_total_vendas DECIMAL(10,2),
47     FOREIGN KEY (id_produto) REFERENCES Dim_Produto(id_produto),
48     FOREIGN KEY (id_cliente) REFERENCES Dim_Cliente(id_cliente),
49     FOREIGN KEY (id_loja) REFERENCES Dim_Loja(id_loja),
50     FOREIGN KEY (id_tempo) REFERENCES Dim_Tempo(id_tempo)
51 ) TABLESPACE usuarios;
52
```

6. Criação dos Comandos ETL - Made by: Rigoni

```
3_PACKAGE.pkg_etl_dimensional_BODY.sql X
MODELAGEM_DW > PROJETO_Scripts_ETL_DW > 3_PACKAGE.pkg_etl_dimensional_BODY.sql > {} pkg_etl_dimensional
1 CREATE OR REPLACE PACKAGE BODY pkg_etl_dimensional AS
2
3 -- Procedure para carregar a Dimensão Produto
4 PROCEDURE etl_dim_produto IS
5 BEGIN
6     INSERT INTO Dim_Produto (id_produto, nm_produto, nm_categoria, nm_marca, vl_preco)
7     SELECT DISTINCT
8         p.produto_id,
9         p.nome_produto,
10        p.categoria,
11        p.marca,
12        ROUND(p.preco, 2) -- Arredondar para duas casas decimais
13     FROM Produto p;
14 END etl_dim_produto;
15
16 -- Procedure para carregar a Dimensão Cliente
17 PROCEDURE etl_dim_cliente IS
18 BEGIN
19     INSERT INTO Dim_Cliente (id_cliente, nm_cliente, nm_email, dt_cadastro)
20     SELECT DISTINCT
21         c.cliente_id,
22         c.nome_cliente,
23         c.email,
24         TO_CHAR(c.data_cadastro, 'DD-MM-YY')
25     FROM Cliente c;
26 END etl_dim_cliente;
27
28 -- Procedure para carregar a Dimensão Tempo
29 PROCEDURE etl_dim_tempo IS
30 BEGIN
31     INSERT INTO Dim_Tempo (id_tempo, dt_periodo, dt_data, dt_mes, dt_ano, dt_trimestre, dt_dia)
32     SELECT DISTINCT
33         v.venda_id,
34         CASE
35             WHEN TO_NUMBER(TO_CHAR(v.data_venda, 'HH24')) BETWEEN 5 AND 11 THEN 'Manhã'
36             WHEN TO_NUMBER(TO_CHAR(v.data_venda, 'HH24')) BETWEEN 12 AND 17 THEN 'Tarde'
37             ELSE 'Noite'
38         END AS dt_periodo,
39         TO_CHAR(v.data_venda, 'DD-MM-YY'),
40         EXTRACT(MONTH FROM v.data_venda),
41         EXTRACT(YEAR FROM v.data_venda),
42         CASE
43             WHEN EXTRACT(MONTH FROM v.data_venda) IN (1, 2, 3) THEN 1
44             WHEN EXTRACT(MONTH FROM v.data_venda) IN (4, 5, 6) THEN 2
45             WHEN EXTRACT(MONTH FROM v.data_venda) IN (7, 8, 9) THEN 3
46             ELSE 4
47         END AS dt_trimestre,
48         EXTRACT(DAY FROM v.data_venda)
49     FROM Venda v;
50 END etl_dim_tempo;
51
52 -- Procedure para carregar a Dimensão Loja
53 PROCEDURE etl_dim_loja IS
54 BEGIN
55     INSERT INTO Dim_Loja (id_loja, nm_loja, ds_endereco, nm_regiao, cd_cep)
56     SELECT DISTINCT
57         l.loja_id,
58         l.nome_loja,
59         SUBSTR(l.localizacao_loja, 1, INSTR(l.localizacao_loja, ',') - 1) AS ds_endereco,
60         SUBSTR(l.localizacao_loja, INSTR(l.localizacao_loja, ',') - 1, 3) + 2,
61         INSTR(l.localizacao_loja, ',') - 1, 4) - INSTR(l.localizacao_loja, ',') - 1, 3) - 2) AS nm_regiao,
62         -- Extrair apenas números e formatar o CEP
63         CASE
64             WHEN LENGTH(REGEXP_REPLACE(SUBSTR(l.localizacao_loja, INSTR(l.localizacao_loja, ',') - 1, 4) + 5, '[^0-9]', '')) = 8 THEN
65                 SUBSTR(REGEXP_REPLACE(SUBSTR(l.localizacao_loja, INSTR(l.localizacao_loja, ',') - 1, 4) + 5, '[^0-9]', ''), 1, 2) || '.' ||
66                 SUBSTR(REGEXP_REPLACE(SUBSTR(l.localizacao_loja, INSTR(l.localizacao_loja, ',') - 1, 4) + 5, '[^0-9]', ''), 3, 3) || '-' ||
67                 SUBSTR(REGEXP_REPLACE(SUBSTR(l.localizacao_loja, INSTR(l.localizacao_loja, ',') - 1, 4) + 5, '[^0-9]', ''), 6, 3) -- Formato XX.XXX-XXX
68             ELSE
69                 NULL -- ou um valor padrão
70         END AS cd_cep
71     FROM Loja l;
72 END etl_dim_loja;
73
74 -- Procedure para carregar a Tabela Fato Vendas
75 PROCEDURE etl_fato_vendas IS
76 BEGIN
77     INSERT INTO Fato_Vendas (id_fato, id_produto, id_cliente, id_loja, id_tempo, qt_vendas, vl_total_vendas)
78     SELECT DISTINCT
79         v.venda_id,
80         v.produto_id,
81         v.cliente_id,
82         vl.loja_id,
83         v.venda_id,
84         v.quantidade,
85         v.valor_total
86     FROM Venda v
87     JOIN Vendas_Loja vl ON v.venda_id = vl.venda_id -- Assumindo que há uma relação entre Venda e Vendas_Loja
88     JOIN Produto p ON v.produto_id = p.produto_id;
89 END etl_fato_vendas;
90 pkg_etl_dimensional;
91 /
```