Considere o seguinte cenário envolvendo a Big Z, uma atacadista de produtos automotivos.

A Big Z deseja criar o banco de dados analítico (Data Warehouse) para analisar as quantidades de seus pedidos. Os dois conjuntos de dados disponíveis, Fonte 1 e Fonte 2, são descritos abaixo.

Fonte 1 - O banco de dados de pedidos da Big Z, conforme Figura 1.

Fonte 2 - Tabela do departamento de recursos humanos, conforme Figura 2.

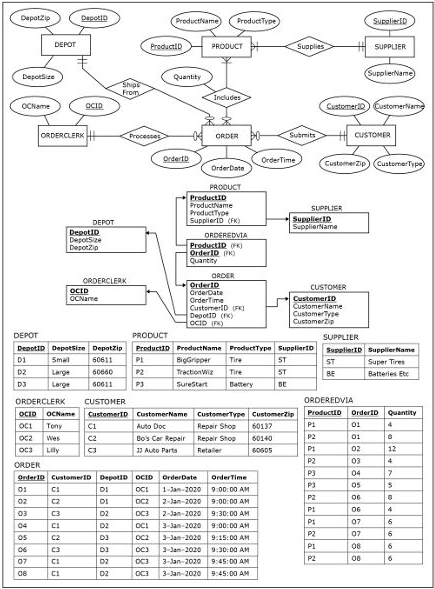


Figura 1 – Fonte 1 - Banco de dados de pedidos da Big Z.

Tabela

Descrição gerada automaticamente

Figura 2 - Fonte 2 - tabelas de Recursos Humanos – RH

O data warehouse deve permitir uma análise das quantidades de pedido, por:

• data, incluindo:

• data completa

• dia da semana

• dia do mês

• mês

• trimestre

• ano

• hora

• produto, incluindo:

• ID de produto

• nome de produto

• tipo de produto

• nome do fornecedor do produto

• cliente, incluindo:

• ID do cliente

• nome do cliente

• tipo de cliente

• zip do cliente (cep)

• depósito (depot), incluindo:

• id do depósito

• tamanho do depósito

• zip do depósito

• atendente de pedidos (orderclerk), incluindo:

• id do atendente

• nome do atendente

• função do atendente (title)

• nível educacional do atendente

• ano de contratação do atendente

Faça:

1. Com base nas fontes e nos requisitos listados acima, crie um modelo dimensional que será usado para o data warehouse modelado dimensionalmente para a Big Z.
2. Usando os dados das Figura 1 e Figura 2 mostre as tabelas populadas no modelo dimensional recém-criado no item 1). Considere as seguintes suposições:
   1. O processo de ETL escolhe a versão mais longa do nome em casos em que exista mais de uma versão de nome da mesma pessoa.
   2. o processo de ETL reconhece que um mesmo pedido foi incorretamente gravado 2 vezes com 2 identificadores de pedidos distintos (exemplos dos pedidos O7 e O8) no banco de dados de pedidos.