

Solução do desafio Luiza Labs - Descrição Geral

A solução apresentada possui um workflow (denominado *Luiza_workflow.py*) desenvolvido em *Luigi* composto por uma tarefa baseada em Python e outra em R 3.4.

A primeira tarefa (*kyConcumer.py*) é responsável pelo manuseio das *streams*. Já a segunda tarefa (*Rtask.r*) é responsável por manusear os pedidos e as *streams* para, então, gerar o arquivo final.

O arquivo final, na verdade, é um conjunto de arquivos relativos a um modelo multidimensional (Figura 1) que permite a flexível análise por pivotagem (Excel e Power Pivot, por exemplo). Caso necessário, a montagem de um único arquivo é trivial.

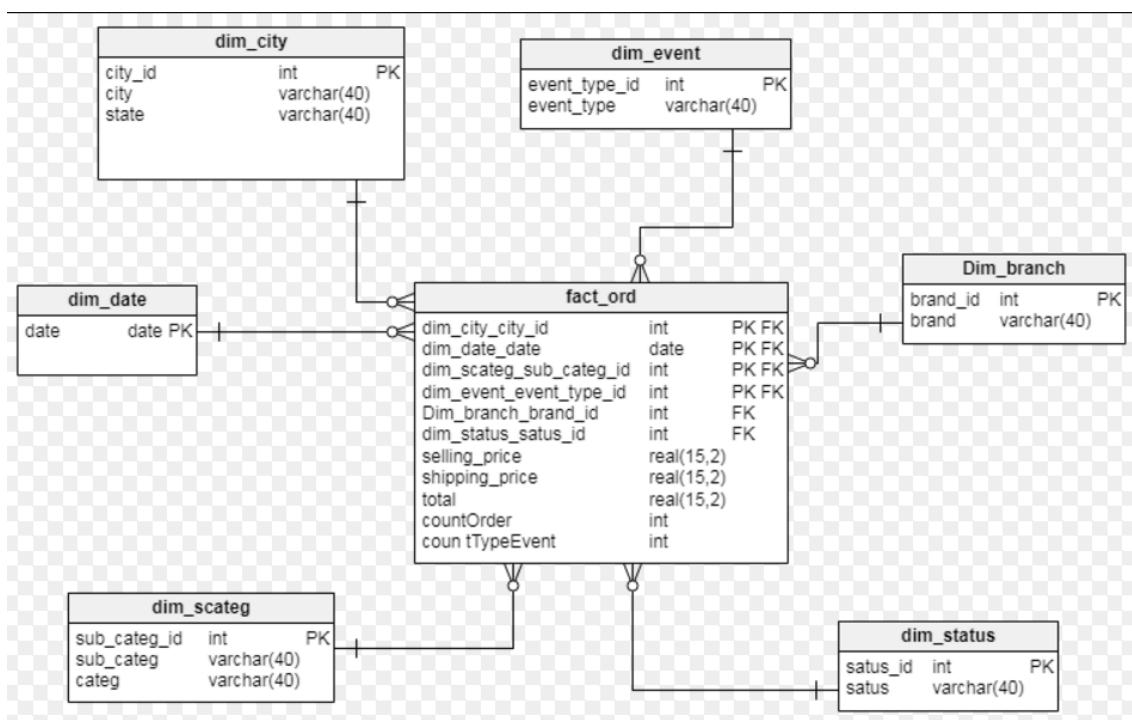


Figura 1 – Modelo dimensional da solução

Cumprе ressaltar que a solução original era totalmente baseada em R 3.4 em uma arquitetura *pipeline*. No entanto, problemas na montagem do *wrapper* Java, R, Maven requerida pelo *Kinesis for R*, fizeram-me optar pela opção Python amplamente documentada.

Att.

J. M. Borovina Josko