



Escola de Engenharia
Universidade do Minho

DEPARTAMENTO DE ENGENHARIA INFORMÁTICA
Mestrado Integrado em Engenharia Informática
Laboratórios de Informática III

Gestão de Vendas de uma cadeia de Distribuição com 3 filiais

GEREVENDAS

Parte JAVA

Grupo 84



Célia Fi-
gueiredo
a67637



Gil
Gonçalves
a67738



Humberto
Vaz
a73236



Ricardo
Lopes
a72062

Braga, 11 de Junho de 2016

1. Introdução

No âmbito da unidade curricular de Laboratórios de Informática III do 2º ano da licenciatura de Engenharia Informática, foi proposto o desenvolvimento de um projeto em linguagem Java, que tem por objectivo ajudar à consolidação dos conteúdos teóricos e práticos e enriquecer os conhecimentos adquiridos na UC de Programação Orientada aos Objectos.

Este projeto considera-se um grande desafio para nós pelo facto de passarmos a realizar programação em grande escala, uma vez que se trata de grandes volumes de dados e por isso uma maior complexidade. Como primeira etapa, o mais importante é definir as estruturas utilizadas no trabalho. Assim sendo iremos explicar ao longo do relatório detalhadamente cada uma e os aspetos mais importantes que definem a base do nosso projeto.

Por fim, explicaremos todas as nossas decisões quanto as resolução das diversas consultas (“queries”), os resultados obtidos para cada uma destas e apresentaremos também uma detalhada explicação das mesmas. Nisto englobaremos algumas estatísticas e alguns testes de “performance” do nosso programa em relação aos tipos de estruturas de dados oferecidas pela linguagem. Estes testes serão baseados na alteração das nossas estruturas iniciais e serão comparados através dos tempos de execução da leitura e de algumas queries.

Conteúdo

1	Introdução	1
2	Arquitetura da Aplicação	3
3	Módulos de dados	5
3.1	Catálogo de Clientes	5
3.2	Catálogo de Produtos	5
3.3	Faturação Global	5
3.4	Gestão da Filial	5
4	Interface do utilizador	6
5	Resultados e comentários sobre os testes de performance	7
6	Conclusão	8

2. Arquitetura da Aplicação

A arquitetura da aplicação a desenvolver é definida por quatro módulos principais: Catálogo de clientes, Catálogo de produtos, Faturação Global e Vendas por Filial, cujas fontes de dados são três ficheiros de texto detalhados abaixo.

No ficheiro **Produtos.txt** cada linha representa o código de um produto vendável no hipermercado, sendo cada código formado por duas letras maiúsculas e 4 dígitos (que representam um inteiro entre 1000 e 1999), como no exemplo:

```
AB9012
XY1185
BC9190
```

O ficheiro de produtos contém cerca de 200.000 códigos de produto.

No ficheiro **Clientes.txt** cada linha representa o código de um cliente identificado no hipermercado, sendo cada código de cliente formado por uma letra maiúscula e 4 dígitos que representam um inteiro entre 1000 e 5000, segue um exemplo:

```
F2916
W1219
F2915
```

O ficheiro de clientes contém cerca de 20.000 códigos de cliente.

O ficheiro **Vendas_1M.txt**, no qual cada linha representa o registo de uma venda efectuada numa qualquer das 3 filiais da Cadeia de Distribuição. Cada linha (a que chamaremos compra ou venda, o que apenas depende do ponto de vista) será formada por um código de produto, um preço unitário decimal (entre 0.0 e 999.99), o número inteiro de unidades compradas (entre 1 e 200), a letra **N** ou **P** conforme tenha sido uma compra **Normal** ou uma compra em **Promoção**, o código do cliente, o mês da compra (1 ... 12) e a filial (de 1 a 3) onde a venda foi realizada, como se pode verificar nos exemplos seguintes:

```
KR1583 77.72 128 P L4891 2 1
QQ1041 536.53 194 P X4054 12 3
OP1244 481.43 67 P Q3869 9 1
JP1982 343.2 168 N T1805 10 2
IZ1636 923.72 193 P T2220 4 2
```

O ficheiro de vendas inicial, **Vendas_1M.txt**, conterà 1.000.000 (1 milhão) de registos de vendas realizadas nas 3 filiais da cadeia de distribuição. Existirão também os ficheiros **Vendas_3M.txt** e **Vendas_5M.txt** utilizados para as questões de performance da aplicação.

A aplicação possui uma arquitectura tal como apresentado na figura seguinte, em que se identificam as fontes de dados, a sua leitura e os módulos de dados a construir:

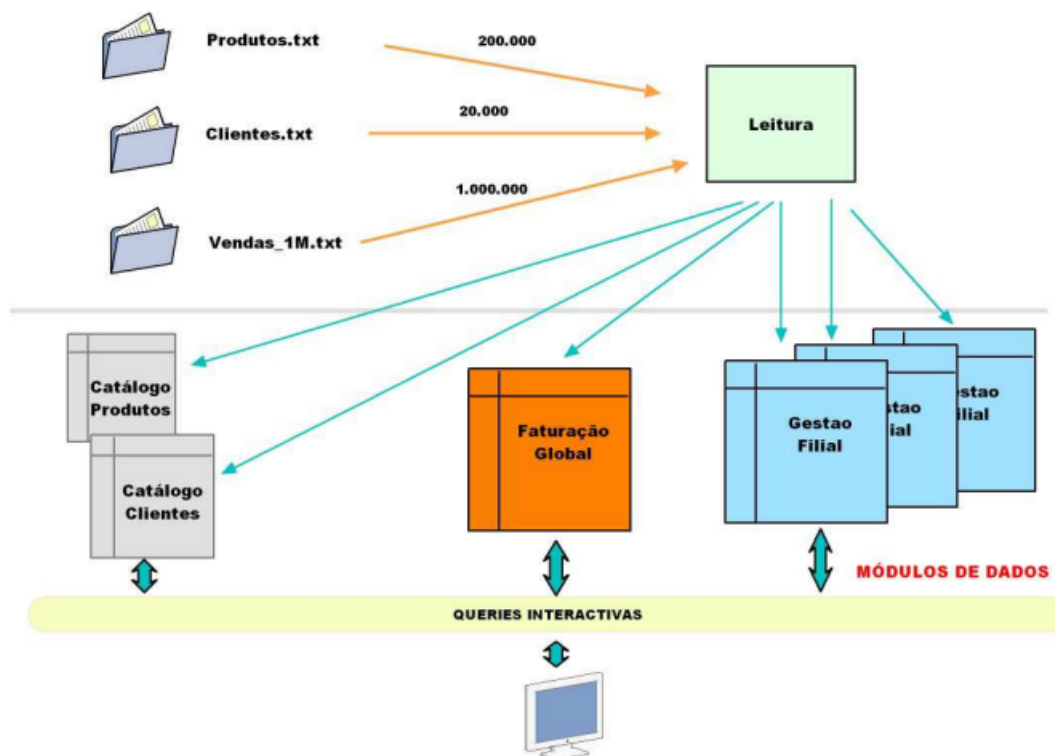


Figura 2.1: Arquitetura da aplicação

3. Módulos de dados

3.1 Catálogo de Clientes

É o módulo de dados onde são guardados os códigos de todos os clientes do ficheiro **Cientes.txt**, organizados por índice alfabético;

```
private Set<Cliente> catalogoClientes;
```

3.2 Catálogo de Produtos

Módulo de dados onde são guardados os códigos de todos os produtos do ficheiro **Produtos.txt**, organizados por índice alfabético, o que irá permitir, de forma eficaz, saber quais são os produtos cujos códigos começam por uma dada letra do alfabeto e saber quantos produtos são contabilizados.

```
private Set<Produto> catalogoProdutos;
```

3.3 Faturação Global

Módulo de dados que contém as estruturas de dados responsáveis pela resposta a questões quantitativas que relacionam os produtos às suas vendas mensais, em modo Normal (N) ou em Promoção (P), para cada um dos casos guardando o número de vendas e o valor total de faturação de cada um destes tipos. Este módulo refecencia todos os produtos, mesmo os que nunca foram vendidos, não contém qualquer referência a clientes, mas é capaz de distinguir os valores obtidos em cada filial.

3.4 Gestão da Filial

Módulo de dados que, a partir dos ficheiros lidos, contém as estruturas de dados adequadas à representação dos relacionamentos, fundamentais para a aplicação, entre produtos e clientes, ou seja, para cada produto, saber quais os clientes que o compraram, quantas unidades cada um comprou, o mês e a filial.

Para a estruturação otimizada dos dados deste módulo de dados tivemos em atenção que pretendemos ter o histórico de vendas organizado por filial para uma melhor análise, nunca esquecendo que existem 3 filiais nesta cadeia.

4. Interface com utilizador

Nesta secção apresenta-se a interface com o utilizador, fazendo algumas considerações sobre as decisões tomadas. Ao iniciar o programa, o utilizador tem um menu que lhe permite escolher 4 opções para ler os ficheiros:

5. Resultados e comentários sobre os testes de performance

Depois de desenvolver e codificar todo o projeto foi-nos proposto realizar alguns testes de performance que consistem em comparar os tempos de execução das queries 8, 9, 10, 11 e 12 usando os ficheiros Vendas_1M.txt (1000 000 vendas), Vendas_3M.txt (3 milhões de vendas) e Vendas_5M.txt (5 milhões de vendas).

6. Conclusão

Uma vez que se tratou de um trabalho de uma dimensão já considerável comparando com o que estávamos habituados envolveu utilização de técnicas particulares e tivemos sempre como objetivo que este trabalho fosse concebido de modo a que seja facilmente modificável, e seja, apesar da complexidade, o mais optimizado possível a todos os níveis.

Inicialmente, tivemos dificuldades nas AVLs pois estávamos a fazer uma AVL para cada módulo. Depois de alguns problemas com o seu balanceamento, acabamos por apostar na utilização da biblioteca standard AVL da GNU, que nos facilitou não só o carregamento dos ficheiros em memória, mas também na realização de algumas queries, devido ao vasto conjunto de úteis funções que a biblioteca contém, evitando assim a repetição de código.

Tivemos dificuldades em conseguir resultados quando usámos o módulo das filiais e era necessário cruzar os dados com as 3 filiais existentes, pois o módulo filial era apenas para uma filial.

Em suma, podemos concluir que embora todas as queries estejam a funcionar corretamente há aspectos na interface com o utilizador que poderiam ser melhorados.