

UNIVERSIDADE DO PORTO
FACULDADE DE ENGENHARIA

Análise de Séries Temporais com Recurso
a Técnicas de Bases de Dados

César Alexandre Fernandes Mendes Garcia

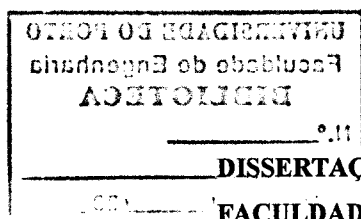
2000



ANÁLISE DE SÉRIES TEMPORAIS COM RECURSO A TÉCNICAS DE BASES DE DADOS

por

César Alexandre Fernandes Mendes Garcia



**DISSERTAÇÃO DE MESTRADO APRESENTADA À
FACULDADE DE ENGENHARIA DA UNIVERSI-
DADE DO PORTO PARA OBTENÇÃO DO GRAU
DE MESTRE EM INTELIGÊNCIA ARTIFICIAL E
COMPUTAÇÃO**

PROF. ORIENTADOR:

PROFESSOR JOÃO CORREIA LOPES

***Faculdade de Engenharia da Universidade do Porto
Outubro de 2000***

UNIVERSIDADE DO PORTO
Faculdade de Engenharia
BIBLIOTECA M
N.º 51694
CDU 681.3(843)
Data 12, 12, 2000

Resumo

A evolução do mercado electrónico em Portugal levou a que neste momento já seja possível ao utilizador comum da *Internet* proceder à aquisição e venda de acções de empresas cotadas na Bolsa de Valores de Lisboa em tempo real (online) através de vários sítios existentes para o efeito.

No âmbito deste trabalho desenvolveu-se uma aplicação em que o eventual investidor, baseado apenas no conhecimento histórico do valor das cotações das acções (séries temporais), com alguma segurança, efectuar previsões sobre o comportamento futuro de determinadas acções e assim encontrar oportunidades de negócio em tempo real.

O trabalho efectuado envolveu a criação de uma base de dados para armazenamento do valor das cotações diárias das acções, a recolha de informação, a verificação e eventual adequação dos valores e a inserção dos mesmos na base de dados.

A *interface* com o utilizador permite a visualização gráfica do comportamento das séries das cotações das acções e a opção de seleccionar pares de acções a analisar inseridas, ou não, dentro do mesmo ramo de actividade e a parametrização do factor de correlação e período a considerar para efeitos de simulação.

A adequação do método utilizado para a análise das séries temporais é nesta dissertação avaliado, através da apresentação de alguns exemplos de simulações de investimento efectuadas, podendo-se concluir que, para um investidor que não possua grandes conhecimentos sobre o funcionamento do mercado de acções, a utilização desta ferramenta cumulativamente com um acompanhamento diário da evolução das cotações, pode permitir a realização de investimentos neste tipo de produto (acções) com um factor de risco mínimo.

Abstract

The evolution of the Portuguese electronic market made it possible for the ordinary Internet user to buy and sell stocks in the Lisbon Stock Exchange online through several specific sites.

In this work an application was developed in which the potential investor, based only on the historical knowledge of the stock quotation (temporal series) can, with some reliability, make predictions about the future variation of certain stocks quotations and so find business opportunities online.

In this work a database was created for the loading of daily stock quotations, information gathering, verification and values fitness and their insertion in the database.

The user interface allows the graphic visualization of stocks quotation series behaviour and the option of selecting pairs of stocks to analyse, whether they are or not in the same area of activity, the measurement of the correlation factor and the period of time for simulation effects.

The adaptation of the method used for the analyses of the temporal series is here evaluated through the presentation of some examples of investment simulation. So, for an investor with low knowledge about the functioning of the stock exchange market the use of this tool, along with a daily following of stocks quotation evolution, may allow the investment in this kind of product (stocks) with a minimum risk factor.

Synthèse

L'évolution du marché électronique au Portugal a poussé, à ce moment, ceux qui utilisent l'Internet à faire l'acquisition et la vente des actions des entreprises de cote de la Bourse de Valeurs de Lisbonne, en temps réel (online) en plusieurs endroits qui existent par l'effet.

Dans le champ d'action de ce travail, on a développé une mise en place où l'éventuel acquéreur, fondé seulement à la connaissance historique de la valeur des cotations des actions (des séries temporelles), puisse avec une certaine assurance, faire des prévisions du comportement futur de quelques actions et prendre la balle au bond.

Le travail réalisé a permis la création d'une banque d'information pour stocker la valeur des cotations quotidiennes des actions, la récolte de l'information, la vérification et la possibilité d'ajuster les valeurs et les mettre à la banque de données.

L'interface avec l'utilisateur permet de visualiser graphiquement le comportement des séries des cotations des actions et l'option de sélection de paires d'actions à analyser introduites, ou non, dans la même partie d'activité et établir le paramètre du facteur de corrélation et la période à regarder pour des effets de simulation.

L'ajustement de la méthode utilisée pour l'analyse des séries temporelles est à cette dissertation évalué, par la présentation de quelques exemples de simulation d'acquisitions réalisées et nous pouvons conclure qu'une personne qui n'ait pas de grandes connaissances du fonctionnement du marché des actions, avec l'utilisation de cet outil cumulativement avec la suite quotidienne de l'évolution des cotations, peut réaliser l'acquisition de cette sorte de produit (des actions) avec une possibilité minimum de risquer.

Agradecimentos

Não podia deixar de expressar o meu profundo agradecimento a todos aqueles que, de alguma forma, contribuíram para realizar e terminar este trabalho.

Os meus mais sinceros agradecimentos ao Professor João Correia Lopes, a quem devo a escolha do tema para esta dissertação, e que foi inexcedível no acompanhamento deste trabalho. Foi graças às suas palavras de estímulo e indicações que me foi dando durante a orientação da dissertação que foi possível finalizar este trabalho.

Ao Professor Eugénio de Oliveira pela compreensão e atenção que teve comigo e a quem devo o facto de ter sido possível realizar este trabalho, mostrando-se sempre disponível e dando-me coragem numa fase algo “complicada” do curso.

À Altina pelas horas em que “estive ausente”, tendo-me sempre encorajado e aceitado esta minha “ausência”, assumindo muitas vezes os meus deveres de pai para com a Joana e a Ana Rita (que ainda só tem um ano ☺).

Por último gostaria de agradecer também a todo o corpo docente do mestrado e aos meus colegas, especialmente ao Carlos Ramos e ao Paulo Dias pela força que sempre transmitiram para poder concluir esta tese.

Índice

Resumo.....	II
Agradecimentos.....	V
Índice.....	IV
1. Introdução.....	1
1.1- Motivação-----	3
1.2- Objectivos-----	4
1.3- Estrutura do relatório-----	5
2. Análise de Séries Temporais.....	6
2.1- Introdução-----	6
2.2- Séries Temporais-----	6
2.3- Factor de Correlação-----	7
2.4- Médias Móveis-----	9
2.5- Resumo-----	10
3. Concepção e Desenvolvimento da Aplicação.....	12
3.1 - Análise de requisitos-----	12
3.1.1- Modelo Entidade/Associação-----	13
3.2 - Base de Dados-----	15
3.3 - K/KDB-----	16
3.3.1 - Introdução-----	16
3.3.2 - KDB-----	17
3.3.2.1- KSL-----	19
3.4 - Recolha de Informação-----	20
3.4.1- Parsing.....	23
3.5- Resumo-----	24

4. Técnicas de Acesso e Manipulação de Dados	26
4.1- Interrogações à Base de Dados	26
4.1.1- JAVA/SQL (ANSI 92)	27
4.1.2- JAVA/KSQL	30
4.1.2.1- Acesso à base de dados	31
4.1.2.2- Perguntas à base de dados	31
4.1.2.3- Cálculo do factor de correlação	33
4.2- Regras de Análise dos Dados	34
4.2.1- Algoritmo ID3	35
4.2.1.1- Séries altamente correlacionadas(valor positivo)	35
4.2.1.2- Séries pouco ou nada correlacionadas	37
4.2.1.3- Séries altamente correlacionadas (valor negativo)	37
4.3- Interpretação dos Resultados	41
4.4- Resumo	45
5. Avaliação da Tecnologia Utilizada	46
5.1- K/KDB	46
5.1.1- K (Desempenho)	47
5.2- Bases de Dados - Factores que influenciam negativamente o desempenho dos SGBD	48
5.3- Transacções em KDB	49
5.4- Conclusão	51
6. Satisfação de Objectivos	53
6.1 - Dificuldade de obtenção de dados	53
6.2 - Resultados obtidos	54
6.2.1- Exemplo de simulação 1	56
6.2.2- Exemplo de simulação 2	58
6.2.3- Exemplo de simulação 3	61
6.2.4- Exemplo de simulação 4	64
6.3- Conclusões	67
6.5 - Resumo	69

7. Conclusão.....	71
7.1- Satisfação de Objectivos -----	70
7.2- Perspectivas de Desenvolvimento Futuro-----	72
7.3- Consideração Final -----	73
BIBLIOGRAFIA.....	75
Anexo 1	77
Anexo 2	92
Anexo 3	99

Índice de Figuras

Figura 2.1- Factor de Correlação, Desvio Padrão e Covariância	7
Figura 3.1- Modelo Entidade/Associação	14
Figura 3.2- Notação usada no modelo Entidade/Associação	15
Figura 3.3- Sistema OLTAP.....	17
Figura 3.4- Recolha de Informação	21
Figura 4.1- Algoritmo ID3 - Árvore de Decisão.....	40
Figura 4.2- Parametrização do factor de correlação e nº dias para efeitos de análise	41
Figura 4.3- Selecção do ramo de actividade.....	42
Figura 4.4- Selecção específica de duas empresas.....	43
Figura 4.5- Gráfico linear da evolução das séries temporais	43
Figura 5.1- Algoritmo OBEYORDER.....	51
Figura 6.1- Gráficos das empresas Centralcer e Sumolis	57
Figura 6.2- Gráficos das empresas Portugal Telecom e Telecel	60
Figura 6.3- Gráficos das empresas BCP e BTA	62
Figura 6.4- Gráficos das empresas Mota & Cª e Teixeira Duarte.....	65

Índice de Tabelas

Tabela 2.1- Períodos de tempo a considerar para cálculo da média móvel	10
Tabela 3.1- Extracto das cotações de acções da Bolsa de Valores de Lisboa para 25-06-1999.....	22
Tabela 4.1- Análise do factor de correlação (valor positivo).....	38
Tabela 4.2- Análise do factor de correlação (valor negativo).....	39
Tabela 5.1- Comparação entre KDB e Sybase.....	48
Tabela 6.1- Preço das acções das empresas Centralcer e Sumolis.....	56
Tabela 6.2- Preço das acções das empresas Portugal Telecom e Telecel.....	60
Tabela 6.3- Preço das acções das empresas BCP e BTA	62
Tabela 6.4- Preço das acções das empresas Mota & C ^a e Teixeira Duarte.....	64
Tabela 6.5- Simulação da Evolução do Investimento	68

Índice de Programas

Programa 3.1- Verificação dos valores nulos	24
Programa 4.1- Acesso à base de dados e manipulação do conjunto de registos	28
Programa 4.2 - Cálculo do Desvio Padrão	28
Programa 4.3- Cálculo da Covariância	29
Programa 4.4 - Métodos implementados na classe c	31
Programa 4.5- Código em KSQL para saber o mínimo de registos das séries.....	32
Programa 4.6 - Código em KSQL para calcular o número de registos	32
Programa 4.7- Instrução em KSQL para calcular as médias móveis de 5 e 21.....	33
Programa 4.8- Verificação do cruzamento das médias móveis	40

Capítulo 1

Introdução

A análise de séries temporais, ou seja, o “estudo do comportamento de uma variável ordenada de uma forma cronológica - semanal, mensal, anual,...-“ [SHA99] é fundamental no apoio à tomada de decisões e definição de estratégias, quer a nível empresarial, financeiro ou outro.

Quase todas as organizações, independentemente do ramo de actividade em que estão inseridas, usam previsões na fase de tomada de decisões.

A previsão mostra-se imprescindível e, uma organização que não corra alguns riscos na definição de estratégias, está de alguma forma condenada ao insucesso. Daí a necessidade de possuir um instrumento que permita calcular com rigor os riscos com base em previsões.

As previsões podem efectuar-se com base quer no conhecimento e intuição pessoal sobre determinada matéria, quer recorrendo a métodos que consistem na análise de dados históricos.

Nestes métodos é assumido que o conhecimento do passado e as experiências obtidas a partir desse conhecimento é útil pois as experiências serão muito provavelmente repetidas no futuro [SHA99].

A grande quantidade de dados armazenados nas bases de dados das mais diferenciadas áreas – economia, medicina, engenharia, etc. - permitem-nos realizar estudos no intuito de, com base no conhecimento adquirido no passado, podermos extrapolar para o futuro a ocorrência de determinados acontecimentos.

A grande questão que se coloca neste tipo de análise é a forma de tratar todos estes dados já que, e apesar dos inúmeros estudos efectuados nesta área, continua a ser extremamente difícil aplicar algoritmos que nos dêem previsões com um grau de certeza absoluto.

Devido ao grande volume de informação existente, não é fácil analisar com precisão todos os dados, residindo a grande dificuldade na capacidade humana para filtrar a informação que de facto é útil e analisá-la da melhor forma [EMP99] sendo por isso admissível que hajam falhas.

A necessidade de obtenção de uma resposta imediata (*online*) e a manipulação de grandes quantidades de informação (que ainda não é uma tarefa fácil apesar da constante evolução da velocidade de processamento das actuais máquinas) leva a que se tentem encontrar soluções de optimização do tempo de resposta das perguntas às bases de dados onde reside a informação.

A manipulação e tratamento da informação é um processo moroso que exige não só tecnologia eficaz como também a aplicação de alguns métodos de análise e estudo cuja computação é demorada.

A crescente evolução de que tem sido objecto a Internet leva a que não seja possível a qualquer organização alhear-se da sua existência visto que este se mostra “um meio de negócio” imprescindível para o sucesso e desenvolvimento económico.

Exemplo disto são as recentes previsões do volume de negócios nos próximos três anos que deverá ser 100 vezes superior aos cerca de cinco milhões de contos que se regista actualmente¹.

Assim, não é de estranhar que recentemente tenham surgido algumas organizações a disponibilizarem meios para a realização de operações de investimento directo via Internet na Bolsa de Valores de Lisboa – casos do BPI, Banco Melo, BES...

A disponibilização de uma ferramenta de análise que permita dar uma resposta em tempo útil a uma possível operação de investimento no mercado de acções da Bolsa de Valores de Lisboa é, essencial e necessária para este tipo de negócio, já que este é um ramo de actividade económica onde a análise das cotações através das respectivas séries de valores se pode mostrar essencial para o sucesso de uma operação de investimento.

¹ Metas do Governo para a Internet até 2003, publicação do jornal Expresso de 19-08-2000

O recurso ao estudo das séries temporais no mercado bolsista tem sido aliás um campo de investigação importante nos últimos anos, quer através da implementação de algoritmos de pesquisa e análise quer da implementação de técnicas de acesso e manipulação de dados.

Neste trabalho pretende-se, através do recurso a técnicas para acesso e manipulação de dados, proceder à análise de séries de acções da Bolsa de Valores de Lisboa e assim tentar encontrar oportunidades de negócio.

1.1 – Motivação

Este trabalho tem duas componentes distintas que acarretam também motivações diferentes:

- O estudo de séries temporais é uma área interessante já que existem diversas áreas de actuação onde é possível aplicar os métodos a que se refere este trabalho e definir, com algum rigor, estratégias de previsão com base nesse estudo.

Mesmo para um não investidor, a previsão do comportamento das acções é uma área de interesse do ponto de vista económico – estudar possibilidades de investimento em acções na Bolsa de Valores de Lisboa, minimizando a possibilidade de risco.

- O estudo de técnicas que permitem o acesso e manipulação de dados com maior rapidez.

O volume de informação armazenado em bases de dados é e tem tendência a ser cada vez mais elevado, o que obriga não só a uma actualização permanente do parque informático que suporte os requisitos exigidos pelos actuais sistemas de gestão de bases de dados (SGBD) como também à implementação de técnicas que permitam dar resposta com eficiência às solicitações feitas às bases de dados.

Esta é também uma área que designaria de ordem afectiva, já que está perfeitamente enquadrada com a minha actividade profissional, sendo por isso extremamente motivante para mim poder desenvolver um trabalho neste campo de actuação.

1.2- Objectivos

Neste trabalho pretende-se, através da implementação de técnicas para acesso e manipulação de dados, proceder à análise das cotações das acções das empresas cotadas na Bolsa de Valores de Lisboa e tentar encontrar oportunidades de negócio. Para tal, foram definidos os seguintes objectivos:

- Análise das séries temporais das cotações das acções, através do estudo do comportamento diário do seu valor e do factor de correlação entre elas;
- Estudo das tendências das séries temporais, através do cálculo das médias móveis (média móvel de 5 e 21);
- Estudo e implementação de técnicas de acesso e manipulação da informação armazenada na base de dados.

O tipo de análise efectuada não engloba todos os parâmetros que julgo necessários e que permitiriam prever com maior certeza como vai ser o comportamento das acções de uma empresa qualquer. Tal não é possível porque:

- Uma análise aprofundada exigia conhecimentos técnicos e científicos da área de Economia, para além das técnicas das Bases de Dados;
- Existem vários factores internos às próprias empresas (dados económicos das empresas, estratégias de gestão, etc.) aos quais não é possível ter acesso;
- Existem factores externos que influenciam, ou podem influenciar, o comportamento, não só de determinado tipo de acções, mas até da(s) própria(s) Bolsa(s) de Valores— exemplo disso são as quedas generalizadas das bolsas muitas vezes influenciadas por factores totalmente alheios ao seu normal funcionamento.

A correlação entre o comportamento das principais bolsas mundiais tem sido aliás objecto de estudo [WRT94], verificando-se uma grande correlação entre as principais bolsas (Nova Iorque e Tóquio), verificando-se sempre uma subida ou descida em função de factores como por exemplo a valorização ou desvalorização do Dólar, etc.

Com este trabalho espero de alguma forma contribuir para o desenvolvimento de novas ferramentas que permitam a um investidor que não possua grandes conhecimentos sobre o funcionamento do mercado de acções e, que mesmo assim gostaria de investir neste tipo de

produto, a possibilidade de realizar alguns investimentos na Bolsa de Valores de Lisboa, minimizando o risco inerente a estes investimentos.

1.3- Estrutura do relatório

Este relatório está organizado em sete capítulos. Neste capítulo é feita uma introdução ao tema desta tese e aos objectivos que levaram à sua elaboração.

No capítulo seguinte é dada uma descrição sobre as séries temporais, o interesse do seu estudo para este trabalho e a forma como, através do seu estudo, calcular as tendências de subida ou descida dos valores das acções.

No capítulo 3 é referida a ferramenta de bases de dados a utilizar neste trabalho, a concepção da base de dados e, alguns pormenores da linguagem utilizada para acesso e manipulação dos dados – o KSQL.

No capítulo 4, através de excertos de código retirados da aplicação desenvolvida, são descritas as técnicas utilizadas para o acesso e manipulação dos dados, referidos os critérios utilizados para efeitos de análise das séries temporais e a interpretação dos resultados.

No capítulo 5 é feita uma análise de desempenho da ferramenta utilizada, suas características e particularidades, e no capítulo 6, através de exemplos de simulação efectuados, é feita uma avaliação dos resultados obtidos.

Finalmente no capítulo 7 é feita a avaliação deste trabalho e referidas as perspectivas de desenvolvimento futuro.

Capítulo 2

Análise de Séries Temporais

Introdução

Neste capítulo pretende-se fazer uma abordagem ao estudo das séries temporais, à importância do estudo do factor de correlação para análise e realização de previsões, metodologia utilizada para o tipo de séries objecto de estudo e ainda o significado e importância das médias móveis para uma correcta interpretação dos resultados obtidos.

2.1 – Séries Temporais

Uma série temporal não é mais que uma sequência de dados ordenada cronologicamente (anual, mensal, semanal, diária, horária, etc.) [SHA99].

O objectivo do estudo de uma série temporal é tentar descobrir padrões nos dados históricos e extrapolar esse padrão para o futuro.

A grande vantagem de utilizar as séries temporais para a análise e previsão é a de não haver necessidade de descobrir ou conhecer os factores que afectam o seu comportamento.

Para este estudo, esta vantagem é ainda mais relevante se considerarmos que nas variações a que estão sujeitas as cotações das acções, não é fácil identificar relações de causa - efeito tornando-se muito difícil, senão mesmo quase impossível, conhecer todos os parâmetros necessários a uma correcta análise. Este método é adequado para fazer este tipo de previsões.

Por outro lado, o que se afigura importante é encontrar uma oportunidade de negócio independentemente da forma como esse objectivo é conseguido, ou seja, o importante é prever o que acontece e não como acontece.

Dos vários componentes que constituem, ou podem constituir, a série, entre os quais se podem referir:

- A sazonalidade da série em que se nota uma repetição do padrão da série de k em k períodos em que k é o factor sazonal;
- O ciclo da série, em que o padrão da série varia de uma forma cíclica;
- A aleatoriedade da série, ou seja, o comportamento da série para o qual não se encontra explicação já que não é possível efectuar qualquer tipo de previsão sobre a mesma – daí talvez a justificação para uma possível falha da previsão;
- A tendência da série que nos mostra o crescimento ou declínio dos valores da série,

neste estudo apenas foi considerado este último componente – a tendência já que o objectivo consiste apenas em encontrar uma oportunidade de negócio em tempo real – *comprar e vender* e não uma previsão para um investimento a médio ou longo prazo.

A tendência de subida ou descida do valor de uma determinada acção a cada instante de tempo é pois o principal objectivo a considerar .

2.2- Factor de Correlação

O factor de correlação é talvez a mais poderosa ferramenta de análise que podemos utilizar no estudo das series temporais.

A correlação [CSC89] mede a relação entre dois conjuntos de dados escalonados de forma a serem independentes da unidade de medida. O seu cálculo devolve-nos a covariância dos dois conjuntos de dados divididos pelo produto dos seus desvios padrão:

$$\sigma_Y^2 = \frac{1}{n} \sum (Y_j - \mu_Y)^2 \qquad \sigma_X^2 = \frac{1}{n} \sum (X_j - \mu_X)^2$$

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \times \sigma_Y} \qquad \text{cov}(X,Y) = \frac{1}{n} \sum_{j=1}^n (X_j - \mu_X)(Y_j - \mu_Y)$$

Figura 2. 1 - Factor de Correlação, Desvio Padrão e Covariância

Em que

$$-1 \leq \text{Correlação}(x, y) \leq 1$$

De facto, o factor de correlação entre duas séries apenas pode variar no intervalo compreendido entre os valores 1 e -1, podendo-se extrair do seu valor uma das seguintes conclusões:

- valor do coeficiente é igual ou próximo de 1
 - Os valores mais altos de um conjunto estão associados aos valores mais altos do outro (correlação positiva), isto é, sempre que se verifica o aumento do valor de X, o valor de Y também aumenta.
- valor do coeficiente é igual ou próximo de -1
 - Os valores mais baixos de um conjunto estão associados aos valores mais altos do outro (correlação negativa), isto é, sempre que se verifica o aumento do valor de X, o valor de Y varia inversamente, ou seja, diminui e vice-versa.
- valor do coeficiente é igual ou próximo de 0
 - Sempre que se verifica o aumento do valor de X, o valor de Y varia de uma forma aleatória, ou seja, não podemos extrair qualquer conclusão.

Exemplificando neste trabalho, se por exemplo, se verificar que existe uma correlação entre as séries das acções da empresa X e da empresa Y fazendo ambas parte da minha carteira de acções (*portfolio*) [SHA99], e as acções da empresa X subiram de cotação e se as acções da empresa Y mantiveram a cotação, então a melhor atitude será comprar acções da empresa Y e vender as acções da empresa X, a não ser que existam factores que eu conheça (por exemplo o lançamento de uma OPA (Oferta Pública de Aquisição) ou de uma OPT (Oferta Pública de Troca) sobre as acções da empresa X) que tenham levado a essa variação e que desaconselhem a operação.

Uma das dificuldades com que se depara este tipo de análise é a selecção dos pares de acções a analisar, já que diversos critérios podem ser seguidos, como por exemplo:

- proceder ao estudo de todas as séries independentemente do ramo de actividade em que a empresa se encontra colocada;

- estudar apenas os pares de acções de empresas pertencentes ao mesmo ramo de actividade ou sector económico;
- qual o período de tempo a considerar para a análise;
- etc.

Atendendo à especificidade das séries analisadas neste trabalho, é de considerar que o passado recente é mais importante que o “histórico” [SHA99] e a variação ao longo de um período de tempo muito grande (anos) poderá servir apenas de mera indicação para uma análise mais aprofundada.

Por outro lado, é aceite que as séries das empresas inseridas no mesmo ramo de actividade estão mais correlacionadas do que as séries de empresas pertencentes a ramos de actividades diferentes [LJJ99], porque pertencendo ao mesmo ramo ou sector, estão sujeitas ao mesmo tipo de variação quando confrontadas com factores externos às próprias empresas.

Dentro do mesmo ramo de actividade há que considerar também o tipo de actividade, isto é, se as empresas estão inseridas em grupos mais homogéneos e menos susceptíveis de variações bruscas (caso da banca ou seguros em que se denota mais consistência e estabilidade nos dados) ou em grupos mais susceptíveis a variações bruscas (caso das novas tecnologias) e este é um facto a ter em consideração quando se pretende investir tendo em vista a minimização do risco.

2.3- Médias Móveis

A média móvel é um indicador que nos mostra a soma do preço do valor das acções durante determinado período de tempo, dividida pelo número de dias do período, sendo recalculada para cada dia do período considerado [FAT00].

A denominação de móvel advém do facto de que a cada novo período de tempo, consideramos um novo dado e abandonamos o mais antigo. Como a quantidade de dados a analisar é sempre a mesma, leva a que se abandone sempre o primeiro dado da série – o mais antigo.

Através da média móvel podemos constatar as mudanças de tendência e o desenvolvimento da série e assim, analisada cumulativamente com o factor de correlação, encontrar pontos de compra ou venda.

Existem diferentes tipos de médias móveis [SBA00]: simples (também conhecidas como aritméticas), exponenciais, triangulares, variáveis e ponderadas, podendo ser aplicadas a qualquer tipo de dados da série, ou seja, ao preço de fecho, ao preço máximo, ao preço mínimo, ao preço de abertura, ao volume de acções transaccionadas em determinado dia..

A diferença entre a aplicação dos diferentes tipos de médias móveis prende-se essencialmente com o período de tempo que se pretende analisar. Enquanto, por exemplo, a média aritmética dá igual peso a cada elemento da série, a média exponencial ou ponderada dá maior peso aos elementos da série numérica mais recente.

O período de tempo a considerar neste trabalho para cálculo das médias móveis é a média móvel de 5 e a média móvel de 21 [SHA99]. Este critério, que é discutível pois podem ser definidos outros períodos de tempo [JJM87], como por exemplo as médias móveis de 4, 9 e 18, enquadra-se dentro dos parâmetros aceites de uma forma generalizada, representados na tabela 2.1 [SBA00]:

Período	Média Móvel
Muito curto prazo	5-13 dias
Curto prazo	14-25 dias
Curto/Médio prazo	26-49 dias
Médio prazo	50-100 dias
Longo prazo	100-200 dias

Tabela 2. 1- Períodos de tempo a considerar para cálculo da média móvel

Da análise das duas médias podemos deduzir que quando se verifica que a primeira (a média de 5) cruza debaixo para cima sobre a segunda (a média de 21), tal significa que foi detectado um ponto de compra e se deve optar por comprar as acções objecto de análise. Inversamente, quando a primeira (a média de 5) cruza a segunda (a média de 21) de cima para baixo, existe um ponto de venda e então devem-se vender as acções [SHA99;JJM87].

2.4- Resumo

Neste capítulo foi feita, ainda que de uma forma sucinta, uma abordagem às séries temporais, foi referido o interesse do estudo das mesmas para alcançar os objectivos propostos

neste trabalho e as técnicas a utilizar para o efeito, como sejam o cálculo do factor de correlação entre pares de acções pertencentes a empresas inseridas no mesmo ramo de actividade, bem como as médias móveis de 5 e 21 dos mesmos pares de acções.

É óbvio que uma análise aprofundada sobre séries temporais, sua aplicabilidade e metodologias possíveis de implementar para a sua análise, é um campo de investigação inesgotável, no entanto esse não é o objectivo essencial deste trabalho e o tipo de análise a efectuar também não obriga a um aprofundamento desta matéria.

Capítulo 3

Concepção e Desenvolvimento da Aplicação

Introdução

Nos capítulos anteriores foram descritos os objectivos deste trabalho e a metodologia a utilizar para análise das séries temporais.

Neste capítulo é descrita a forma como foi concebida e desenvolvida a aplicação de suporte à previsão de compra e venda de acções. Será descrita a base de dados utilizada para efeitos de armazenamento dos valores diários das transacções e cotações da Bolsa de Valores de Lisboa, e serão referidas as ferramentas utilizadas para a concepção e desenvolvimento das várias fases de desenvolvimento deste trabalho.

No final será descrita a implementação do módulo cuja única função é diariamente descarregar os valores das transacções e das cotações da BVL.

3.1 – Análise de requisitos

Para poder estudar o comportamento ou variação do valor das acções, é essencial e imprescindível possuir um histórico sobre o qual possamos trabalhar. Neste estudo, o histórico não é mais que o valor diário das cotações das acções na Bolsa de Valores de Lisboa e, pontualmente, os valores dos dividendos¹, bem como os *splits*².

¹ - valor pago (por acção) periodicamente pela empresa aos seus accionistas proveniente dos lucros obtidos pela mesma.

Estes valores foram guardados numa base de dados concebida para o efeito, dado que não é possível aceder directamente a uma base de dados externa, como por exemplo à Bolsa de Valores de Lisboa, e a existência de um histórico é fundamental para este estudo.

A concepção da base de dados utilizada teve em consideração os requisitos necessários à posterior manipulação dos dados julgados necessários (para cada empresa cotada na Bolsa de Valores) para o estudo das séries temporais, como sejam:

- Quantidade de acções transaccionadas em determinado período;
- Cotação máxima das acções em determinado período;
- Cotação mínima das acções em determinado período;
- Valor dos dividendos pagos e data do pagamento;
- Valor do factor de *Split* (caso exista);
- Data em que são anunciados os pagamentos de dividendos;
- Data em que são anunciados os *Splits*.

A concepção da base de dados foi de certa forma condicionada pelos dados que é possível recolher pois a principal fonte de informação disponível é a Bolsa de Valores de Lisboa.

3.1.1 – Modelo Entidade/Associação

Com base nos requisitos e condicionalismos definidos no ponto anterior, o modelo entidade/associação da base de dados a implementar foi elaborado segundo uma notação representada por *Raghu Ramakrishnan*[RAR98] e encontra-se na Figura 3.1:

² - alteração do valor das acções após junção ou separação; por exemplo, a cada 5 acções passa a corresponder uma (factor de 0,2) levando o valor das acções a passar instantaneamente a 5x o valor actual; ou, a cada acção passam a corresponder 10 (factor de *split* 10) levando o valor das acções a passar instantaneamente a 1/10 do valor actual.

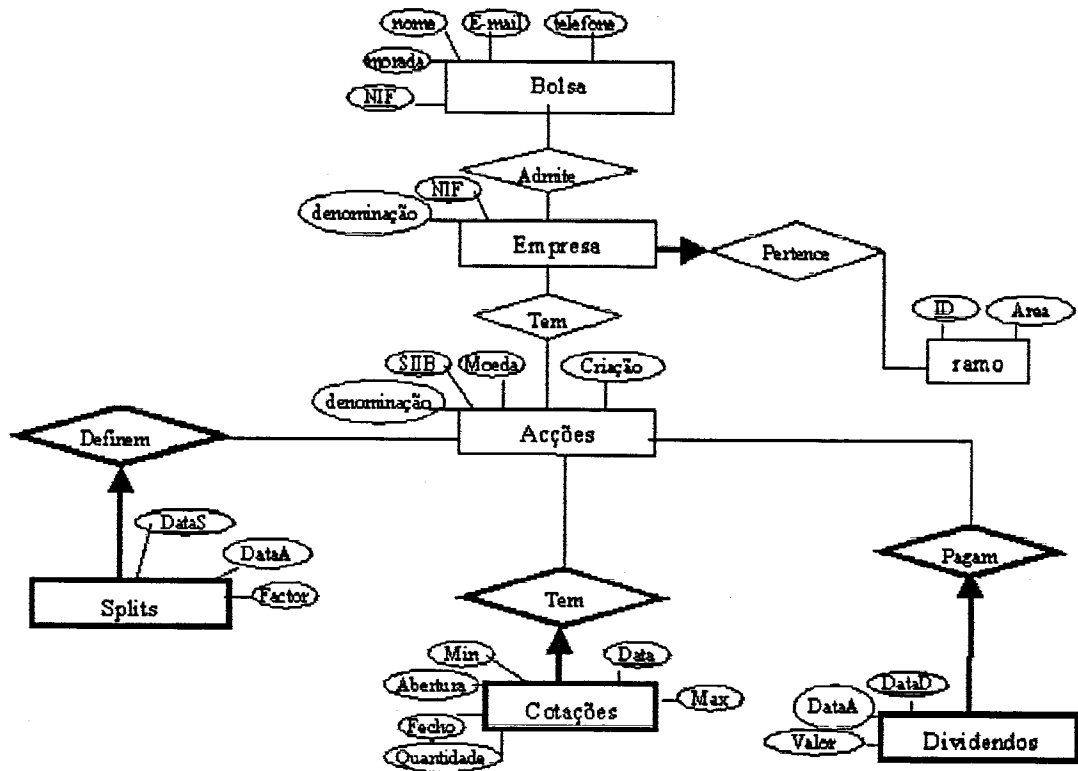


Figura 3.1- Modelo Entidade/Associação

No modelo Entidade/Associação representado na figura 3.1 estão descritas as entidades e associações entre as mesmas, bem como a informação a armazenar na base de dados respeitante a cada uma delas.

No modelo estão também representadas as restrições de integridade e as entidades fracas que se caracterizam pela dependência que têm das entidades donas, isto é, uma entidade fraca só pode ser identificada univocamente à custa da chave de outras entidades (dona) – neste modelo temos as entidades Dividendos, Cotações e Splits – que só existem porque existe a entidade dona – Acções. Na Figura 3.2 é pode-se ver a forma e o significado da representação utilizada no modelo Entidade/Associação.

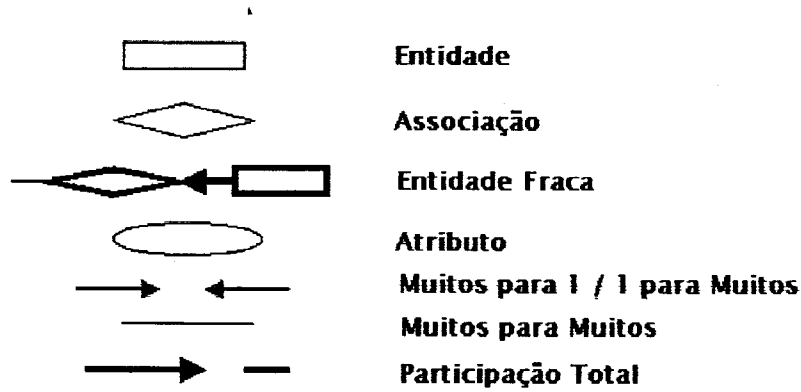


Figura 3.2 - Notação usada no modelo Entidade/Associação

3.2– Base de Dados

Para a realização deste trabalho foram utilizados dois sistemas de bases de dados distintos:

- Para o armazenamento e relacionamento da informação foi utilizado o MS Access² pela facilidade que oferece na manipulação de dados desde a importação de dados armazenados em tabelas, à formatação, criação de tabelas e relacionamentos já que, o objectivo era apenas ter uma base de dados onde armazenar a informação e à qual fosse possível aceder através de KDB/KSQL.
- Para interrogações à base de dados foi utilizada a KDB, pelas características que a mesma possui para a manipulação do tipo de dados armazenados (séries temporais), conforme é descrito no ponto seguinte.

² Sobre o MS Access, dada a sua generalização, penso que não é necessário introduzir aqui uma explicação sobre as características deste produto da Microsoft..

3.3– K/KDB

3.3.1 – Introdução

Uma das características de um sistema de base de dados é permitir ao utilizador fazer interrogações por forma a obter a informação desejada. Para tal poderá ser usada a linguagem SQL (*Structured Query Language*), sendo neste caso a resposta devolvida sob a forma de tuplos.

Esta resposta é devolvida contudo de uma forma que poderemos considerar bidimensional (cada linha contém uma instância de dados relativamente às categorias definidas nas colunas).

O problema que se nos coloca é que, neste modelo não se tira vantagem da ordem das linhas, isto é, pode-se de facto ordenar os dados a extrair (“*select <campo> from <tabela> order by <campo> ASC*”) mas depois não é possível explorar essa mesma ordenação.

Por exemplo, pode-se seleccionar todas as cotações das acções e ordenar as mesmas pelo respectivo preço, mas para saber qual o 20º preço mais alto, não é possível fazê-lo directamente. Para o conseguir teria que, através de sub-perguntas ir encontrar esse valor – o que técnica e eficazmente é de todo inviável, ou então socorrer-me de uma outra linguagem de programação de forma a poder manipular o resultado da consulta efectuada em SQL.

Outros exemplos se poderiam dar, como:

- encontrar o lugar que ocupa determinado elemento na ordenação efectuada,
- realizar o cálculo de uma média móvel,
- realizar o cálculo do factor de correlação,
- proceder à comparação anual entre os valores das transacções,
- etc.

Este tipo de problemas, associados ao facto de que as bases de dados actuais cada vez possuem armazenadas maiores quantidades de informação, leva a que, apesar do modelo relacional ser óptimo para a independência dos dados, mostra-se bastante moroso quando o problema

a tratar é, por exemplo, relacionado com séries temporais e neste caso interessa não só aceder aos dados ordenados segundo critérios predeterminados mas, e essencialmente, poder posteriormente proceder à sua análise (cálculo do factor de correlação, médias móveis, etc.).

3.3.2– KDB

KDB é um sistema de base de dados implementado em linguagem K³ e que reúne não só as características de um sistema tradicional de base de dados (com algumas particularidades como veremos mais à frente), com as características dos sistemas OLAP⁴ e OLTP⁵ a que, a *KX Systems* (entidade que desenvolveu o sistema) designou por sistema OLTP.

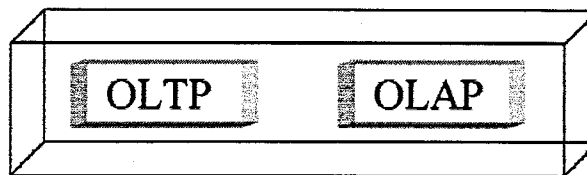


Figura 3.3 - Sistema OLTP

Especificamente concebida para uma utilização sobre grandes volumes de dados, em que a informação precisa de ser rapidamente organizada, analisada, armazenada e distribuída, tal como qualquer outro sistema de base de dados, KDB suporta como linguagem de acesso o ANSI-SQL II (vulgarmente denominado SQL 92).

Além do KSQL que é uma linguagem definida especificamente para responder às exigências que são colocadas quer aos sistemas OLAP quer aos sistemas OLTP, podendo a informação ser originária de outros sistemas de bases de dados.

A particularidade deste sistema reside no facto de, contrariamente à maioria dos sistemas de bases de dados (não me estou a referir como é óbvio aos outros sistemas OLAP, pois

³ K é uma linguagem de alto nível derivada (baseada) da APL, de NIAL e de LISP e que se caracteriza essencialmente pela forma como foram desenvolvidas e aperfeiçoadas as técnicas de manipulação de *arrays*.

⁴ OLAP (*online analytical processing*) são sistemas desenhados de forma a permitirem ao utilizador extrair e manipular de uma forma rápida grandes quantidades de dados armazenados em sistemas de bases de dados que contêm muita informação.

⁵ OLTP (*online transaction processing*) são aplicações desenvolvidas tipicamente em versão cliente/servidor que visam facilitar a consulta e manipulação de grandes quantidades de dados de uma forma tão rápida quanto possível.

por exemplo o S-Plus⁶ neste aspecto é muito semelhante), a estrutura dos dados é o *arrable* (array-table).

O *arrable* é uma tabela em que a ordem dos dados pode ser facilmente explorada. Por exemplo, um campo de um registo da BD, pode conter uma série temporal, em que a ordem da série corresponde à ordem dos registos.

A exemplo do que acontece noutras bases de dados utilizando o SQL, com o KSQL também é possível efectuar as operações clássicas de agregação, junção, selecção, etc.. Para além destas operações “clássicas” é ainda possível definir funções para manipulação dos *arrays*, como por exemplo *reverse* (função que inverte a ordem dos elementos) ou *rank* (que nos dá o lugar ocupado por determinado elemento), etc.

A importação de dados para KDB pode ser feita quer via ODBC, quer via JDBC [KXS00] - isto no caso de se pretender analisar dados provenientes de outros sistemas pois, como qualquer outro sistema de base de dados, KDB também tem a sua própria estrutura organizativa da informação, sendo as tabelas armazenadas sob a forma *CSV* (*comma separated value*) que é um formato de troca típico.

Como atrás referi, a manipulação dos dados pode ser efectuada quer através de SQL, quer através de KSQL. Para poder extrair toda a potencialidade do sistema, a linguagem que utilizei neste trabalho foi o KSQL – na parte correspondente à análise e estudo das séries temporais, tendo a interface sido desenvolvida em JAVA pelos seguintes motivos:

- O *driver* JDBC fornecido pela *KX Systems* que permite fazer a ligação ao servidor de base de dados K é do Tipo IV (100% *pure Java*) [MSO00] o que, em termos de compatibilidade e eficiência de acesso à base de dados torna o mesmo mais fácil e rápido.
- JAVA é uma linguagem que, além da portabilidade possui uma API que permite com relativa facilidade a implementação de uma interface [JZK98].

⁶ O S-Plus é um sistema derivado da linguagem S, desenvolvido na ATT Bell Laboratories e cujo desenvolvimento actualmente é da responsabilidade da MathSoft Inc.. Informação detalhada pode ser consultada em:

- <http://mathsoft.com/splus/>
- <http://math.umbc.edu/~arghya/Splus/splus.html>

3.3.2.1 – KSQL

O KSQL é muito parecido com o SQL, quer no tipo de operações que permite realizar e que atrás já foram referidas, quer na construção de instruções. A principal diferença consiste em que, em KDB as tabelas são ordenadas, ou seja, ao instanciar uma base de dados em KDB (*load <base_de_dados>*) as tabelas são automaticamente ordenadas e o KSQL explora esse facto.

Uma base de dados relacional é baseada no conceito matemático de relação. As tabelas são conjuntos de tuplos não existindo portanto, qualquer ordem pré-estabelecida para os dados devolvidos.

A única forma de efectuar uma consulta em SQL especificando a ordem pela qual os dados são devolvidos pelo servidor é através da instrução *order by* (*select <campo> from <tabela> order by <campo>*).

KDB é, neste aspecto, mais adequado já que os dados se encontram ordenados⁷ e é esta a vantagem que é posteriormente explorada pelo KSQL.

Socorrendo-se de funções já implementas na própria linguagem ou permitindo a sua implementação em linguagem K e em ficheiros com a extensão “.t” (no capítulo 4 esta questão será abordada de uma forma mais pormenorizada), perguntas à base de dados que em SQL não são possíveis, em KSQL fazem-se de uma forma directa, como por exemplo:

- *Select 5 avgs price by stock, date.month from trade* – que nos devolve directamente a média móvel de 5 por *stock* e por mês, ou
- *Select last 10 first price by stock from 'price' desc trade* – que nos devolve de uma forma directa o 10º preço mais alto para cada *stock*,

Outra vantagem do KSQL relativamente ao SQL é a simplificação de algumas perguntas como por exemplo, a forma como são abordadas as relações entre tabelas (chaves estrangeiras), permitindo simplificar as interrogações.

Exemplificando:

⁷ Deve-se ter em atenção o que atrás foi referido relativamente à KDB, ou seja, este sistema de base de dados foi implementado tendo por objectivo a análise de uma forma tão rápida quanto possível dos dados, sendo a sua área de aplicação essencialmente empresas de Telecomunicações, Bancos de Investimento, Companhias de Seguros, Universidades e outras instituições que têm necessidade de analisar grandes quantidades de informação, e é neste contexto que se faz esta afirmação.

Supondo que temos três tabelas *OrderItens* (*orderid*, ..), *Customers* (*custid*, *name*,...) e *Orders* (*orderid*, *custid*, ...) em que:

- *Customers* tem como chave primária *custid*;
- *OrderItens* tem *orderid* como chave estrangeira;
- *Orders* tem *orderid* como chave primária e *custid* como chave estrangeira

Em vez da instrução em SQL

```
... from OrderItens,Orders,Customers where
OrderItens.orderid = Orders.orderid and
Orders.custid = Customers.custid and
Customers.cname = ...
```

em KSQL esta instrução reduz-se a:

```
... from OrderItens where orderid.custid.cname = ....
```

É óbvio que esta notação só é possível porque na implementação da base de dados, a exemplo do que acontece em SQL, são referenciadas as chaves – primária e estrangeira – de cada tabela:

```
create table Orders(
  custid char(5) references Customers,
  orderid char(6),
  qty int,
  primary key(custid,orderid))
```

3.4 – Recolha de Informação

Como foi já referido, para a realização deste trabalho era imprescindível obter um histórico das cotações das acções das empresas e este era necessariamente o primeiro passo para o desenvolvimento deste projecto.

A principal “fonte de informação” a que tive acesso foi à Bolsa de Valores de Lisboa (<http://www.bvl.pt>).

De toda a informação disponível neste sítio Web, interessava essencialmente recolher a informação constante das páginas:

- "http://www2.bvl.pt/por/cotacoes_acc.html", e
- "http://www2.bvl.pt/por/transaccoes_acc.html"

que contêm respectivamente o valor diário das cotações das empresas cotadas em bolsa e o valor das transacções de acções registadas diariamente em relação a cada uma delas.

Relativamente ao índice BVL 30, a informação disponível na BVL está representada sob a forma de gráfico que pode ser visualizado e eventualmente descarregado.

Como a informação disponibilizada pela BVL se encontra no formato *HTML*, e não sendo possível recolher a informação noutro tipo de formato (penso que neste aspecto a BVL poderia proporcionar a informação em ficheiro com formato que facilitasse o seu uso), implicou que o desenvolvimento do módulo *BVLDOWN* responsável por descarregar diariamente o conteúdo das páginas tivesse os seguintes requisitos:

- Aceder via Internet (TCP-IP) às páginas atrás referenciadas;
- Descarregar o conteúdo das páginas para um ficheiro temporário;
- Fazer a verificação (*parsing*) do conteúdo do ficheiro temporário;
- Acrescentar os dados a duas tabelas criadas para o efeito – tabela das transacções e tabela das cotações.

Na figura 3.4 pode-se ver as diversas fases da recolha de informação, de acordo com os requisitos atrás referidos:

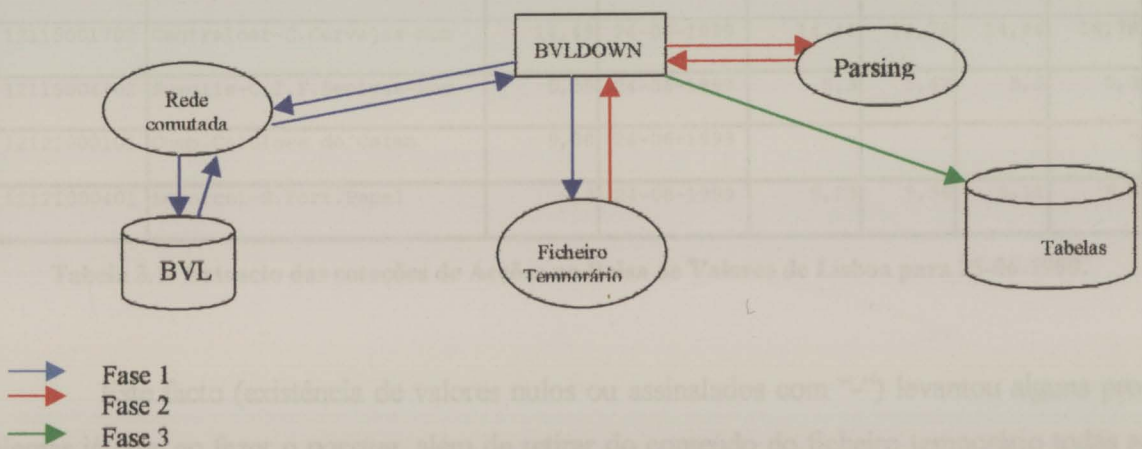


Figura 3.4 - Recolha de Informação

De todos os requisitos apontados, apenas vou abordar a parte do desenvolvimento referente ao *parsing* visto ter sido esta a fase foi onde residiu a maior dificuldade de implementação.

Para uma posterior análise dos dados, e atendendo às restrições impostas a nível de concepção das tabelas⁸, foi necessário ter em atenção o seguinte:

- Não poderia haver valores nulos, ou seja, independentemente da cotação diária das acções já que para muitas empresas nem sempre existem transacções diárias, deveria sempre constar um registo respeitante a esse dia, sendo o preço de fecho igual ao do dia anterior em que houve qualquer transacção.
- Para cada empresa cotada, independentemente de ter havido ou não transacções em determinado dia, o campo respeitante ao preço de abertura das acções deveria estar sempre preenchido.

3.4.1 – Parsing

Se analisarmos cuidadosamente o conteúdo das páginas da BVL podemos notar que nos dias em que não se verificarem quaisquer transacções relativamente a cada empresa, nos respectivos campos vem assinalado “-“ ou mesmo um valor nulo “”.

Cod. SIIB	Descrição	Cot.Ant.	Data	Abertura	Máximo	Mínimo	Fecho
12115001302	Unicer-União Cerv.-Nom	18,85	24-06-1999	18,85	18,87	18,61	18,85
12115001702	Centralcer-C.Cervejas-Nom	14,49	24-06-1999	14,44	14,78	14,44	14,78
12115004402	Sumolis-C.I.F.Bebidas-Nom	5,35	24-06-1999	5,3	5,43	5,3	5,3
12121000101	Comp.Celulose do Caima	9,38	24-06-1999	-	-	-	-
12121000401	Soporcel-S.Port.Papel	9,3	24-06-1999	9,23	9,38	9,11	9,3

Tabela 3.1- Extracto das cotações de Acções na Bolsa de Valores de Lisboa para 25-06-1999.

Este facto (existência de valores nulos ou assinalados com “-“) levantou alguns problemas já que, ao fazer o *parsing*, além de retirar do conteúdo do ficheiro temporário todas as *tags* que caracterizam a linguagem HTML e do tratamento dos valores nulos, foi necessário ter

⁸ De notar que, não sendo possível consultar on-line a evolução das transacções e cotações da Bolsa as actualizações são efectuadas apenas uma vez por dia (útil), normalmente após as 20:00 horas

em atenção a admissão a cotação de novas empresas – o que implicou a introdução de novos códigos SIIB e consequente enquadramento das mesmas no ramo de actividade.

Como é óbvio, apesar de todo o cuidado posto no desenvolvimento desta fase, é indispensável um acompanhamento pessoal para a inserção dos dados na base de dados, pois além da eliminação dos valores nulos há a necessidade de verificar a correcção dos dados.

Neste ponto deparei com outro tipo de problemas alheios ao normal funcionamento da bolsa, nomeadamente:

- empresas que apesar de cotadas nunca tiveram qualquer movimento (por exemplo as acções da SAD Sporting Club de Braga nunca tiveram qualquer transacção)
→ seria de manter as mesmas na base de dados ou eliminá-las?
- empresas que temporariamente vêm as suas acções suspensas – quer porque foram objecto de lançamento de ofertas públicas de aquisição quer porque a administração da BVL considerou (por qualquer motivo), que as devia suspender

A opção que tomei para realizar o *parsing* foi a seguinte:

Ao efectuar a leitura do ficheiro temporário e depois de retiradas as *tags* que necessariamente há que retirar, verifica se a nova linha da tabela começa por um código SIIB. Neste caso percorre o ficheiro até encontrar uma nova linha e, se encontrar um valor nulo ou “-“ substitui o mesmo por zero, inserindo posteriormente esta linha na tabela correspondente separando os campos por “ ”(tab):

```

e=s.indexOf("121");    //o código SIIB começa por 121,221,321,329
t=s.indexOf("221");
tres=s.indexOf("321");
quatro=s.indexOf("329");

// o campo SIIB tem sempre 11 caracteres de forma que se for o SIIB muda de
linha e acrescenta a data
if((p=s.length())==11) && ((e==0) || (t==0) || (tres==0) || (quatro==0) ) &&
teste)
{
    Calendar gc= Calendar.getInstance();
    Date data = new Date();
    // para preencher o campo respeitante à data em que é feito o download
    String minhadata = data.getDate() + "-" + (gc.get(Calendar.MONTH) + 1) + "-"
+ gc.get(Calendar.YEAR);
    bw.newLine();
    bw.write(minhadata+ "      " + s+ "      ");
}
else{
    // se não for SIIB e estiver com "-" altera para 0, senão substitui a virgula
por ponto
    if(teste){
        if((f=s.indexOf("-")==0){
            bw.write(s.replace('-', '0') + "      ");
        } else {
            bw.write(s.replace(',', '.') + "      ");
        }
    }
}
}

```

Programa 3.1- Verificação dos valores nulos

A inserção dos dados na base de dados é efectuada posteriormente através da importação das tabelas de texto. Nesta fase é verificada a correcção dos dados e efectuadas, quando necessário, as correspondentes alterações.

3.5– Resumo

Neste capítulo foram abordados os requisitos do sistema, o tipo de ferramenta utilizada para a manipulação e análise dos dados, e a linguagem associada (o KSQL) e a recolha de informação.

Relativamente ao modelo Entidade/Associação, o modelo apresentado é aquele que considero ideal para o tratamento deste tipo de informação. Neste trabalho, por vários motivos que à frente serão explicados, deparei com grande dificuldade na obtenção de alguns tipos de dados razão pela qual a base de dados não corresponde exactamente ao modelo descrito.

Relativamente à KDB foi dada uma visão genérica da mesma, os objectivos e vocação deste sistema de base de dados, tendo sido tratados apenas os pormenores que a tornam diferente de outros sistemas e propositadamente esquecidos alguns que, apesar de diferentes, penso não serem importantes para o trabalho efectuado – refiro-me especificamente à forma como são geridas as transacções⁹ em que nitidamente a *KX Systems* optou por explorar a atomicidade das operações dispensando assim o tradicional “gestor de transacções” [ADS99].

Relativamente ao KSQL, esta linguagem foi abordada também de uma forma sucinta porque no próximo capítulo na descrição da Interrogações à base de dados é dada, através de exemplos concretos, uma explicação mais pormenorizada desta linguagem.

Foi também abordado o ponto referente à recolha de informação, tendo descrito apenas a fase que requereu mais atenção na fase de desenvolvimento. Tendo sido esta a primeira fase deste trabalho, e sendo necessário proceder ao seu desenvolvimento no mais curto espaço de tempo possível, pormenores havia que provavelmente numa análise mais cuidada seriam tratados de outra forma, no entanto o objectivo que se impunha era apenas a criação de um “histórico” e esse objectivo foi alcançado.

⁹ No capítulo 5 será abordada esta questão para efeitos de comparação com outros sistemas de bases de dados.

Capítulo 4

Técnicas de Acesso e Manipulação de Dados

Introdução

Neste capítulo são abordadas as etapas de desenvolvimento deste trabalho na parte respeitante à análise dos dados e às técnicas utilizadas para o efeito, nomeadamente a utilização de funções implementadas em linguagem K e a utilização de KSQL para acesso e manipulação dos dados.

4.1 – Interrogações à Base de Dados

Para poder efectuar a análise dos dados, é necessário interrogar a base de dados. A exemplo do que sucede na grande maioria dos sistemas de bases de dados relacionais, é usada a linguagem SQL para efectuar esse tipo de interrogações e, como se trata de uma linguagem standard, são mínimas as derivações ou extensões que a mesma pode sofrer em função do sistema utilizado.

A linguagem JAVA utilizada neste trabalho, permite também o acesso a bases de dados relacionais através de *drivers* concebidos para o efeito – JDBC [KXS00], respeitando a norma SQL - ANSI 92. Neste trabalho foi utilizada, quer esta norma e consequentemente o *SQL*, quer o *KSQL* já referido no capítulo 3.

4.1.1 – JAVA/SQL (ANSI 92)

Como foi referido no capítulo 2, a análise e estudo das séries temporais é, neste trabalho, baseada essencialmente no estudo do comportamento das séries através do respectivo coeficiente de correlação.

Para conseguirmos chegar a esse valor há, no entanto, a necessidade de seguir determinados passos e pela ordem descrita:

1. Saber quantos registos diários existem na base de dados respeitantes a cada uma das empresas (acções) a analisar – o factor de correlação apenas pode ser calculado entre séries com o mesmo número de valores;
2. Realizar a pergunta à base de dados para obter a série respeitante à primeira empresa¹;
3. Repetir o procedimento anterior para a segunda empresa a analisar;
4. Para cada uma das séries calcular o respectivo desvio padrão;
5. Calcular a covariância das séries;
6. Calcular o factor de correlação.

Relativamente aos três primeiros passos apenas vou referir o aspecto respeitante à ligação ao servidor de bases de dados e execução da instrução em SQL, pois é neste aspecto que existem diferenças significativas entre a generalidade dos sistemas de bases de dados e o K.

Como atrás foi referido, o K aceita instruções quer em SQL (ANSI 92) quer em KSQL. Para aceder à base de dados através de instruções feitas em SQL, e atendendo que esta é uma linguagem standard e consequentemente qualquer servidor de bases de dados relacionais deve respeitar esta norma, o acesso é feito da “forma tradicional” isto é, acedendo via JDBC e manipulando posteriormente o conjunto de registos devolvidos (*recordset*). Exemplificando:

¹ Apesar de em SQL se poder realizar a pergunta apenas uma vez e retornar as duas séries, o que aumentaria sem qualquer dúvida a rapidez de execução já que evitava uma nova consulta à base de dados, o facto de nem sempre o número de registos relativos a cada série ser igual implica um tratamento diferenciado para cada uma das séries.

```

Class.forName("k.jdbc"); //carrega o driver
Connection c = DriverManager.getConnection ("jdbc:kx://localhost:1");
Statement e = c.createStatement();
ResultSet r = e.executeQuery(<expressão em SQL>);
while(r.next()) { //vai carregar o vector siibVector com todos os
elementos retornados pelo servidor
Object valor = r.getObject(1);
siibVector.addElement(Double.valueOf(valor.toString()) );
c.close(); // fecha a ligação
}

```

Programa 4.1- Acesso à base de dados e manipulação do conjunto de registos

Inevitavelmente para cada série é necessário percorrer todo este ciclo o que, em termos de eficácia de execução acarreta a perda de algum tempo e, por isso, prejudica o desempenho da aplicação.

Como os cálculos estatísticos referidos nos pontos 4, 5 e 6, não fazem parte da norma ANSI 92 não é possível através de instruções em SQL realizá-los.

Assim, para efectuar o cálculo do desvio padrão² implementei uma função que devolve esse valor recebendo como argumento um vector que contem toda a série:

```

public double calculadevpad(Vector xy){
double devfinal=0;
double somaquadrado=0;
double somavalores=0;
double divisor=0;
for (int i = 0; i < xy.size(); i++)
{
somaquadrado=somaquadrado+
Math.pow( (Double.valueOf(xy.elementAt(i).toString())
.doubleValue()),2);
somavalores=somavalores+(Double.valueOf(
xy.elementAt(i).toString()).doubleValue());
}
somaquadrado=xy.size()*somaquadrado;
divisor=Math.pow(xy.size(),2);
devfinal = Math.sqrt((somaquadrado-Math.pow(somavalores,2))/divisor);
return (devfinal);
}

```

Programa 4.2- Cálculo do Desvio Padrão

² O servidor K devolve sempre um objecto do tipo de dados armazenados na base de dados, obrigando portanto a algumas conversões entre diferentes tipos de dados (casting) para se poder manipular os dados em Java.

Antes da execução destas funções é também necessário verificar que o comprimento dos dois vectores é igual e respeita ao mesmo período, ou seja, não se pode analisar dados de uma série cujo primeiro valor seja de uma data diferente do 1º valor da outra série.

Para efectuar o cálculo da covariância, implementei também uma função que recebe como argumento dois vectores que contêm as séries temporais de cada uma das empresas em análise, e devolve o valor da covariância:

```
public double calculavariancia(Vector A, Vector B){
    double media1=0;
    double media2=0;
    double subtotal=0;
    double parcial=0;
    double covariancia=0;
    for (int i = 0; i < A.size(); i++){
        subtotal = subtotal +
        (Double.valueOf(A.elementAt(i).toString()).doubleValue());
    }
    media1=subtotal/A.size();
    subtotal=0;
    for (int i = 0; i < B.size(); i++){
        subtotal = subtotal +
        (Double.valueOf(B.elementAt(i).toString()).doubleValue());
    }
    media2=subtotal/B.size();
    for (int i = 0; i < A.size(); i++){
        parcial=parcial + (media1-
        (Double.valueOf(A.elementAt(i).toString()).doubleValue()))*(media2-
        (Double.valueOf(B.elementAt(i).toString()).doubleValue()));
    }
    covariancia=parcial/A.size();
    return(covariancia);
}
```

Programa 4.3 - Cálculo da Covariância

Na posse destes dois valores (o desvio padrão e a covariância), é possível determinar directamente o factor de correlação:

```
correlacao=Math.abs(covariancia/(desvio1*desvio2));
```

Relativamente a este ponto, é importante salientar a necessidade de execução de todos estes passos para se poder calcular o factor de correlação entre duas séries temporais.

Se analisarmos a parte de código descrita verifica-se que, além da conversão previamente efectuada do tipo de dados (*casting*), existem alguns ciclos destinados a percorrer os vectores que contêm as séries que inevitavelmente reduzem a eficiência e rapidez de execução da aplicação.

A análise que é efectuada com base nos dados calculados anteriormente será objecto de um estudo mais aprofundado no ponto 4.3.

4.1.2 – JAVA/KSQL

A utilização do servidor de base de dados K e da linguagem KSQL altera de uma forma muito significativa todo o processo de cálculo descrito no ponto anterior. Se há passos que inevitavelmente é preciso percorrer como em SQL, outros há que nos são devolvidos directamente pelo servidor : o desvio padrão, a covariância e o factor de correlação.

Como foi referido no capítulo 3, o servidor K aceita instruções implementadas em SQL e em KSQL. Além destas é possível implementar guiões em linguagem K.

Estes últimos mostram-se de grande utilidade pois é através deles que nos é possível efectuar este tipo de cálculos estatísticos.

Para que o servidor reconheça as funções implementadas pelo programador, ao fazer a inicialização do servidor, é preciso indicar-lhe também quais as funções que irá utilizar (isto no caso de se estarem a utilizar funções exteriores ao servidor pois a linguagem já possui algumas funções embutidas).

Exemplificando, se pretendo que o servidor me devolva o factor de correlação entre duas séries, devo iniciar previamente o servidor de funções K (“load funcoes.k -i <porta>”).

No ficheiro *funcoes.k* (em anexo), estão definidas as várias funções necessárias ao cálculo, não só do factor de correlação, como também das médias móveis que são utilizadas posteriormente para análise dos resultados.

Utilizando KSQL, os seis passos descritos no ponto anterior reduzem-se a apenas quatro evitando ainda o ajustamento dos vectores por forma a que contenham o mesmo número de valores:

1. Saber quantos registos diários existem na base de dados respeitantes a cada uma das empresas (acções) a analisar;
2. Realizar a pergunta à base de dados para obter a série respeitante à primeira empresa;
3. Repetir o procedimento anterior para a segunda empresa a analisar;
4. Calcular o factor de correlação.

Pela diferença que existe relativamente ao JAVA/SQL, penso que é importante descrever com algum pormenor cada um destes pontos já que é essencialmente neste aspecto que a KDB difere dos outros sistemas de bases de dados relativamente à rapidez de execução.

4.1.2.1 – Acesso à base de dados

Contrariamente ao que acontece em JAVA/SQL, para aceder à base de dados não é necessário passar por todos os procedimentos descritos no ponto anterior (código 4.1):

A forma mais simples que a *Kx Systems* encontrou para ultrapassar esta questão foi através da implementação de uma classe (classe *c*), que de uma forma directa nos permite aceder à base de dados³:

```
k.c c = new k.c("<servidor>", <porta>) //cria um objecto do tipo c
```

e para interrogar a base de dados apenas é necessário invocar o método da classe através da instrução:

```
Object[] nome = (Object[])c.k(instrução em SQL ou KSQL)
```

O resultado da interrogação à base de dados e a manipulação do mesmo é dado também através da invocação de alguns métodos implementados na classe *c*:

```
String[] n = (String[])nome[0] - que nos dá o nome das colunas resultado  
Object[] d=(Object[])nome[1] - que nos dá o valor das colunas  
String[] t = (String[])nome[2] - que nos dá o tipo de dados associado a cada  
coluna  
c.n(d[0])) - que nos dá o comprimento do vector
```

Programa 4.4 – Métodos implementados na classe *c*

4.1.2.2 – Perguntas à base de dados

Neste ponto a vantagem de utilização de KSQL em vez de SQL é notória porque, enquanto em SQL temos necessidade de perguntar à base de dados quais os valores das séries e posteriormente, através da implementação de um ciclo, ler todo o *recordset* e carregar um vec-

³ Pressupõe ter activado previamente o servidor .

tor de forma a que o número de elementos dos dois vectores (que contêm as séries), seja igual, em KSQL todo este processo se faz directamente.

Tendo conhecimento do número de registos de cada série existentes na base de dados, ao efectuar a pergunta sobre os valores das acções de uma empresa, apenas vamos perguntar pelo número de registos que de facto interessam.

Neste caso concreto, e na necessidade de “deitar fora” alguns registos, são eliminados sempre os registos mais antigos uma vez que não sendo igual o número de registos na base de dados, apenas duas situações podem ter acontecido:

- Ou a admissão a cotação na Bolsa de Valores de Lisboa foi efectuada em data diferente para as duas empresas em análise
 - Neste caso os registos respeitantes à primeira empresa a ser cotada até à data de admissão da segunda empresa, para efeitos de comparação das séries deixa de ter interesse pois não temos dados que nos permitam concluir qual o comportamento da segunda empresa.
- Ou uma das empresas deixou de estar cotada em Bolsa ou suspensa indefinidamente
 - Neste caso não há qualquer interesse em proceder à análise comparativa.

Exemplificando em KSQL:

Para saber qual o mínimo de registos na base de dados relativamente à empresa A e à empresa B, pode-se perguntar directamente através da instrução (ver código 4.4)

```
String minimelemento= "select min x,Descricao from (select count by
Descricao from Basc where Descricao in ('" + empA + "','"+ empB +"''))";
Object[]rmin = (Object[])c.k(minimelemento); // executa a pergunta
deitafora= (((int[])dmin[0])[0]); // guarda o valor
```

Programa 4.5 - Código em KSQL para saber o mínimo de registos das séries

Como é óbvio é preciso saber quantos registos existem no histórico relativamente a cada empresa. Essa pergunta é muito semelhante à que seria formulada em SQL:

```
String todos ="select count by Descricao from Basc where Descricao in ('" +
empA + "','"+ empB +"'')";
Object[]cor11 = (Object[])c.k(todos); // executa a instrução
Object[]d11 = (Object[])cor11[1]; // guarda o valor
```

Programa 4.6 - Código em KSQL para calcular o número de registos

Na posse destes dois valores (o menor e o valor total dos registos), é possível interrogar a base de dados de forma a serem devolvidos apenas os registos necessários para a análise.

Como o objectivo da pergunta é saber os valores a partir dos quais vão ser calculadas as médias móveis de 5 e 21, neste passo é possível explorar as características da linguagem e proceder simultaneamente à execução da pergunta e ao cálculo daquelas duas médias:

```
String pergunta = "select by " + (((int[])d11[1])[1]-deitafora) + " drop  
Datad, " + (((int[])d11[1])[1]-deitafora) + " drop Preco, "  
+ (((int[])d11[1])[1]-deitafora) + " drop 5 avgs Preco, "  
+ (((int[])d11[1])[1]-deitafora) + " drop 21 avgs Preco from Basc where  
Descricao = '" + empB + "'";
```

Programa 4.7 - Instrução em KSQL para calcular as médias móveis de 5 e 21

Nesta instrução a palavra reservada *drop* significa que o servidor ignora os primeiros $((int[])d11[1])[1]-deitafora$ registos e que as médias devem ser calculadas apenas sobre os registos devolvidos, ou seja, não é acumulada na soma, para efeitos de cálculo da média, a soma dos registos que foram ignorados.

4.1.2.3– Cálculo do factor de correlação

Para o cálculo do factor de correlação temos necessidade de nos socorrer de funções implementadas em linguagem K. Esta é aliás uma função que não está directamente ligada à especificidade que um servidor de base de dados deve ter enquanto tal, pois esta é a área de aplicação dos sistemas OLAP.

O facto de em KSQL ser possível efectuar alguns tipos de operações estatísticas, como no exemplo anterior, determinado tipo de funções têm que estar previamente implementadas em linguagem K.

Para aceder à função que nos permite obter directamente o factor de correlação de duas séries temporais, o procedimento não difere muito daquele que é necessário para aceder à base de dados e que foi descrito anteriormente.

Da mesma forma que é executada uma instrução (em SQL ou KSQL), pode-se chamar directamente a função pretendida:

```
Object xy= correlac.k("cor", (double[])d2[1], (double[])d1[1])
```

em que:

- *cor* é o nome da função previamente definida em linguagem K para cálculo do factor de correlação e
- *(double[])d2[1], double[])d1[1])* são os dois vectores que contêm as séries temporais,

sendo depois o valor dado directamente:

```
correlacao="" + (Double.valueOf(xy).toString()).doubleValue()
```

Comparativamente com o procedimento seguido em JAVA/SQL, os pontos 4, 5 e 6 descritos em 4.2.1 são substituídos por uma simples instrução o que, indubitavelmente, melhora a eficácia do sistema pois uma vez que os dados são devolvidos pelo servidor já de uma forma ordenada e ajustados à análise que se pretende efectuar (o comprimento das duas séries), não é necessário percorrer os ciclos descritos em 4.2.1.

Outras funcionalidades da KDB e as potencialidades do KSQL, serão abordadas no capítulo 5 pois a eficiência do sistema não passa apenas pela resposta do servidor às perguntas que lhe são endereçadas.

Como foi referido no capítulo 3, optei por armazenar os dados num sistema de base de dados diferente, daí que todas as funcionalidades do sistema que não tinham directamente a ver com consultas, não foram exploradas neste trabalho, no entanto será dada uma perspectiva das mesmas no próximo capítulo.

4.2– Regras de Análise dos Dados

No capítulo 2 foi já abordada a forma de análise e estudo das séries temporais e o objectivo do mesmo. No entanto, uma análise aprofundada dos valores, quer das médias móveis, quer do factor de correlação obriga a seguir alguns critérios tendo em vista não só alcançar os objectivos (encontrar oportunidades de negócio), como também melhorar a eficiência do sistema na procura da melhor solução.

Uma das fórmulas para melhorar a eficácia na procura da melhor solução é a adopção de um algoritmo que se mostre compatível com o desempenho do sistema usado, ou seja, que

permita a implementação de um programa de execução rápida e fosse, de certo modo, de fácil compreensão.

4.2.1 – Algoritmo ID3

A escolha de um algoritmo que melhor se adapte ao tipo de análise a efectuar não é directa, isto é, não é possível concluir previamente que determinado algoritmo é melhor ou tem melhor desempenho na procura de uma solução do que outro.

Atendendo a que o problema a analisar neste trabalho não era passível de gerar muitas ramificações na árvore de decisão (Figura 4.1), que é um dos condicionalismos que caracterizam o algoritmo ID3 (pode levar a enormes árvores de decisão que não levam necessariamente a melhores bases de regras [hna+199]) optei pela sua utilização pelos seguintes motivos:

- É um algoritmo adequado a problemas do “tipo diagnóstico”;
- Proporciona uma forma automática de desenvolver regras de produção relativamente eficientes;
- Permite uma compreensão do raciocínio seguido com relativa facilidade.

Neste trabalho as regras de análise das séries temporais que foram implementadas tendo em vista a oportunidade de negócio foram as descritas nas secções seguintes:

4.2.1.1 – Séries altamente correlacionadas⁴(valor positivo):

1ª Hipótese:

- A média móvel de 5 da empresa A “andava por cima” da média móvel de 21 da empresa A – significa que existe uma tendência de subida do valor das acções.
- A média móvel de 5 da empresa B “ andava por cima” da média móvel de 21 da empresa B - significa que existe uma tendência de subida do valor das acções.

⁴ O significado exacto de altamente correlacionadas é muito subjectivo pois se tecnicamente um valor superior a 0,5 já significa que existe muita correlação entre as duas séries, no caso das séries de cotações de empresas depende muito do potencial investidor ou, da sua capacidade de investimento, a quantificação deste valor. Na aplicação este valor é parametrizável.

→ Não é verificada a oportunidade de negócio pois ambas têm tendência de subida e não teria muita lógica vender acções quando se sabe previamente que as mesmas estão a valorizar.

2ª Hipótese:

- A média móvel de 5 da empresa A “andava por baixo” da média móvel de 21 da empresa A – significa que existe uma tendência de descida do valor das acções.
- A média móvel de 5 da empresa B “ andava por baixo” da média móvel de 21 da empresa B - significa que existe uma tendência de descida do valor das acções.

→ Não é verificada a oportunidade de negócio pois ambas têm tendência de descida e também não teria muita lógica estar a adquirir acções quando se sabe previamente que as mesmas estão a desvalorizar.

3ª Hipótese:

- A média móvel de 5 da empresa A “andava por baixo” da média móvel de 21 da empresa A – significa que existe uma tendência de descida do valor das acções.
- A média móvel de 5 da empresa B “ andava por cima” da média móvel de 21 da empresa B - significa que existe uma tendência de subida do valor das acções.

→ Neste caso existe uma oportunidade de negócio pois se tenho conhecimento prévio que o valor das acções da empresa A vão subir e, se tenho no meu *portfolio* acções da empresa B, a melhor opção é vender as acções da empresa B e adquirir acções da empresa A, uma vez que ao vender as acções da empresa B já estou a “ganhar dinheiro” .

4ª Hipótese:

- A média móvel de 5 da empresa A “andava por cima” da média móvel de 21 da empresa A – significa que existe uma tendência de subida do valor das acções.
- A média móvel de 5 da empresa B “ andava por baixo” da média móvel de 21 da empresa B - significa que existe uma tendência de descida do valor das acções.

→ Neste caso também se verifica uma oportunidade de negócio pois se tenho conhecimento prévio que o valor das acções da empresa B vão subir e se tenho no meu *portfolio* acções da empresa A, a melhor opção é vender as ac-

ções da empresa A e adquirir acções da empresa B pois a venda das acções da empresa A já foram rentabilizadas.

4.2.1.2– Séries pouco ou nada correlacionadas

Se as séries estão pouco ou nada relacionadas entre si (parti do princípio de que apresentando o factor de correlação um valor inferior a 0,5 ou superior a $-0,5$, se poderia considerar não existir correlação entre as séries) , isso significa que o comportamento entre elas é totalmente aleatório e por isso não é possível fazer nenhuma análise baseada no estudo das respectivas séries temporais.

4.2.1.3– Séries altamente correlacionadas (valor negativo)

No caso de as duas séries se encontrarem altamente correlacionadas mas tendo o factor de correlação um valor negativo ($< -0,5$), podemos concluir que o comportamento das séries é totalmente oposto, isto é, quando uma sobe a outra desce e nessa altura temos as seguintes hipóteses:

1ª e 2ª Hipóteses:

- A média móvel de 5 da empresa A “andava por cima” da média móvel de 21 da empresa A e a média móvel de 5 da empresa B “andava por cima” da média móvel de 21 da empresa B ou, inversamente, as médias móveis de 5 de ambas as empresas “andavam por baixo” das médias móveis de 21 das respectivas empresas:

→ Nada é aconselhado pois se ambas apresentavam uma tendência de subida ou descida - o que era pouco provável já que contrariava o valor indicado pelo factor de correlação, uma eventual oportunidade de negócio só se verificaria se tivesse a certeza de as acções de uma empresa iriam valorizar mais que as da outra, mas essa seria uma previsão baseada no conhecimento pessoal.

3ª Hipótese:

- A média móvel de 5 da empresa A “andava por cima” da média móvel de 21 da empresa A – significa que existe uma tendência de subida do valor das acções.

- A média móvel de 5 da empresa B “andava por baixo” da média móvel de 21 da empresa B - significa que existe uma tendência de descida do valor das acções.

→ Neste caso pode-se verificar uma oportunidade de negócio pois pela análise das tendências sei que as acções da empresa B vão continuar a descer enquanto que as acções da empresa A vão continuar a subir e se tenho no meu *portfolio* acções da empresa B que apresentam uma tendência para descida, a melhor opção é vender as acções da empresa B e adquirir da empresa A.

4ª Hipótese:

- A média móvel de 5 da empresa A “andava por baixo” da média móvel de 21 da empresa A – significa que existe uma tendência de descida do valor das acções.
- A média móvel de 5 da empresa B “ andava por cima” da média móvel de 21 da empresa B - significa que existe uma tendência de subida do valor das acções.

→ Neste caso também se verifica uma oportunidade de negócio pois pela análise das tendências sei que as acções da empresa A vão continuar a descer enquanto que as acções da empresa B vão continuar a subir e se tenho no meu *portfolio* acções da empresa A que apresentam uma tendência para descida, a melhor opção é vender as acções da empresa A e adquirir acções da empresa B.

Esta análise descritiva das várias opções que é possível tomar pode-se resumir nos quadros seguintes:

Factor de Correlação > 0,5				Compra acções da Empresa A e vende as acções da empresa B	Compra acções da Empresa B e vende as acções da empresa A	Aguarda melhor oportunidade de negócio
A tende subida	A N/ tende subida	B tende subida	B N/ tende subida			
X		X				X
X			X		X	
	X		X			X
	X	X		X		

Tabela 4.1 - Análise do factor de correlação (valor positivo)

Factor de Correlação < - 0,5				Compra acções da Empresa A e vende as acções da empresa B	Compra acções da Empresa B e vende as acções da empresa A	Aguarda melhor oportu- nidade de negócio
A tende subida	A tende des- cida	B tende subi- da	B tende des- cida			
X		X				X
X			X	X		
	X		X			X
	X	X			X	

Tabela 4.2 - Análise do factor de correlação (valor negativo)

Face ao apuramento das diversas situações possíveis constantes das duas tabelas anteriores, é agora possível desenhar a árvore de decisão que vai posteriormente facilitar a implementação do algoritmo (representada na Figura 4.1):

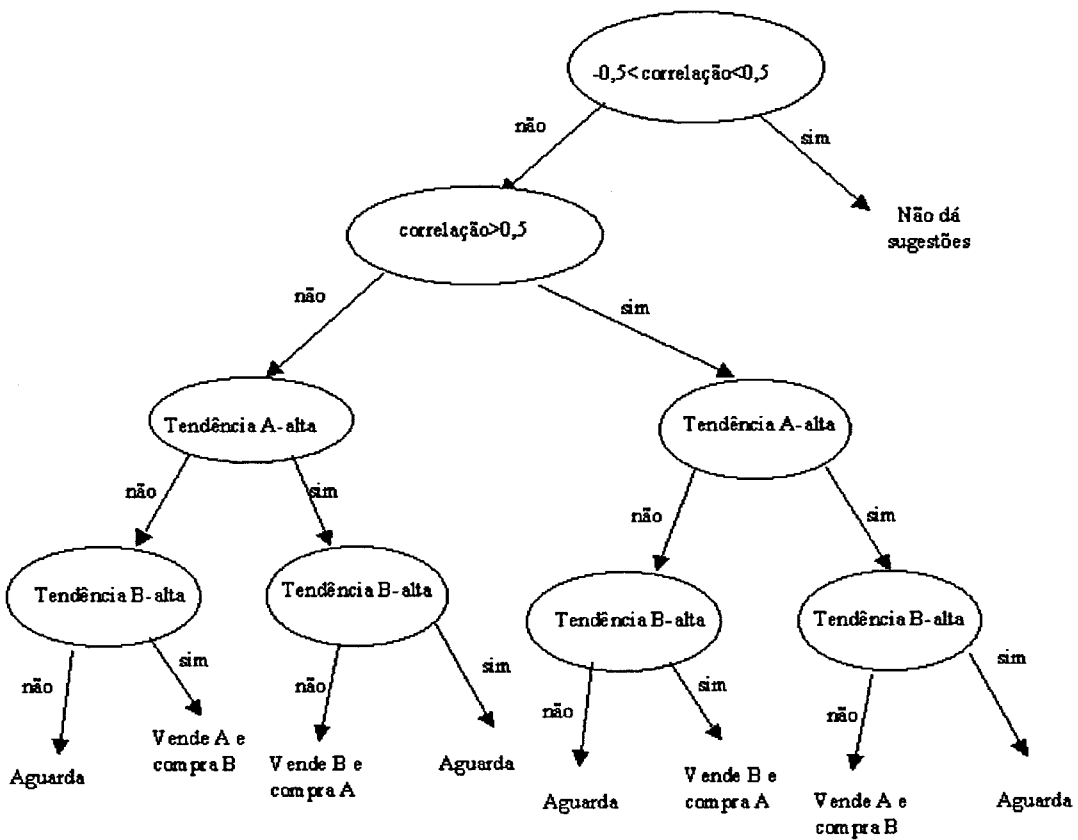


Figura 4.1 - Algoritmo ID3- Árvore de decisão

Na implementação deste algoritmo foram também exploradas as características do KSQL e da KDB já que, obtido o resultado das perguntas efectuadas à base de dados, é possível sem a necessidade de nos socorrermos de outras quaisquer instruções ou funções, proceder à implementação das regras:

```

// se a média de 5 for superior à média de 21
if (((double[])d2[2])[c.n(d2[2]) - m - 1] > ((double[])d2[3])[c.n(d2[3]) - m - 1]){
    media5vec2="Alta";
}else{
    media5vec2="Baixa";
}

double minhacorrelacao=(Double.valueOf((xy).toString()).doubleValue());
if(media5vec1="Alta" && media5vec2="Baixa" && minhacorrelacao>0.5){
    //compra acções da empresa B e vende as acções da empresa A
    comentariol= "Deve comprar as acções da firma\n" +tempB + " e vender as
    acções da firma\n"+tempA;
    rendimento= (((capitalA*((double[])d1[1])[c.n(d1[1]) - m - 1] ) +
    (capitalB * ((double[])d2[1])[c.n(d2[1]) - m - 1])) + capitali);
}
    
```

Programa 4.8 - Verificação do cruzamento das médias móveis

em que:

$((double[])d2[2])[c.n(d2[2])-m-1]$ – tem o valor respeitante à média móvel de 5 da empresa A

$((double[])d2[3])[c.n(d2[3])-m-1]$ – tem o valor respeitante à média móvel de 21 da empresa A

$(Double.valueOf(xy.toString()).doubleValue())$ – tem o valor do coeficiente de correlação

Estes valores foram devolvidos pelo servidor aquando da execução da pergunta e, como se pode verificar pelo exemplo atrás descrito, a manipulação dos dados é conseguida através do acesso directo ao “vector resultado” da pergunta.

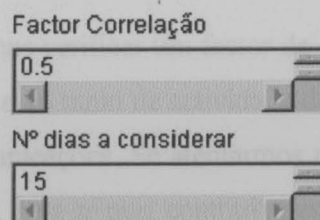
4.3– Interpretação dos Resultados

Para uma correcta análise dos resultados obtidos, como foi referido anteriormente, seria necessário o conhecimento de muitos outros factores que directa ou indirectamente influenciam o comportamento de cada acção na Bolsa de Valores.

Na ausência desses elementos, na aplicação desenvolvida tentei de alguma forma dar uma imagem de credibilidade aos resultados conseguidos apenas pela análise das tendências das séries temporais e do factor de correlação.

Na interface disponibilizada ao utilizador são dadas algumas opções de selecção que, penso eu, facilitam uma análise mais correcta ao utilizador, como sejam:

1. Escolha do factor de correlação que deve ser considerado para efeitos de análise



Factor Correlação

0.5

Nº dias a considerar

15

Figura 4.2 – Parametrização do factor de correlação e nº dias para efeitos de análise

Por restrição foi imposto que não seriam efectuadas quaisquer análises cujo factor de correlação entre as séries estivesse fora do intervalo $-0,5 < \text{Correlação} < 0,5$ já que, a análise de séries temporais em que o factor de correlação se situe fora deste intervalo pode conduzir a erro e distorção dos resultados.

No entanto, se o utilizador assim o entender pode verificar o comportamento das séries através do gráfico.

2. Número de dias anteriores que a aplicação deverá considerar para efeitos de simulação, ou seja, tendo por base os “palpites” resultantes da análise, a aplicação vai comprar e vender acções de acordo com a tendência das séries. O objectivo é verificar a exactidão dos “palpites” no decurso de determinado período.
3. Selecção de empresas por ramos de actividade

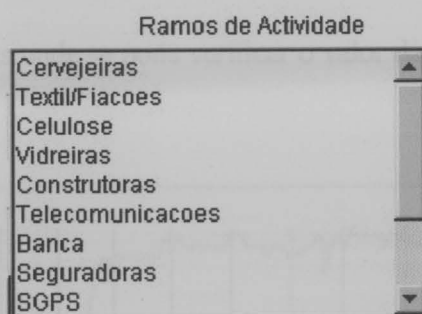


Figura 4.3 - Selecção do ramo de actividade

A opção pela inserção de todas as empresas cotadas na Bolsa de Valores de Lisboa num ramo de actividade específico visa não só uma mais fácil identificação das próprias empresas como também uma análise comparativa entre empresas inseridas na mesma área de actividade uma vez que, conforme já foi descrito no capítulo 2, existe uma maior correlação entre empresas inseridas no mesmo ramo diminuindo o risco deste tipo de análise.

Exemplificando, podemos verificar um factor de correlação elevado ($>0,7$) entre duas empresas, Empresa A pertencente ao ramo de actividade da Banca e Empresa B pertencente ao ramo de actividade de Telecomunicações. Se atentarmos na evolução que se tem verificado a nível da Bolsa nestes dois ramos, nota-se que, contrariamente ao ramo da Banca que é um ramo relativamente estável, o ramo das Telecomunicações tem grandes oscilações (no anexo 1 podem ser vistos exemplos).

Um investimento baseado na comparação entre estas duas empresas poderá, ou não, ser seguro, dependendo mais do conhecimento exterior das próprias empresas e da intuição pessoal (que por vezes é a melhor análise !).

4. Selecção de empresas arbitrariamente:

Investimento inicial em: 28-05-2000

Empresa A- Empresa B-

-1000 acções

Total: 40590.0

Figura 4.4 - Selecção específica de duas empresas

Como já foi referido é possível, mesmo que o factor de correlação esteja fora dos limites atrás descritos, verificar o comportamento das séries de duas empresas seleccionadas de forma arbitrária. No caso do factor de correlação entre as empresas seleccionadas ser superior a 0,5 ou inferior a -0,5, nesse caso, apesar de poderem não pertencer ao mesmo ramo de actividade é efectuada a análise das séries.

5. Gráfico de linhas onde se pode verificar o valor das acções, e os meses correspondentes.

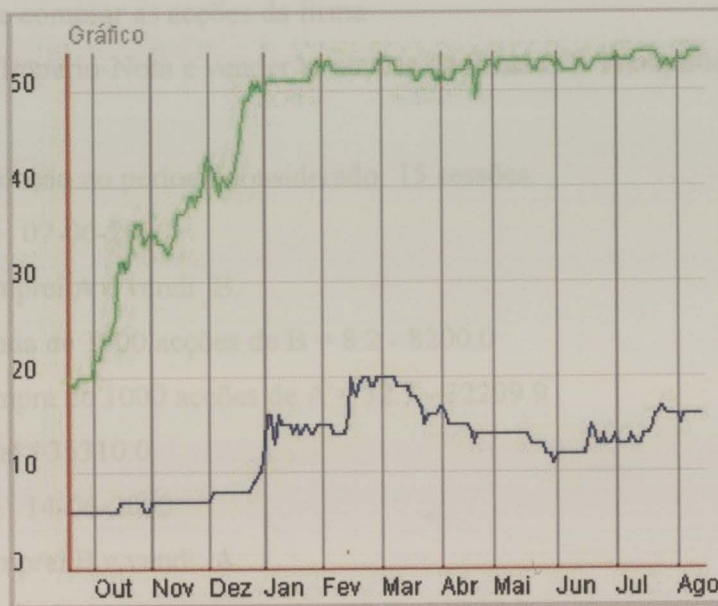


Figura 4.5 - Gráfico linear da evolução das séries temporais

6. Nas situações em que a aplicação procede à análise, é facultado ao utilizador um “resumo” da simulação efectuada e é dado o “palpite” sobre a evolução futura das acções:

Investimento inicial em : 28-05-2000

- 1000 acções de A a 32.42 - 32420.0

- 1000 acções de B a 8.17 - 8170.0

Total : 40590.0

Capital em Reserva: 59410.0

O Factor de correlação entre as Séries é: 0.891

A tendência de mercado para a empresa

C S Tranquilidade-Nom é de Alta

A tendência de mercado para a empresa

C S Imperio-Nom é de Baixa

A minha sugestão é:

Deve comprar as acções da firma

C S Imperio-Nom e vender as acções da firma C S Tranquilidade-Nom

Evolução no período considerado: 15 sessões:

Em : 07-06-2000

Comprei A e vendi B.

Venda de 1000 acções de B = 8.2 - 8200.0

Compra de 1000 acções de A = 32.3 - 32299.9

Total =35310.0

Em : 14-06-2000

Comprei B e vendi A.

Venda de 1000 acções de A = 32.7 - 32700.0

Compra de 1000 acções de B = 7.95 - 7950.0

Total =60060.0

Capital total final

1000.0 accções de A a 32.7

1000.0 acções de B a 7.65

total: 100410.0

Em : 16-06-2000

4.4– Resumo

Neste capítulo foi descrito o trabalho efectuado relativamente à análise das séries temporais tendo sido dado especial ênfase à vertente de bases de dados.

A ferramenta utilizada (KDB) e as suas características foram neste capítulo descritas com algum pormenor através de exemplos retirados do trabalho subjacente a esta tese.

Neste capítulo merece também especial destaque a descrição de como pode ser efectuada a manipulação dos dados a partir de uma simples pergunta à base de dados. Para isso muito contribui a linguagem KSQL essencialmente implementada para efeitos de análise (sistemas OLAP).

Capítulo 5

Avaliação da Tecnologia Utilizada

Introdução

Neste capítulo pretende-se, ainda que de uma forma não exaustiva, realçar as vantagens na utilização de um sistema de base de dados que reuna as características de um sistema de base de dados tradicional com os sistemas OLAP e OLTP e, simultaneamente analisar o desempenho da KDB.

5.1 – K/KDB

Nos actuais sistemas de bases de dados comerciais como por exemplo *Oracle*, *Microsoft*, *Informix*, etc., é possível definir novos tipos de dados e assim realizar perguntas à base de dados de forma a atingir os mesmos objectivos que conseguimos com o KSQL.

Exemplos disto são o *Oracle Express da Oracle* [ORA00], o *OLAP Services do Microsoft SQL Server*, ou a *Essbase* [AEO98]— uma base de dados OLAP, etc.

No entanto, existem alguns pormenores que diferenciam a KDB dos outros sistemas de bases de dados:

- A simplicidade colocada na sua implementação – apenas 170 Kb;
- Escalabilidade – apenas limitada pela capacidade da máquina (memória RAM¹);

¹ A memória RAM deve ser pelo menos quatro vezes superior ao tamanho (em memória) da maior coluna da tabela

- A tecnologia utilizada – o *arrable* e a possibilidade de implementação de funções em linguagem K para a resolução deste tipo de problemas (análise de séries temporais);
- Forma como são tratadas as transacções – sem locks nem deadlocks.

5.1.1 – K (Desempenho)

Uma das grandes virtudes desta ferramenta de bases de dados é, sem dúvida, "o espaço" que ocupa (todo o código fonte foi escrito em apenas 300.000 bytes) incluindo:

- Interface;
- Acesso às funções implementadas em linguagem K;
- Servidor.

Desenvolvida em linguagem C, é uma linguagem extremamente rápida permitindo por exemplo "carregar" em memória 3 milhões de registos em apenas 2 segundos (numa máquina IBM 990) [ADS99].

O facto de ao iniciar o sistema ser carregada em memória toda a base de dados leva a que o acesso e manipulação dos dados se efectue de uma forma muito mais rápida do que se tivesse que aceder a disco sempre que fosse necessário efectuar quaisquer transacções.

Além das funções de servidor de base de dados, o K suporta também o protocolo TCP-IP, o que lhe permite dispensar outras linguagens de programação para efeitos de acesso aos dados, reduzindo assim o tempo gasto com conversões de protocolo e comunicações.

Como exemplo pode-se referir alguns valores gastos pelo servidor para devolver os dados a algumas perguntas feitas em KSQL [KXS99]:

<code>select sum x from t</code>	- 30 MRPS ²
<code>select sum x by f from t</code>	- 6 MRPS
<code>select sum x by f, g from t</code>	- 1.5 MRPS

² MRPS milhões de registos por segundo. Estes valores foram conseguidos numa máquina com um processador a 450MHZ.

Comparativamente com outros sistemas de bases de dados, os valores alcançados são significativamente inferiores. No caso concreto da Sybase, os valores publicados no Dr.Dobb's Jornal jan/98 página 82 em que é feita uma comparação entre kdb/jdbc e Sybase/jdbc que utilizam ambas drives do tipo IV, são os seguintes³:

	Kdb	Sybase
jdbc-compliant	Sim	Não
driver size	20KB	700KB
download(56.6)	5 segundos	200 segundos
Connect	.1 segundos	1.1 segundos
select 10000	.6 segundos	100 segundos
procedure call	.004 segundos	.05 segundos
update/insert	.018 segundos	.05 segundos

Tabela 5.1- Comparação entre KDB e Sybase

Se compararmos apenas a performance da KDB usando o driver JDBC ou o driver KDBC verifica-se que num ciclo de leitura de 800 000 registos, com JDBC demora o dobro do tempo do que KDBC – 30 000 registos/segundo contra 15 000 registos/segundo.

5.2– Bases de Dados – Factores que influenciam negativamente o desempenho dos SGBD

Os grandes problemas com que se depara qualquer sistema de base de dados em termos de eficiência são essencialmente os seguintes [KXS00]:

³ (Nota: Os tempos respeitantes a kdb foram obtidos num processador a 450MHZ. Os tempos respeitantes a Sybase foram obtidos num processador a 200MHZ.)

- Controlo da concorrência das transacções
- Capacidade de análise e processamento – sistemas OLAP
- Volume de informação armazenada e a analisar que, no caso de bases de dados distribuídas implicam também eficiência de resposta por parte da rede.

Relativamente ao primeiro ponto – controlo das transacções, independentemente da forma utilizada em cada sistema de bases de dados para efectuar esse controlo, a necessidade de manter a consistência e integridade da base de dados comprometem de facto a rapidez e eficiência do sistema.

É neste contexto que a seguir é descrita a forma como são geridas as transacções da KDB.

5.3– Transacções em KDB

Tratar cada transacção de forma independente envolve custos (em termos de desempenho) já que torna inevitavelmente o sistema mais lento.

A *Kx Systems* optou por tratar as transacções de forma agrupada, isto é, quando é feita a escrita em disco das modificações das transacções, a operação envolve todas as transacções que estão pendentes até aquele momento, sendo as mesmas caracterizadas pelo tipo de transacção e pelos parâmetros correspondentes.

Uma vez efectuada a escrita em disco das modificações das transacções, é sempre possível a recuperação do estado inicial da base de dados através da simples leitura do ficheiro de “log”, considerando-se assim que as transacções são feitas de forma segura.

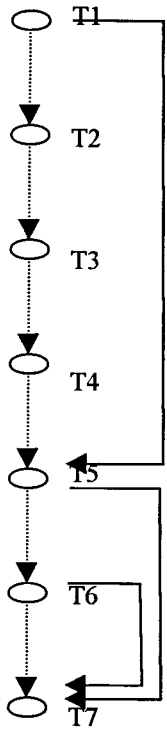
Para conseguir que as transacções sejam efectuadas em série, a *Kx Systems* socorreu-se de um algoritmo - OBEYORDER, que faz uso de um predicado já implementado (CONFLIT), que mais não faz que não seja determinar quando é que pode ocorrer conflito entre duas transacções que estão a ser efectuadas em série.

Este algoritmo permite explorar uma das características do K – *Multithreading* e assim paralelizar algumas transacções sem ter que se socorrer de bloqueios para controlo de concorrência, como veremos mais a seguir.

O modo de funcionamento do algoritmo *OBEYORDER* consiste no seguinte:

1. Conforme as transacções vão chegando, vão sendo registadas num ficheiro - *log* em "*batches*";
2. É construído um grafo cujos nós são as transacções que se encontram registadas no *log*. Se entre algumas transacções existem conflitos, então deve-se fazer uma ligação entre esses nós. Exemplificando, se a transacção *t* entra em conflito com a transacção *t'* e *t* está sequencialmente à frente de *t'* no ficheiro *batch* (ficheiro que nos dá a ordem porque deverão ser efectuadas as transacções), então é feita uma ligação directa entre *t* e *t'*.
3. Isto leva a que seja efectuado um novo grafo $G = (T, E)$ em que *T* representa a totalidade das transacções a efectuar e *E* representa as transacções que têm conflitos.
4. Executa *T* em paralelo, respeitando contudo a ordem derivada das transacções representadas por *E*.

A Figura 5.2 mostra como funciona este algoritmo:



As linhas a tracejado representam as transacções de acordo com a ordem de chegada.
As linhas ligadas representam as transacções que estão em conflito.

Em primeiro lugar tratam-se as transacções que estão em conflito, isto é, as transacções respeitantes aos nós (T2,T5), (T5,T7) e (T6,T7).

Depois são efectuadas as transacções de acordo com a ordem dos nós – T1,T2,T3,T4,T6, as quais podem ser executadas em paralelo. Depois de executadas estas transacções, apenas permanecem as transacções respeitantes aos nós T5 e T7.

É então executada a transacção respeitante ao nó T5 e só depois a transacção respeitante ao nó T7.

Figura 5.1- Algoritmo OBEYORDER

Como se pode verificar pela figura anterior, as transacções são assim colocadas em série, ou seja, de acordo com a ordem de chegada (T1->T2->T3->T4->T5->T6->T7) sendo posteriormente executadas em paralelo de acordo com a ordem dos nós, isto é, primeiro as transacções que são passíveis de gerar qualquer conflito e depois aquelas que não estão em conflito.

5.4- Conclusão

Seria presunção pensar que as potencialidades atribuídas à KDB são a melhor opção quer para o acesso quer para a manipulação de dados. Apesar da rapidez de execução, também tem as suas limitações como penso que ficou demonstrado ao nível das transacções.

No que concerne às transacções, o modelo adoptado não se pode dizer que seja original – tratar as mesmas de uma forma sequencial de acordo com a ordem de chegada e processá-las em paralelo a fim de evitar bloqueios.

Relativamente ao controlo de concorrência, inserindo-se a KDB no tipo de ferramentas vocacionadas para a manipulação rápida de grandes quantidades de dados, é de prever que, e

atendendo que a informação armazenada nas *Datawarehouses* é de certo modo estática, os acessos às mesmas serão feitos apenas para leitura, razão pela qual não deve haver preocupações com o controlo de concorrência [ART00].

Relativamente às comparações efectuadas, é preciso ter em consideração que, e não percebo porquê, os processadores utilizados para efeitos de medida de desempenho não eram iguais, havendo no entanto factores para os quais a velocidade do processador não é significativa- por exemplo descarregar o conteúdo de uma página ou um ficheiro.

Os critérios seguidos pela *Kx Systems* para a implementação deste sistema apesar de não serem novidade já que muita da tecnologia usada foi baseada em conceitos já desenvolvidos em *APL*, permitem de alguma forma pensar neste sistema como uma solução para certo tipo de problemas e especificamente as séries temporais.

Outros sistemas assentes na mesma linguagem, como por exemplo a *IQL (Interactiv Query Language)* [IQL00], uma ferramenta implementada em *Dyalog/APL*, permitem também a conciliação entre os tradicionais sistemas de bases de dados e as funcionalidades OLAP.

Capítulo 6

Obtenção de Resultados

Introdução

Neste capítulo são abordadas as dificuldades com que me deparei na realização deste trabalho, nomeadamente a obtenção de dados, os aspectos que não foram implementados e os resultados alcançados.

6.1 – Dificuldade de obtenção de dados

Neste trabalho a recolha de informação, mais especificamente dos valores das cotações diárias das acções, foi efectuada apenas no sítio da Bolsa de Valores de Lisboa.

Como já anteriormente foi referenciado, e porque a actualização das páginas é efectuada uma vez por dia, os dados que foi possível recolher apenas correspondem aos valores diários.

Este foi, em minha opinião, um pormenor que limitou de forma muito significativa a análise efectuada, pelos seguintes motivos:

- No que concerne à análise das séries e tendo em consideração o tipo de análise em que me baseei, poderia obter resultados substancialmente diferentes pois durante o dia ocorrem várias oscilações dos valores das cotações, isto é, pode o preço de fecho mostrar uma tendência de alta e durante o dia terem ocorrido variações que contrariavam esta tendência. Estas variações afectariam os valores das médias móveis de 5 e 21 e consequentemente a opção de compra ou venda de determinada acção, durante a sessão bolsista.

- Em relação ao desempenho da base de dados, em 14 de Setembro de 2000 apenas constavam da base de dados 30 400 registos que correspondiam às cotações compreendidas entre 14 de Setembro de 1999 e 14 de Setembro do corrente ano. Em termos de performance, embora tenha apresentado alguns valores comparativos entre o K e outros sistemas de bases de dados, os dados de que dispunha eram manifestamente insuficientes para poder proceder à análise em termos de eficiência e poder com base nos mesmos extrair conclusões, sob pena de estar de alguma forma a reproduzir dados errados.

Apesar de haver alguns sítios onde este tipo de informação é supostamente disponibilizada em tempo real (existe sempre algum atraso), a forma como é disponibilizada a informação varia de sítio para sítio o que implicava o desenvolvimento da aplicação efectuada para realizar diariamente o download dos valores das acções e, consequentemente a alteração de toda a calendarização prevista para a realização deste trabalho.

6.2– Resultados obtidos

Esta é talvez a parte deste relatório que mais curiosidade poderá despertar pois retrata com fidelidade os resultados obtidos com a aplicação dos métodos atrás descritos (factor de correlação, médias móveis, algoritmo de pesquisa) às séries de acções da BVL.

Para uma melhor compreensão dos passos dados pela aplicação para a análise diária dos preços das cotações, optei por incluir quatro exemplos que traduzem também algumas das possíveis situações de análise, comparando os valores reais com o pseudo investimento efectuado na simulação:

1. Séries altamente correlacionadas positivamente – teste efectuado entre empresas inseridas dentro do mesmo ramo de actividade e cujo resultado final se traduziu numa boa oportunidade de negócio;
2. Séries altamente correlacionadas positivamente – teste efectuado entre empresas inseridas dentro do mesmo ramo de actividade e cujo resultado final se traduziu numa má oportunidade de negócio;

3. Séries altamente correlacionadas negativamente – teste efectuado entre empresas inseridas dentro do mesmo ramo de actividade e cujo resultado final se traduziu numa boa oportunidade negócio;
4. Séries altamente correlacionadas positivamente – teste efectuado entre empresas inseridas dentro do mesmo ramo de actividade mas de ramo de actividade diferente do analisado nos pontos anteriores e cujo resultado final se traduziu numa boa oportunidade negócio;

De notar que, nos exemplos apresentados, o factor de correlação indicado é o respeitante aos valores de toda a série e não apenas aos valores constantes das tabelas correspondentes a cada exemplo em que apenas são referidos os valores respeitantes aos períodos de tempo constantes das mesmas e, pelos valores das tabelas o factor de correlação seria muito diferente (para o exemplo 1 seria de $-0,37494$, o que estava totalmente errado).

De seguida são apresentados os resultados da simulação efectuada para cada um destes quatro casos atrás identificados.

6.2.1 – Exemplo de simulação 1

Empresas analisadas: Centralcer (Empresa A) e Sumolis (Empresa B)

Factor de Correlação: 0.8108

Período considerado: De 21-07-2000 a 11-08-2000

Nas duas tabelas seguintes podemos verificar o valor real do preço de fecho das acções das empresas:

Data do Fecho	Centralcer (A) Preço (Euros)	Sumolis (B) Preço (Euros)
21-Jul-00	54,00	16,40
24-Jul-00	53,00	16,06
25-Jul-00	52,40	17,88
26-Jul-00	53,00	17,65
27-Jul-00	53,00	16,50
28-Jul-00	52,95	16,94
31-Jul-00	53,00	16,65
01-Ago-00	53,00	16,73
02-Ago-00	53,00	16,39
03-Ago-00	53,00	15,95
04-Ago-00	53,00	16,24
07-Ago-00	54,45	16,60
08-Ago-00	54,45	16,50
09-Ago-00	54,45	16,27
10-Ago-00	54,45	16,35
11-Ago-00	54,45	16,46

Tabela 6.1- Preços das acções das empresas Centralcer e Sumolis

Para uma melhor visualização da evolução ao longo do período considerado (21 de Julho de 2000 a 11 de Agosto de 2000) do valor das acções, nos gráficos da Figura 6.1 pode-se verificar as respectivas variações.

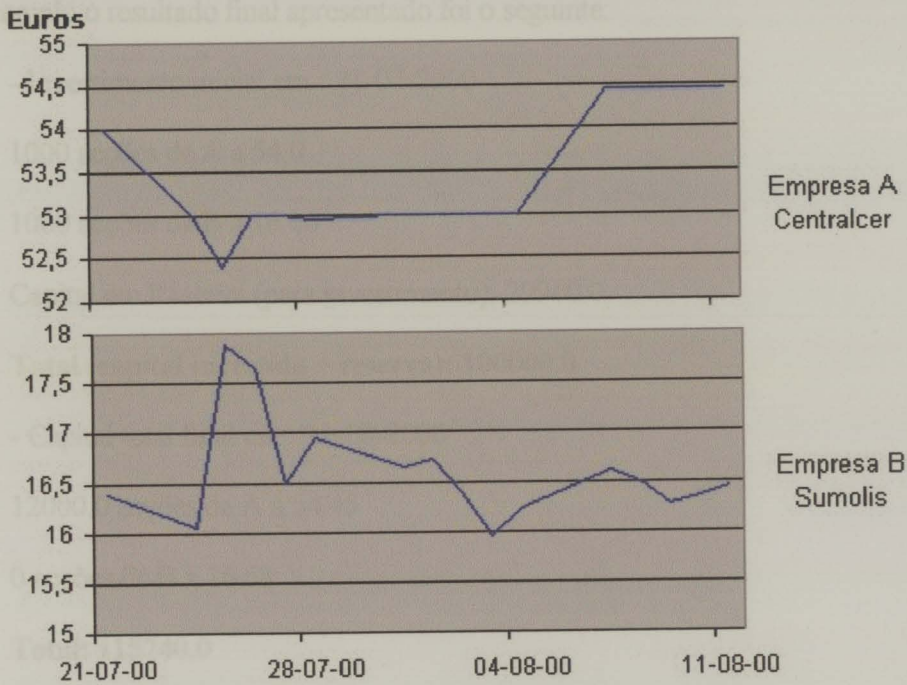


Figura 6.1- Gráficos das empresas Centralcer e sumolis

A análise e transacções efectuadas pela aplicação foram os seguintes:

Movimentos

Em : 25-07-2000
 Comprei A e vendi B.
 Em : 26-07-2000
 Comprei A.
 Em : 27-07-2000
 Comprei A.
 Em : 28-07-2000
 Comprei A.
 Em : 31-07-2000
 Comprei A.
 Em : 01-08-2000
 Comprei A.
 Em : 02-08-2000
 Comprei A.
 Em : 03-08-2000
 Comprei A.
 Em : 04-08-2000
 Comprei A.
 Em : 07-08-2000
 Comprei A.
 Em : 08-08-2000
 Comprei A.

Comentário

Como se pode verificar no gráfico, no dia 25 verificou-se uma subida acentuada do valor das acções da Sumolis.

Seguindo o critério descrito anteriormente, era previsível que o valor das acções da Centralcer subisse também, de forma que vendeu a totalidade das acções que detinha da Sumolis e começou a investir na Centralcer.

A partir do dia 8 o valor das acções da Centralcer estabilizou e, não se verificando nenhuma tendência, não é dada qualquer opção de compra.

Neste exemplo o resultado final apresentado foi o seguinte:

- Investimento inicial em : 21-07-2000

1000 acções de A a 54.0

1000 acções de B a 16.06

Capital em Reserva (para investimento): 29940.0

Total (capital investido + reserva): 100000.0

- Capital total final em : 11-08-2000

12000.0 acções de A a 54.45

0 acções de B a 16.68

Total: 115740.0

Valor que teria sem a aplicação do método: $(1000 \times 54.45 + 1000 \times 16.68 + 29940$
(Capital em reserva)) = **101070**

6.2.2– Exemplo de simulação 2

Empresas analisadas: Portugal Telecom (Empresa A) e Telecel (Empresa B)

Factor de Correlação: 0.871

Período considerado: De 30-06-2000 a 11-08-2000

Nas duas tabelas seguintes podemos verificar o valor real do preço de fecho das acções das empresas:

Data do Fecho	Portugal Telecom (A) Preço (Euros)	Telecel (B) Preço (Euros)
30-Jun-00	11,66	15,50
01-Jul-00	13,10	16,96
03-Jul-00	11,76	15,90
04-Jul-00	11,90	16,25
05-Jul-00	11,90	15,95
06-Jul-00	12,03	15,99
07-Jul-00	11,95	15,81
10-Jul-00	12,33	16,00
11-Jul-00	12,42	15,90
12-Jul-00	12,70	15,95
13-Jul-00	12,85	16,10
14-Jul-00	12,97	16,23
18-Jul-00	13,50	16,76
19-Jul-00	13,25	16,60
20-Jul-00	13,10	16,40
21-Jul-00	13,63	16,83
24-Jul-00	13,31	16,41
25-Jul-00	13,80	16,25
26-Jul-00	13,12	16,08
27-Jul-00	13,00	15,80
28-Jul-00	12,99	15,96
31-Jul-00	11,75	15,90
01-Ago-00	11,75	16,00
02-Ago-00	12,15	16,02
03-Ago-00	12,35	16,19
04-Ago-00	11,81	16,03
07-Ago-00	11,90	16,12
08-Ago-00	11,65	15,90
09-Ago-00	11,38	15,98
10-Ago-00	11,89	15,81
11-Ago-00	12,07	15,75

Tabela 6.2- Preço das acções das empresas Portugal Telecom e Telecel

Para uma melhor visualização da evolução ao longo do período considerado (30 de Junho de 2000 a 11 de Agosto de 2000) do valor das acções, nos gráficos da Figura 6.2 pode-se verificar as respectivas variações.

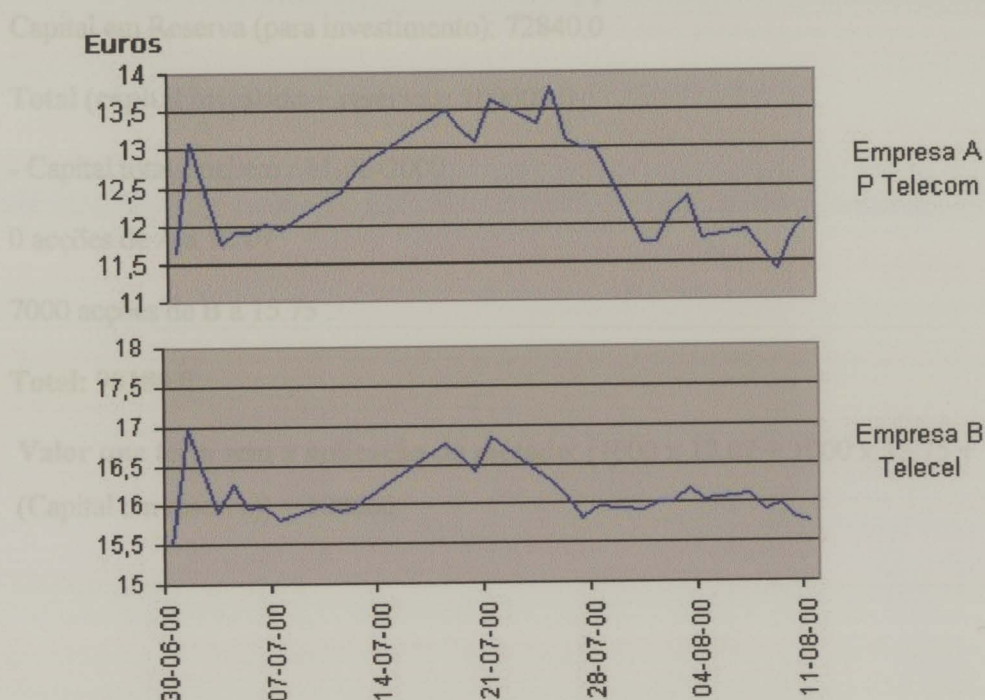


Figura 6.2- Gráficos das empresas Portugal Telecom e Telecel

A análise e transacções efectuadas pela aplicação foram os seguintes:

Movimentos

Em : 01-07-2000
 Comprei B e vendi A.
 Em : 13-07-2000
 Comprei B.
 Em : 14-07-2000
 Comprei B.
 Em : 18-07-2000
 Comprei B.
 Em : 28-07-2000
 Comprei B.
 Em : 01-08-2000
 Comprei B.

Comentário

Como se pode verificar, apesar do período considerado ser relativamente longo, apenas foram efectuadas seis transacções.

Se verificarmos no gráfico, no dia 1 verificou-se uma subida acentuada do valor das acções de ambas as empresas.

A opção continuada de compra de acções da Telecel foi correcta pois era previsível que o valor das acções subisse e, se atentarmos nos valores diários descritos na tabela 6.2 verifica-se essa subida.

A venda das acções da Portugal Telecom no dia 1 também foi uma boa opção pois tinha comprado no dia anterior mais barato.

Neste exemplo o resultado final apresentado foi o seguinte:

- Investimento inicial em : 30-06-2000

1000 acções de A a 11.66

1000 acções de B a 15.5

Capital em Reserva (para investimento): 72840.0

Total (capital investido + reserva): 100000.0

- Capital total final em : 11-08-2000

0 acções de A a 12.07

7000 acções de B a 15.75

Total: 98180.0

Valor que teria sem a aplicação do método: $(1000 \times 12.07 + 1000 \times 15.75 + 72840$
(Capital em reserva)) = 100660

6.2.3– Exemplo de simulação 3

Empresas analisadas: Banco Comercial Português (Empresa A) e Banco Totta & Açores (Empresa B)

Factor de Correlação: -0.804

Período considerado: De 21-07-2000 a 11-08-2000

Nas duas tabelas seguintes podemos verificar o valor real do preço de fecho das acções das empresas:

Data do Fecho	BCP (A) Preço (Euros)	BTB (B) Preço (Euros)
21-Jul-00	54,00	27,00
24-Jul-00	53,00	26,40
25-Jul-00	52,40	26,35
26-Jul-00	53,00	26,01
27-Jul-00	53,00	26,00
28-Jul-00	52,95	25,10
31-Jul-00	53,00	25,10
01-Ago-00	53,00	26,29
02-Ago-00	53,00	26,27
03-Ago-00	53,00	25,99
04-Ago-00	53,00	25,20
07-Ago-00	54,45	25,20
08-Ago-00	54,45	24,64
09-Ago-00	54,45	25,00
10-Ago-00	54,45	24,71
11-Ago-00	54,45	24,60

Tabela 6.3- Preço das acções das empresas BCP e BTB

Para uma melhor visualização da evolução ao longo do período considerado – 21 de Julho de 2000 a 11 de Agosto de 2000 - do valor das acções, nos gráficos da Figura 6.3 pode-se verificar as respectivas variações.

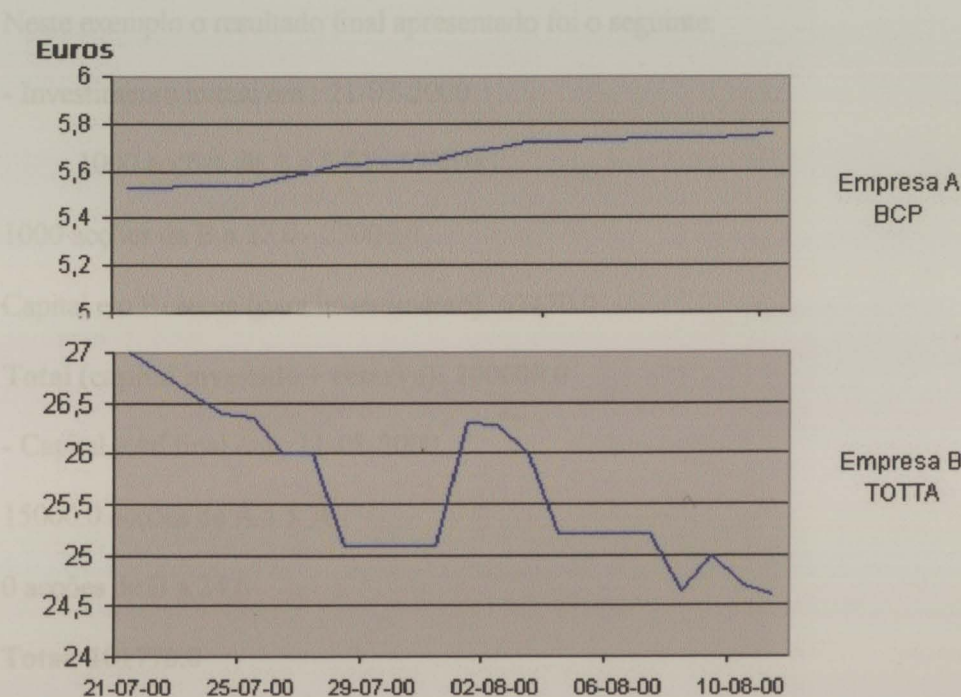


Figura 6.3- Gráficos das empresas BCP e BTB

A análise e transacções efectuadas pela aplicação foram os seguintes:

Movimentos	Comentário
Em : 21-07-2000 Comprei A e vendi B.	
Em : 24-07-2000 Comprei A .	Como se pode verificar no gráfico, o valor
Em : 25-07-2000 Comprei A .	das acções do Banco Comercial Português
Em : 26-07-2000 Comprei A .	sofreram uma evolução constante ao longo
Em : 27-07-2000 Comprei A .	do período considerado.
Em : 28-07-2000 Comprei A .	O valor das acções do Banco Totta & Açores,
Em : 31-07-2000 Comprei A .	pelo contrário, mostra claramente uma
Em : 01-08-2000 Comprei A .	tendência de descida.
Em : 02-08-2000 Comprei A .	A opção de compra de acções do Banco
Em : 03-08-2000 Comprei A .	Comercial Português mostrou-se assim uma
Em : 04-08-2000 Comprei A .	boa opção, sendo que no fim é dado o conselho
Em : 07-08-2000 Comprei A .	para a aquisição de acções deste banco.
Em : 08-08-2000 Comprei A .	
Em : 09-08-2000 Comprei A .	

Neste exemplo o resultado final apresentado foi o seguinte:

- Investimento inicial em : 21-07-2000

1000 acções de A a 5.53 - 5530.0

1000 acções de B a 27.0 - 27000.0

Capital em Reserva (para investimento): 67470.0

Total (capital investido + reserva): 100000.0

- Capital total final em : 11-08-2000

15000.0 acções de A a 5.76

0 acções de B a 24.6

Total: 101770.0

Valor que teria sem a aplicação do método: $(1000 \times 5.76 + 1000 \times 24.6 + 67470$
(Capital em reserva)) = 97830

6.2.4– Exemplo de simulação 4

Empresas analisadas: Mota & Companhia (Empresa A) e Teixeira Duarte Eng Construcões (Empresa B)

Factor de Correlação: 0.817

Período considerado: De 30-06-2000 a 11-08-2000

Nas duas tabelas seguintes podemos verificar o valor real do preço de fecho das acções das empresas:

Data do Fecho	Mota & C ^a (A) Preço (Euros)	Teixeira Duarte (B) Preço (Euros)
30-Jun-00	12,75	18,65
01-Jul-00	13,50	18,00
03-Jul-00	13,00	18,60
04-Jul-00	13,50	18,70
05-Jul-00	13,60	18,67
06-Jul-00	13,45	18,59
07-Jul-00	13,32	18,58
10-Jul-00	13,50	18,53
11-Jul-00	13,10	18,15
12-Jul-00	13,40	18,00
13-Jul-00	13,40	18,20
14-Jul-00	13,50	18,19
18-Jul-00	13,47	18,20
19-Jul-00	13,19	18,20
20-Jul-00	13,14	18,00
21-Jul-00	12,91	18,20
24-Jul-00	12,91	18,20
25-Jul-00	12,80	18,19
26-Jul-00	12,98	18,10
27-Jul-00	12,90	18,00
28-Jul-00	12,90	18,05
31-Jul-00	13,04	18,05
01-Ago-00	12,99	18,50
02-Ago-00	12,99	18,50
03-Ago-00	12,97	18,54
04-Ago-00	12,89	18,51
07-Ago-00	12,90	18,30
08-Ago-00	12,60	18,46
09-Ago-00	12,98	18,50
10-Ago-00	12,97	18,50
11-Ago-00	12,92	18,40

Tabela 6.4- Preço das acções das empresas Mota & C^a e Teixeira Duarte

Para uma melhor visualização da evolução ao longo do período considerado (30 de Junho de 2000 a 11 de Agosto de 2000) do valor das acções, nos gráficos da Figura 6.4 pode-se verificar as respectivas variações.

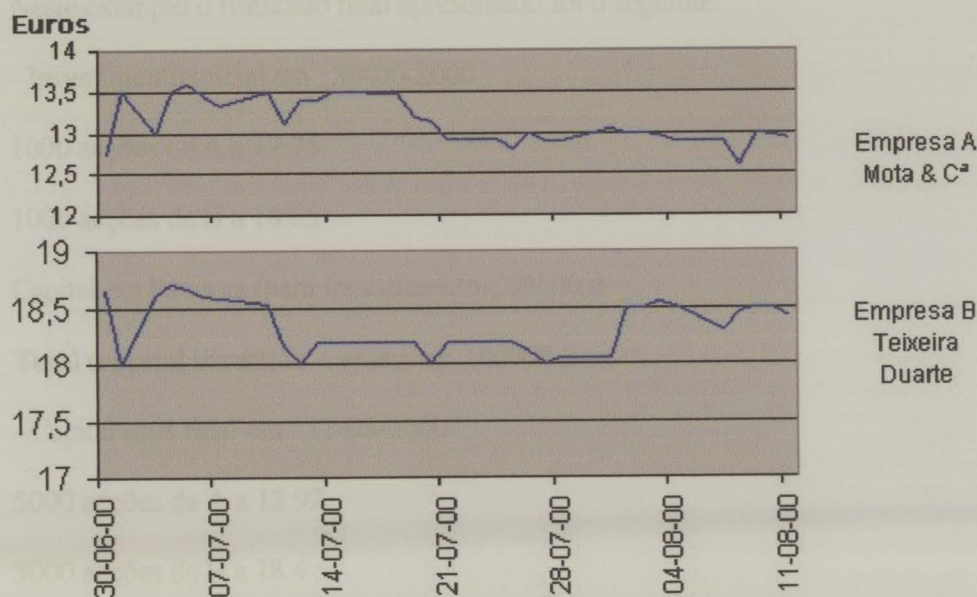


Figura 6.4- Gráfico das empresas Mota & Cª e Teixeira Duarte

A análise e transacções efectuadas pela aplicação foram os seguintes:

Movimentos	Comentário
Em : 01-07-2000 Comprei B e vendi A.	Como se pode verificar, durante o mês de Julho a opção de compra foi sempre em acções da empresa B – Teixeira Duarte – e, se verificarmos no gráfico, podemos constatar que houve de facto um crescimento do valor das acções desta empresa, embora o mesmo só se tivesse verificado mais tarde – no dia 1 de Agosto.
Em : 13-07-2000 Comprei B.	
Em : 14-07-2000 Comprei B.	Assim a opção de venda das acções da empresa B em 3 de Agosto revelou-se uma boa opção pois vendeu a preço superior ao da aquisição.
Em : 18-07-2000 Comprei B.	
Em : 19-07-2000 Comprei B.	Analogamente a opção de aquisição das acções da empresa A a partir de Agosto também se mostrou uma boa opção, que aliás é recomendada pela aplicação.
Em : 20-07-2000 Comprei B.	
Em : 21-07-2000 Comprei B.	
Em : 03-08-2000 Comprei A e vendi B.	
Em : 04-08-2000 Comprei A e vendi B.	
Em : 07-08-2000 Comprei A e vendi B.	
Em : 08-08-2000	

Comprei A e vendi B.
Em : 09-08-2000
Comprei A e vendi B.

Se compararmos os resultados entre os investimentos efectuados e o valor que teríamos sem a aplicação, o resultado final é favorável.

Neste exemplo o resultado final apresentado foi o seguinte:

- Investimento inicial em : 30-06-2000

1000 acções de A a 12.75

1000 acções de B a 18.65

Capital em Reserva (para investimento): 68600.0

Total (capital investido + reserva): 100000.0

- Capital total final em : 11-08-2000

5000 acções de A a 12.92

3000 acções de B a 18.4

Total: 102880.0

Valor que teria sem a aplicação do método: $(1000 \times 12.92 + 1000 \times 18.4 + 68600$
(Capital em reserva)) = 99920

6.3- Conclusões

Não é fácil extrair uma conclusão definitiva sobre a adequabilidade do método usado baseada nos quatro exemplos descritos nos pontos anteriores. Há no entanto alguns aspectos que considero importante salientar:

Para todas as simulações efectuadas foram definidas as seguintes condições:

1. Inicialmente eram adquiridas 1000 acções das empresas objecto de análise;
2. Não havia restrições para a compra, ou seja, se aparecesse uma oportunidade de negócio procedia à aquisição;
3. Não poderia comprar ou vender mais que 1 000 acções por cada transacção.

Como é óbvio estas restrições são impostas de acordo com “a natureza” e situação económica do investidor, isto é:

- se é calculista, que é talvez o caso descrito, em que o risco inerente ao investimento efectuado é de certa forma calculado ;
- se é pessimista, e aí possivelmente não seria um investidor em acções mas talvez em fundos de investimento ou outros produtos do género;
- ou então optimista e possivelmente venderia todas as acções que detivesse e investiria todo o dinheiro em cada transacção.

A alteração destes critérios poderia eventualmente levar à produção de resultados diferentes, mas penso que não teria interesse jogar tudo “numa cartada” pois como se pode verificar (exemplo de simulação 2) nem sempre se consegue um bom negócio!

Pelos testes efectuados e cujos resultados (alguns exemplos) podem ser vistos no anexo 1, a probabilidade de sucesso é de facto maior quando as empresas a analisar estão inseridas dentro do mesmo ramo de actividade.

Contudo, em certos ramos de actividade o risco é muito maior dada a variação desconforme que se verifica nas séries.

Estas variações poderão ter várias explicações que fogem da simples análise através do estudo do factor de correlação e tendência – por exemplo as acções da PT Multimédia que so-

freram variações superiores a 100 euros (17.8 euros em 17 de Novembro de 1999 e 139 euros em 10 de Março de 2000) – tornando-se necessário ter um conhecimento exterior sobre os motivos que levaram a estas variações.

No exemplo de simulação 2 verifica-se também que, apesar de estarem reunidas todas as condições para que as sucessivas operações de compra e venda dessem lugar à obtenção de lucro, o mesmo não se verificou, o que obrigava (para ter sucesso), a uma observação permanente da evolução do preço destas acções. Se repararmos no caso concreto deste exemplo, na evolução verificada ao longo das 6 transacções efectuadas temos:

Transacção	Capital na data da transacção
Em : 01-07-2000 Comprei B e vendi A.	(Capital em reserva + 2000 acções de B a 16.96) = 102900
Em : 13-07-2000 Comprei B .	(Capital em reserva + 3000 acções de B a 16.1) = 101180
Em : 14-07-2000 Comprei B .	(Capital em reserva + 4000 acções de B a 16.23) = 101570
Em : 18-07-2000 Comprei B .	(Capital em reserva + 5000 acções de B a 16.76) = 103690
Em : 28-07-2000 Comprei B .	(Capital em reserva + 6000 acções de B a 15.96) = 99690
Em : 01-08-2000 Comprei B .	(Capital em reserva + 2000 acções de B a 16.00) = 99930

Tabela 6.5- Simulação da Evolução do investimento

Convém salientar também que a aplicação dá como sugestão em 11 de Agosto que “Deve aguardar para fazer negócio” pois não se verifica qualquer oportunidade e que se tivesse realizado a simulação apenas respeitante ao período de 30-06-2000 até 18-07-2000, obteria um lucro de 3 690 euros.

No anexo 2 estão incluídos os valores diários das cotações das acções de ambas as empresas no período de 14 de Setembro de 1999 e 11 de Agosto de 2000 e, o estudo do factor de correlação entre ambas as séries considerados diferentes períodos de tempo para efeitos de análise.

Como se pode verificar pelos diferentes valores calculados do factor de correlação, poder-se-ão levantar duas questões:

Será que sendo o valor do factor de correlação entre as duas empresas no período de 24-07-2000 a 11-08-2000 de 0,493940, este valor recomendaria qualquer estudo para efeitos de investimento?

Qual a importância do histórico neste tipo de análise ?

A resposta à segunda pergunta já foi dada pois sem histórico não era possível efectuar qualquer tipo de previsão (a não ser de intuito pessoal).

Como também se pode verificar nos outros exemplos apresentados, o facto de possuir um histórico leva-nos a verificar a solidez da existência, ou não, de uma correlação entre as cotações das empresas.

O método utilizado neste trabalho e um acompanhamento diário da evolução dos preços das cotações sem descuidar uma análise dos investimentos efectuados pode de facto levar à obtenção de alguns lucros.

6.4– Resumo

Neste capítulo foram abordadas algumas das dificuldades que senti na obtenção de dados, os resultados obtidos neste trabalho e abordadas algumas perspectivas de desenvolvimento futuro.

Relativamente aos resultados obtidos apenas foi abordado um dos objectivos deste trabalho – análise de séries temporais e oportunidades de negócio – sendo propositadamente omitido o outro objectivo – técnicas de bases de dados – porque relativamente a este último penso que foi detalhadamente explicado no capítulo 4 todo o trabalho realizado.

Infelizmente a quantidade de dados armazenados na base de dados não é significativa para poder efectuar comparações de desempenho e eficiência com outros sistemas de base de dados tendo-me limitado por isso a transcrever alguns testes realizados e que considero elucidativos das capacidades desta ferramenta para este tipo de aplicações.

Capítulo 7

Conclusão

Introdução

Depois de descrito o trabalho efectuado nesta tese, neste capítulo é analisada a satisfação dos objectivos propostos para a elaboração deste trabalho, os resultados obtidos e apresentadas perspectivas de desenvolvimento futuro.

7.1 – Satisfação de Objectivos

Neste trabalho pretendeu-se, através da implementação de técnicas para acesso e manipulação de dados, proceder à análise das cotações diárias das acções de empresas cotadas na Bolsa de Valores de Lisboa e, deste modo, tentar encontrar oportunidades de negócio.

Atendendo a que o que se pretendia era fazer previsões sobre o comportamento futuro do valor de pares de acções era inevitável possuir um histórico sobre o qual se pudessem efectuar essas previsões – uma base de dados onde constasse essa informação.

Foi sobre os valores das acções que foi realizado todo este trabalho: a implementação da base de dados (a forma mais eficiente de armazenagem de dados, não só pelos critérios que é possível definir para o seu armazenamento, como também pela facilidade que oferece para uma posterior manipulação dos mesmos), a recolha da informação, definição de critérios de análise e manipulação dos dados.

Os critérios definidos para a realização da análise dos dados e que nortearam todo o trabalho desenvolvido foram os seguintes:

1. Análise das séries temporais das cotações das acções através do estudo do comportamento das respectivas séries;
2. Estudo e implementação de técnicas de acesso e manipulação da informação armazenada na base de dados

Relativamente às séries temporais foi referido o interesse do estudo das mesmas para alcançar os objectivos propostos neste trabalho, e as técnicas utilizadas para o efeito, isto é, o cálculo do factor de correlação entre pares de acções pertencentes a empresas inseridas no mesmo ramo de actividade e as médias móveis de 5 e 21 dos mesmos pares de acções .

A técnica utilizada foi escolhida em função do trabalho a realizar e do período de tempo disponível para o efectuar.

Como é referido na secção seguinte, uma das eventuais perspectivas de desenvolvimento passa pelo refinamento dos critérios de análise das séries temporais, no entanto o estudo das médias móveis que nos mostram as tendências é imprescindível neste tipo de análise, ou seja, de previsão.

A implementação de técnicas de acesso e manipulação passou inicialmente pela escolha do sistema de base de dados a utilizar em virtude de ser necessário conseguir uma ferramenta que permitisse rapidez de execução e facilidade na manipulação dos dados. Essa escolha recaiu sobre a KDB.

Sobre esta ferramenta foi feita uma descrição, com realce para os pormenores que a tornam diferente de outros sistemas, nomeadamente as suas características OLAP e facilidade que oferece na manipulação de grandes quantidades de dados.

Esta ferramenta utiliza como linguagem para acesso e manipulação de dados além do SQL, o KSQL, tendo sido abordada com algum pormenor esta última linguagem através de exemplos de partes do código utilizado no desenvolvimento deste trabalho.

Sobre esta ferramenta foi também feita uma análise comparativa com outros sistemas de bases de dados e focados os aspectos em que eventualmente poderá haver vantagens na utilização deste sistema e nos problemas com que eventualmente nos poderemos deparar na sua utilização.

Finalmente foram apresentados alguns resultados obtidos com a aplicação dos critérios atrás definidos. Neste aspecto foram descritas as dificuldades com que me deparei na realização deste trabalho e que de certa forma influenciaram o resultado final.

Por exemplo se repararmos no gráfico representativo dos valores das acções da Portugal Telecom e da Telecel (Anexo 2), é de assumir que em 18 de Novembro de 1999 deve ter havido um *split* das acções da Telecel e, no dia 27 de Novembro de 1999 um *split* das acções da Portugal Telecom (mas, infelizmente, não foi possível recolher essa informação)

Concerteza que se dispusesse de toda a informação necessária a uma análise mais aprofundada das séries temporais, os resultados poderiam ser substancialmente diferentes pois permitir-me-iam fazer os necessários ajustamentos sempre que ocorresse um pagamento de dividendos ou um *split*.

Relativamente à interpretação dos resultados poder-se-á perguntar porque é que a metodologia utilizada não dá sempre lucro?

A resposta a esta pergunta foi dada na interpretação dos resultados uma vez que, e seguindo as regras definidas no capítulo 4, a opção de compra ou venda nas várias simulações efectuadas estava correcta.

Tal como é referido nas perspectivas de desenvolvimento futuro, a implementação de outros mecanismos de análise poderá produzir resultados diferentes.

Neste trabalho encontrar oportunidades de negócio era não só um objectivo mas também um motivo para alcançar outro objectivo – utilização de técnicas de acesso e manipulação de dados para análise de séries temporais – e pode ser um princípio para um futuro desenvolvimento.

7.2– Perspectivas de desenvolvimento futuro

Em que acções devo investir?

Esta é a pergunta que possivelmente todos gostariam de ver respondida e para a qual eu não tenho resposta.

Como anteriormente foi frisado, todo o trabalho foi desenvolvido tendo em vista a comparação entre duas séries de valores de acções diferentes e, com base nessa comparação, tentar encontrar oportunidades de negócio.

Este é pois um campo por onde poderá evoluir este trabalho, sendo que os objectivos terão que ser necessariamente diferentes – enquanto neste trabalho se pretendia dar uma resposta em tempo real a uma solicitação do utilizador e para o conseguir se tentou explorar as características do sistema de bases de dados utilizado, uma evolução do tipo da referida, dificilmente poderia dar uma resposta em tempo real.

No entanto o sistema de base de dados utilizado no desenvolvimento deste trabalho continua a ser a ferramenta ideal dado que permitiria com relativa rapidez de execução o estudo do comportamento de todas as séries de acções (através do cálculo dos respectivos factores de correlação e médias móveis) e a procura da melhor solução.

Eventualmente uma das técnicas a utilizar na pesquisa dessa solução poderia passar pela implementação de agentes e outros algoritmos de pesquisa que não o ID3 utilizado neste trabalho, visto que como foi referido oportunamente este algoritmo pode levar a complexas árvores de procura o que o tornaria ineficiente.

Uma outra hipótese de desenvolvimento futuro e que não foi implementada neste trabalho prende-se com a forma como foram estudadas as séries das acções, isto é, neste trabalho pegou-se em toda a série e aplicaram-se directamente os métodos de análise – factor de correlação e médias móveis (e através destas as tendências), sem ter a preocupação de analisar os valores constantes da série.

Num estudo mais aprofundado poder-se-iam aplicar conjuntamente outras técnicas de análise das séries, o que poderia conduzir à produção de soluções um pouco diferentes.

7.3– Consideração final

Nesta dissertação um dos objectivos era encontrar oportunidades de negócio em investimentos efectuados no mercado de acções da Bolsa de Valores de Lisboa sem que existisse por parte do investidor a necessidade de um conhecimento prévio quer do funcionamento deste tipo

de mercado, quer de factores de ordem económica que levam, ou podem levar, à variação das tendências de subida ou descida dos valores das cotações das acções.

Os resultados obtidos apontam para que um investidor com o perfil descrito no parágrafo anterior, possa, com a ajuda de uma ferramenta e metodologia utilizadas neste trabalho, efectuar investimentos neste tipo de produto (acções) minimizando o risco inerente a este tipo de operações financeiras.

BIBLIOGRAFIA

- [ADS99] - Arthur Whitney, Dennis Shasha, Stevan Apter, *High Volume transaction Processing*, Kx Systems , 1999
- [AEG92] - A. Ronald Gallant, Peter E. Rossi, George Tauchen, *Stock Prices and Volume*, The Review of Financial Studies ,Volume 5, número 2, 1992
- [AEO98] - Arbor Essbase , *OLAP Server 5*, InstantDoc ID: 3786 Setembro 1998, pag. 93, URL: <http://ntmag.com/Articles/>
- [AGA99] - Aryya Gangopadhyay, *Differences between olap and oltp*, URL: <http://research.umbc.edu/~gangopad/dware/sld003.htm>, atualização em 31-1-1999
- [CSC89] - Chapra, S.C., Canale, R.P., *Numerical Methods of Engineers*, McGraw-Hill; New York; 1989
- [DRL99] - The Journal of APL, Dogon Research Limited, <Http://www.apl.demon.co.uk/aplandj/olap.html>, editada em 31-Jan-99
- [EMP99] - Eamonn J. Keogh e Michael J. Pazzani, SIGIR99
- [FAT00] - *Fundamentos de análise técnica e seus instrumentos* , URL: <http://www.geocities.com/WallStreet/2500/przpanap.htm>, 10-05-2000
- [HNA+199] - Henry Naftulin, *Organizing Information for the Internet and E-commerce*, ACM SIGIR'99
- [HNA99] - Henry Naftulin, UC Davis Computer Science Dept., Sybase,Inc., ACM SIGIR'99.
- [JJM87] - John J. Murphy, *Technical Analysis of the Futures Markets*, New York Institute of Finance-70 Pine Street-New York, New York 10270
- [JZK98] - John Zukowski, *Mastering Java 1.2*, SYBEX, 1998, ISBN 0-7821-2180-2
- [KDE99] - Kalen Delaney, *Inside SQL Server: SQL Server 7.0 Transaction Log*, SQL Server Magazine Online, Maio 1999 , URL: <http://www.microsoft.com>
- [KXC00] - *The C - K Interface* - Kx Systems, 1999
- [KXD00] - Ficheiro DDJ.java, Dr.Dobb's Journal, de 19-03-2000

-
- [KXS00] - *KDP Programming Guide*, Kx Systems , 2000
- [KXS99] - *KSQL Reference Manual*, Kx Systems, 1999
- [LJJ99] - Louis K.C.Chan, Jason Karseski, Josef Lakonishok, *On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model*, The review of Financial Studies , v 12 n5, 1999
- [MGD92] - Mc Goveran D., C.J. Date, *Guide to sybase and sql server* , Hardcover Addison- Wesley Pub Co, ISBN 020155710X, 1992
- [MSO98] - Mukul Sood, Examining JDBC Drivers, Dr. Dobb's Journal January 1998, URL:<http://www.ddj.com/articles/1998/9801/9801i/9801i.htm>
- [ORA00] - <http://www.oracle.com/database/features/index.html?expressprod.html> em 26-9-2000
- [QJR86] - Quinlan J. R. [1986]. *Introduction of Decision Trees, Machine Learning*, Vol 3, pp. 81-106
- [RAR98] - Raghu Ramakrishnan, *Database Management Systems*, McGRAW-Hill International Editions, 1998, ISBN 0-07-115508-2
- [RSN97] - Roger Snowden , *Application Tuning*, URL:<Http://www.oreview.com>
- [SBA00] - Steven B. Achelis, *Technical Analysis from A to Z*, URL:<Http://www.equis.com/free/taaz/index.html>
- [SHA99] - Dennis Shasha, *Time Series in Finance: the array database approach*, Abril de 1999, ACM SIGMOD
- [SSM98] - Concept Learning and Decision Trees -Serengul Smith - School of Computing Scien-ce Middlesex University Data da última revisão: Setembro 1998- <http://www.cs.mdx.ac.uk/staffpages/serengul/table.of.contents.htm>
- [WBH99] - *Transaction Management in Enterprise Application Server 3.0*, www.sysbase.com
- [WRT94] - Weng-Ling Lin, Robert F. Engle, Takatoshi Ito, *The review of Financial Studies*, International Transmission of Stock Returns and Volatility – v/7 n°3 1994

Anexo 1

1 – Introdução

Neste anexo são incluídos alguns exemplos que traduzem os resultados obtidos através da implementação das regras de análise das séries temporais descritas no capítulo 4.

Em todas as simulações foram estipulados os mesmos critérios que já tinham sido definidos nos exemplos descritos no capítulo 6, ou seja:

- Inicialmente procedia à aquisição de 1 000 acções de cada empresa;
- Não tinha restrições de compra sempre que aparecesse uma oportunidade de negócio nem que isso implicasse reforços de capital;
- Não podia adquirir ou vender mais de 1000 acções de cada vez.

Em cada exemplo é analisada a vantagem ou desvantagem da aplicação da técnica de análise abordada neste trabalho, através da comparação do valor final quando aplicada esta técnica e o valor que teria se apenas se limitasse a deter as acções em carteira durante o mesmo período de tempo.

Exemplo 1 – Novabase e Pararede

Neste exemplo, talvez por uma das empresas objecto de análise – a Novabase SGPS – ainda só recentemente ter sido admitida a cotação na Bolsa de Valores de Lisboa, apesar das duas empresas estarem inseridas no mesmo ramo de actividade, o factor de correlação entre elas é negativo.

Isto pode-se explicar talvez pelo facto de uma das séries a estudar não possuir um grande histórico (número de registos na base de dados).

No entanto não deixa de ser interessante realizar a simulação e observar que o resultado final se traduz num lucro significativo.

Empresas analisadas: **Novabase SGPS Nom** e **Pararede SGPS**

Factor de Correlação: -0.684

Período considerado: De 21-07-2000 a 11-08-2000

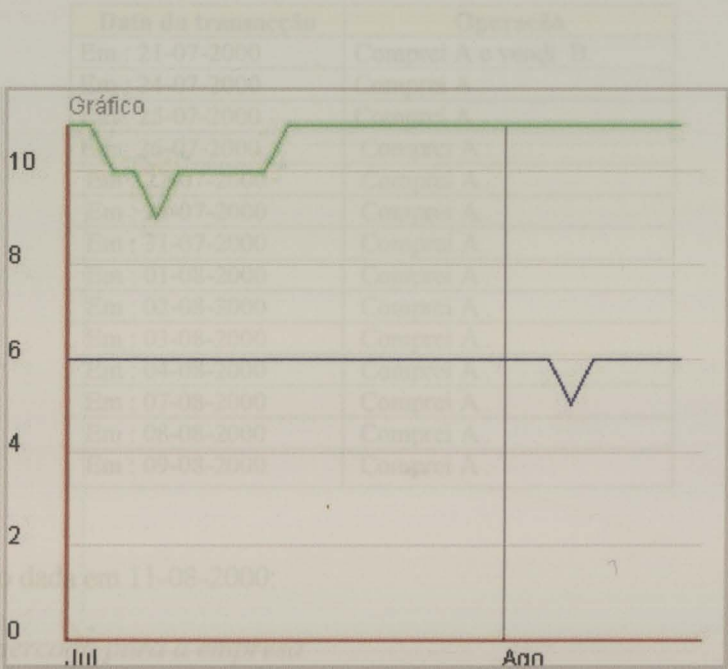


Figura A.1- Gráfico representativo da evolução das cotações da Novabase e Pararede

Investimento inicial em : 21-07-2000

- 1000 acções de A a 11.47 - 11470.0

- 1000 acções de B a 6.41 - 6410.0

Capital em Reserva (para investimento): 82120.0

Total (capital investido + reserva): 100000.0

Capital total final em : 11-08-2000

5000 accções de A a 11.99

0 acções de B a 6.14

Total (capital investido - reforços + valor em carteira): 110040.0

Valor que teria sem a aplicação do método: $(1000 \times 11.99 + 1000 \times 6.14 + 82120$
(Capital em reserva)) = **100250**

Evolução no período considerado: 15 sessões:

Data da transacção	Operação
Em : 21-07-2000	Comprei A e vendi B.
Em : 24-07-2000	Comprei A .
Em : 25-07-2000	Comprei A .
Em : 26-07-2000	Comprei A .
Em : 27-07-2000	Comprei A .
Em : 28-07-2000	Comprei A .
Em : 31-07-2000	Comprei A .
Em : 01-08-2000	Comprei A .
Em : 02-08-2000	Comprei A .
Em : 03-08-2000	Comprei A .
Em : 04-08-2000	Comprei A .
Em : 07-08-2000	Comprei A .
Em : 08-08-2000	Comprei A .
Em : 09-08-2000	Comprei A .

Sugestão dada em 11-08-2000:

A tendência de mercado para a empresa

Novabase SGPS Nom é de Alta

A tendência de mercado para a empresa

Pararede SGPS é de Baixa

A minha sugestão é:

Deve comprar as acções da firma

Novabase SGPS Nom e vender as acções da firma

Pararede SGPS

Exemplo 2 - C S Mundial Confiança Nom e C S Tranquilidade Nom

Neste exemplo são analisadas duas empresas inseridas dentro do mesmo ramo de actividade e que demonstra que apesar do curto período de tempo utilizado na análise, a vantagem da utilização da técnica abordada neste trabalho é significativa.

Empresas analisadas: **C S Mundial Confiança Nom** e **C S Tranquilidade Nom**

Factor de Correlação: 0.758

Período considerado: De 30-06-2000 a 11-08-2000

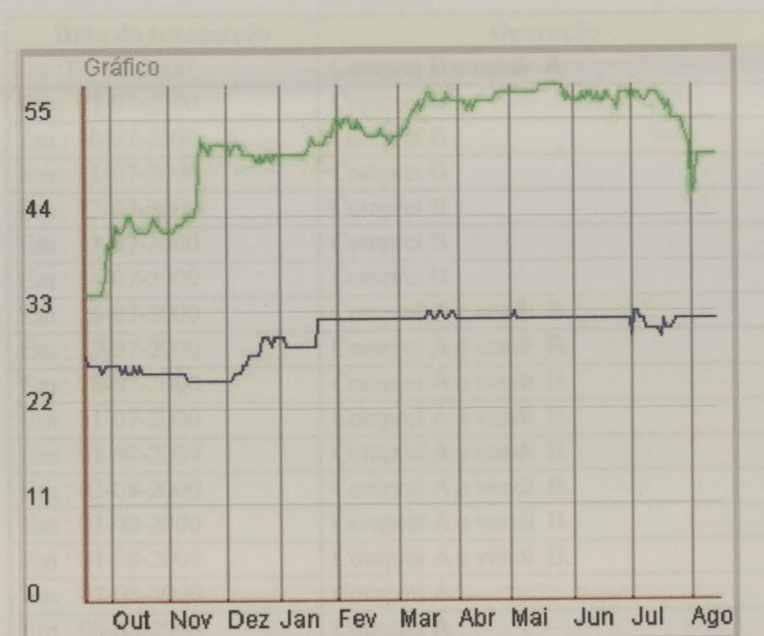


Figura A.2- Gráfico representativo da evolução das cotações da Mundial Confiança e Tranquilidade

Investimento inicial em : 30-06-2000

- 1000 acções de A a 58.0 - 58000.0

- 1000 acções de B a 32.8 - 32800.0

Capital em Reserva (para investimento): 9200.0

Total (capital investido + reserva): 100000.0

Capital total final em : 11-08-2000

11000 acções de A a 51.0

0 acções de B a 32.26

Total (capital investido - reforços + valor em carteira): 101600.0

Valor que teria sem a aplicação do método: $(1000 \times 51.0 + 1000 \times 32.26 + 9200 \text{ (Capital em reserva)}) = 92460$

Evolução no período considerado: 30 sessões:

Data da transacção	Operação
Em : 01-07-2000	Comprei B e vendi A.
Em : 07-07-2000	Comprei B .
Em : 10-07-2000	Comprei B .
Em : 11-07-2000	Comprei B .
Em : 12-07-2000	Comprei B .
Em : 13-07-2000	Comprei B .
Em : 14-07-2000	Comprei B .
Em : 26-07-2000	Comprei A e vendi B.
Em : 27-07-2000	Comprei A e vendi B.
Em : 28-07-2000	Comprei A e vendi B.
Em : 31-07-2000	Comprei A e vendi B.
Em : 01-08-2000	Comprei A e vendi B.
Em : 02-08-2000	Comprei A e vendi B.
Em : 03-08-2000	Comprei A e vendi B.
Em : 04-08-2000	Comprei A e vendi B.
Em : 07-08-2000	Comprei A .
Em : 08-08-2000	Comprei A .
Em : 09-08-2000	Comprei A .

Sugestão dada em 11-08-2000:

A tendência de mercado para a empresa

C S Mundial Confiança Nom é de Baixa

A tendência de mercado para a empresa

C S Tranquilidade Nom é de Alta

A minha sugestão é:

Deve comprar as acções da firma

C S Mundial Confianca Nom e vender as acções da firma

C S Tranquilidade Nom

- 1000 ações de A a 10.15 = 10150.0

Exemplo 3 – Mota e Companhia e Teixeira Duarte Eng Construções

Capital em Reserva (para investimento): 77000.0

Total (capital investido + reserva): 100000.0

Neste exemplo a simulação é efectuada entre duas empresas inseridas dentro do mesmo ramo de actividade, sendo a análise efectuada por um período de tempo de cerca de 10 meses.

38000.0 ações de B a 18.4

Empresas analisadas: Mota e Companhia e Teixeira Duarte Eng Construcoes

Factor de Correlação: 0.817

Período considerado: De 19-10-1999 a 11-08-2000

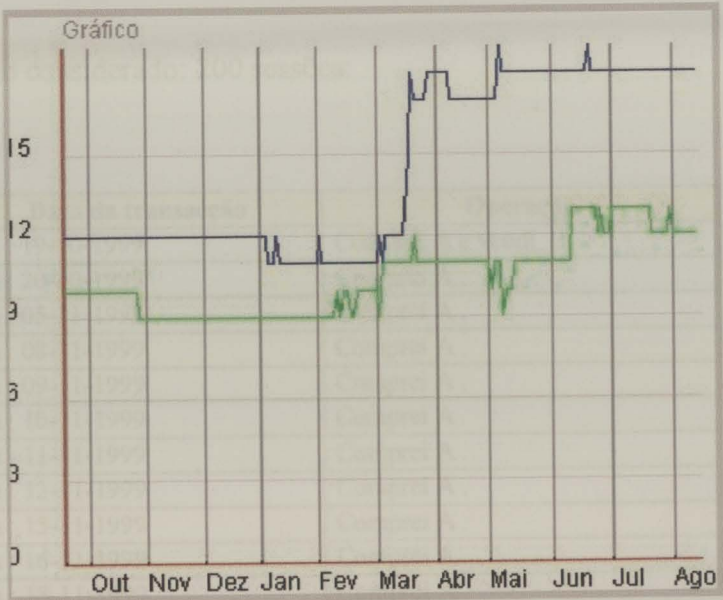


Figura A.3- Gráfico representativo da evolução das cotações da Mota & C e T. Duarte

Investimento inicial em :

19-10-1999

- 1000 acções de A a 10.15 - 10150.0

- 1000 acções de B a 12.79 - 12790.0

Capital em Reserva (para investimento): 77060.0

Total (capital investido + reserva): 100000.0

Capital total final em : 11-08-2000

5000.0 accções de A a 12.92

38000.0 accções de B a 18.4

Total (capital investido - reforços + valor em carteira): 321730.0

Valor que teria sem a aplicação do método: $(1000 \times 12.92 + 1000 \times 18.4 + 77060 \text{ (Capital em reserva)}) = 108380.00$

Evolução no período considerado: 200 sessões:

Data da transacção	Operação
Em : 19-10-1999	Comprei A e vendi B.
Em : 20-10-1999	Comprei A .
Em : 05-11-1999	Comprei A .
Em : 08-11-1999	Comprei A .
Em : 09-11-1999	Comprei A .
Em : 10-11-1999	Comprei A .
Em : 11-11-1999	Comprei A .
Em : 12-11-1999	Comprei A .
Em : 15-11-1999	Comprei A .
Em : 16-11-1999	Comprei A .
Em : 18-11-1999	Comprei A .
Em : 20-11-1999	Comprei A .
Em : 30-11-1999	Comprei B e vendi A.
Em : 02-12-1999	Comprei B e vendi A.
Em : 04-12-1999	Comprei B e vendi A.
Em : 22-12-1999	Comprei B e vendi A.
Em : 24-12-1999	Comprei B e vendi A.
Em : 05-01-2000	Comprei B e vendi A.
Em : 06-01-2000	Comprei B e vendi A.
Em : 09-01-2000	Comprei B e vendi A.
Em : 10-01-2000	Comprei B e vendi A.
Em : 16-01-2000	Comprei B e vendi A.

Em : 18-01-2000	Comprei B e vendi A.
Em : 18-01-2000	Comprei B e vendi A.
Em : 20-01-2000	Comprei B e vendi A.
Em : 20-01-2000	Comprei B .
Em : 21-01-2000	Comprei B .
Em : 25-01-2000	Comprei B .
Em : 26-01-2000	Comprei B .
Em : 28-01-2000	Comprei B .
Em : 28-01-2000	Comprei B .
Em : 02-02-2000	Comprei A e vendi B.
Em : 03-02-2000	Comprei A e vendi B.
Em : 04-02-2000	Comprei A e vendi B.
Em : 07-02-2000	Comprei A e vendi B.
Em : 08-02-2000	Comprei A e vendi B.
Em : 10-02-2000	Comprei B e vendi A.
Em : 11-02-2000	Comprei B e vendi A.
Em : 14-02-2000	Comprei B e vendi A.
Em : 15-02-2000	Comprei B e vendi A.
Em : 16-02-2000	Comprei B e vendi A.
Em : 17-02-2000	Comprei B .
Em : 19-02-2000	Comprei B .
Em : 22-02-2000	Comprei B .
Em : 22-02-2000	Comprei B .
Em : 23-02-2000	Comprei B .
Em : 24-02-2000	Comprei B .
Em : 26-02-2000	Comprei B .
Em : 28-02-2000	Comprei B .
Em : 29-02-2000	Comprei B .
Em : 01-03-2000	Comprei B .
Em : 14-03-2000	Comprei B .
Em : 15-03-2000	Comprei B .
Em : 16-03-2000	Comprei B .
Em : 18-03-2000	Comprei B .
Em : 20-03-2000	Comprei B .
Em : 23-03-2000	Comprei B .
Em : 23-03-2000	Comprei A e vendi B.
Em : 19-04-2000	Comprei A e vendi B.
Em : 20-04-2000	Comprei A e vendi B.
Em : 25-04-2000	Comprei A e vendi B.
Em : 08-05-2000	Comprei A e vendi B.
Em : 09-05-2000	Comprei A e vendi B.
Em : 11-05-2000	Comprei A e vendi B.
Em : 11-05-2000	Comprei A e vendi B.
Em : 13-05-2000	Comprei A e vendi B.
Em : 15-05-2000	Comprei A e vendi B.
Em : 16-05-2000	Comprei A e vendi B.
Em : 18-05-2000	Comprei A e vendi B.
Em : 30-05-2000	Comprei B e vendi A.
Em : 30-05-2000	Comprei B e vendi A.
Em : 31-05-2000	Comprei B e vendi A.
Em : 01-06-2000	Comprei B e vendi A.

Em : 03-06-2000	Comprei B e vendi A.
Em : 05-06-2000	Comprei B e vendi A.
Em : 06-06-2000	Comprei B e vendi A.
Em : 07-06-2000	Comprei B e vendi A.
Em : 08-06-2000	Comprei B e vendi A.
Em : 09-06-2000	Comprei B e vendi A.
Em : 12-06-2000	Comprei B e vendi A.
Em : 13-06-2000	Comprei B e vendi A.
Em : 14-06-2000	Comprei B .
Em : 01-07-2000	Comprei B .
Em : 13-07-2000	Comprei B .
Em : 14-07-2000	Comprei B .
Em : 18-07-2000	Comprei B .
Em : 19-07-2000	Comprei B .
Em : 20-07-2000	Comprei B .
Em : 21-07-2000	Comprei B .
Em : 03-08-2000	Comprei A e vendi B.
Em : 04-08-2000	Comprei A e vendi B.
Em : 07-08-2000	Comprei A e vendi B.
Em : 08-08-2000	Comprei A e vendi B.
Em : 09-08-2000	Comprei A e vendi B.

Sugestão dada em 11-08-2000:

A tendência de mercado para a empresa

Mota e Companhia é de Baixa

A tendência de mercado para a empresa

Teixeira Duarte Eng Construcoes é de Alta

A minha sugestão é:

Deve comprar as acções da firma

Mota e Companhia e vender as acções da firma

Teixeira Duarte Eng Construcoes

Exemplo 4 – Centralcer C Cervejas Nom e Sumolis C I F Bebidas Nom

Capital em Reserva (para investimento): 60910.0

Total (capital investido + reserva): 100000.0

Neste exemplo é novamente repetido o exemplo de simulação 6.2.1 (página 77) mas utilizando para simulação um período mais dilatado – 200 sessões.

Como se pode verificar no período final, enquanto que no exemplo anterior apenas procedia à aquisição de acções da empresa A, neste exemplo e porque detém em carteira acções da empresa B, procede à sua venda.

Empresas analisadas: **Centralcer C Cervejas Nom** e **Sumolis C I F Bebidas Nom**

Factor de Correlação: 0.8108

Período considerado: De 19-10-1999 a 11-08-2000

Evolução no período considerado: 200 sessões

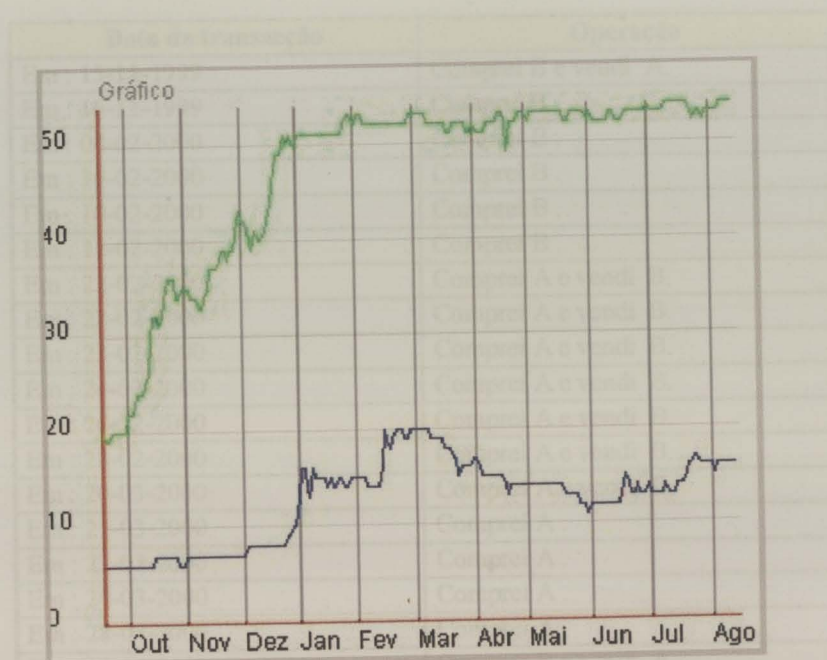


Figura A.4- Gráfico representativo da evolução das cotações da Centralcer e Sumolis

Investimento inicial em :19-10-1999

- 1000 acções de A a 31.98 - 31980.0

- 1000 acções de B a 7.11 - 7110.0

Capital em Reserva (para investimento): 60910.0

Total (capital investido + reserva): 100000.0

Capital total final em : 11-08-2000

11000.0 accções de A a 54.45

27000.0 acções de B a 16.68

Total (capital investido - reforços + valor em carteira): 286670.0

Valor que teria sem a aplicação do método: $(1000 \times 54.45 + 1000 \times 16.68 + 60910 \text{ (Capital em reserva)}) = 132040.00$

Evolução no período considerado: 200 sessões:

Data da transacção	Operação
Em : 15-11-1999	Comprei B e vendi A.
Em : 16-11-1999	Comprei B .
Em : 08-02-2000	Comprei B .
Em : 10-02-2000	Comprei B .
Em : 10-02-2000	Comprei B .
Em : 11-02-2000	Comprei B .
Em : 22-02-2000	Comprei A e vendi B.
Em : 22-02-2000	Comprei A e vendi B.
Em : 23-02-2000	Comprei A e vendi B.
Em : 24-02-2000	Comprei A e vendi B.
Em : 26-02-2000	Comprei A e vendi B.
Em : 28-02-2000	Comprei A e vendi B.
Em : 20-03-2000	Comprei A e vendi B.
Em : 21-03-2000	Comprei A .
Em : 23-03-2000	Comprei A .
Em : 24-03-2000	Comprei A .
Em : 28-03-2000	Comprei A .
Em : 13-04-2000	Comprei B e vendi A.
Em : 14-04-2000	Comprei B e vendi A.
Em : 18-04-2000	Comprei B e vendi A.
Em : 29-04-2000	Comprei B e vendi A.
Em : 03-05-2000	Comprei B e vendi A.
Em : 04-05-2000	Comprei B e vendi A.
Em : 04-05-2000	Comprei B e vendi A.

Em : 05-05-2000	Comprei B e vendi A.
Em : 08-05-2000	Comprei B e vendi A.
Em : 09-05-2000	Comprei B e vendi A.
Em : 11-05-2000	Comprei B e vendi A.
Em : 11-05-2000	Comprei B .
Em : 16-05-2000	Comprei B .
Em : 18-05-2000	Comprei B .
Em : 19-05-2000	Comprei B .
Em : 21-05-2000	Comprei B .
Em : 22-05-2000	Comprei B .
Em : 24-05-2000	Comprei B .
Em : 24-05-2000	Comprei B .
Em : 26-05-2000	Comprei B .
Em : 30-05-2000	Comprei B .
Em : 31-05-2000	Comprei B .
Em : 23-06-2000	Comprei B .
Em : 26-06-2000	Comprei B .
Em : 27-06-2000	Comprei B .
Em : 28-06-2000	Comprei B .
Em : 29-06-2000	Comprei B .
Em : 30-06-2000	Comprei B .
Em : 01-07-2000	Comprei B .
Em : 03-07-2000	Comprei B .
Em : 04-07-2000	Comprei B .
Em : 05-07-2000	Comprei B .
Em : 06-07-2000	Comprei B .
Em : 07-07-2000	Comprei B .
Em : 10-07-2000	Comprei B .
Em : 11-07-2000	Comprei B .
Em : 12-07-2000	Comprei B .
Em : 13-07-2000	Comprei B .
Em : 25-07-2000	Comprei A e vendi B.
Em : 26-07-2000	Comprei A e vendi B.
Em : 27-07-2000	Comprei A e vendi B.
Em : 28-07-2000	Comprei A e vendi B.
Em : 31-07-2000	Comprei A e vendi B.
Em : 01-08-2000	Comprei A e vendi B.
Em : 02-08-2000	Comprei A e vendi B.
Em : 03-08-2000	Comprei A e vendi B.
Em : 04-08-2000	Comprei A e vendi B.
Em : 07-08-2000	Comprei A e vendi B.
Em : 08-08-2000	Comprei A e vendi B.

Sugestão dada em 11-08-2000:

A tendência de mercado para a empresa

Centralcer C Cervejas Nom é de Alta

A tendência de mercado para a empresa

Sumolis C I F Bebidas Nom é de Alta

A minha sugestão é:

Deve aguardar para fazer negócio

Anexo 2

Neste anexo são representadas as tabelas respeitantes ao preço das cotações das acções das empresas Portugal Telecom e Telecel e, aos factores de correlação entre as duas séries de valores considerando diferentes períodos de tempo.

Tabela de Preços das cotações

Data	Preço – Portugal Telecom	Preço - Telecel
14-Set-99	37,80	115,00
17-Set-99	37,50	113,75
20-Set-99	38,00	114,25
21-Set-99	39,10	113,96
22-Set-99	38,86	113,99
23-Set-99	38,16	111,96
24-Set-99	39,47	111,29
27-Set-99	39,50	110,75
28-Set-99	39,60	112,40
01-Out-99	39,08	115,88
04-Out-99	39,00	119,25
06-Out-99	39,94	122,50
07-Out-99	40,00	126,95
08-Out-99	40,10	129,29
11-Out-99	39,49	128,00
12-Out-99	40,00	127,29
13-Out-99	40,22	129,22
14-Out-99	39,80	130,01
15-Out-99	39,67	129,00
18-Out-99	39,30	128,98
19-Out-99	38,58	127,67
20-Out-99	39,00	127,40
21-Out-99	39,45	129,48
22-Out-99	39,30	132,25
25-Out-99	40,00	134,50
26-Out-99	40,21	131,10
27-Out-99	41,70	129,44
28-Out-99	41,55	127,70
29-Out-99	41,70	123,30
02-Nov-99	42,40	123,40
03-Nov-99	43,50	119,20

04-Nov-99	44,25	117,99
05-Nov-99	44,89	118,70
08-Nov-99	44,47	121,00
09-Nov-99	43,10	120,99
10-Nov-99	42,05	118,00
11-Nov-99	42,55	116,52
12-Nov-99	43,64	114,80
15-Nov-99	44,62	114,50
16-Nov-99	45,30	118,35
18-Nov-99	44,30	12,70
20-Nov-99	44,50	12,94
22-Nov-99	46,50	12,87
23-Nov-99	46,65	12,80
24-Nov-99	45,38	12,50
25-Nov-99	46,80	12,24
27-Nov-99	9,47	12,70
29-Nov-99	9,59	13,44
30-Nov-99	9,62	13,35
02-Dez-99	9,55	13,14
04-Dez-99	9,98	13,82
06-Dez-99	10,32	14,60
07-Dez-99	10,40	16,13
10-Dez-99	10,61	16,26
11-Dez-99	10,90	18,26
14-Dez-99	10,83	16,70
14-Dez-99	10,84	16,03
15-Dez-99	10,90	16,70
16-Dez-99	10,31	16,06
17-Dez-99	10,23	16,02
20-Dez-99	10,47	16,91
21-Dez-99	10,53	16,50
22-Dez-99	10,41	16,79
23-Dez-99	10,41	17,00
27-Dez-99	10,93	17,40
28-Dez-99	10,89	17,30
29-Dez-99	10,86	16,85
30-Dez-99	10,86	17,03
03-Jan-00	10,89	17,31
04-Jan-00	10,66	17,22
05-Jan-00	10,00	16,54
06-Jan-00	9,40	16,02
09-Jan-00	10,15	16,24
10-Jan-00	10,60	16,70
11-Jan-00	10,97	16,31
12-Jan-00	10,77	16,13
15-Jan-00	10,80	16,06
16-Jan-00	11,42	16,06
17-Jan-00	12,45	17,17
18-Jan-00	12,18	16,60
19-Jan-00	12,59	17,25
20-Jan-00	12,35	17,53

21-Jan-00	11,85	18,15
25-Jan-00	11,70	18,39
26-Jan-00	12,29	19,00
28-Jan-00	12,70	19,00
29-Jan-00	12,40	19,01
31-Jan-00	12,60	18,75
01-Fev-00	12,20	18,30
02-Fev-00	12,44	18,30
03-Fev-00	12,24	18,47
04-Fev-00	12,49	19,40
07-Fev-00	13,35	21,65
08-Fev-00	13,82	22,15
9-Fev-00	14,01	20,50
10-Fev-00	14,40	21,59
11-Fev-00	14,35	20,21
14-Fev-00	14,86	21,05
15-Fev-00	14,65	20,30
16-Fev-00	13,98	20,50
17-Fev-00	13,78	20,00
19-Fev-00	14,24	20,31
20-Fev-00	14,22	21,40
22-Fev-00	14,44	20,59
23-Fev-00	14,35	21,45
24-Fev-00	15,10	21,70
26-Fev-00	15,10	22,50
28-Fev-00	15,26	23,55
29-Fev-00	14,96	23,95
01-Mar-00	14,96	23,95
03-Mar-00	14,90	24,50
04-Mar-00	15,60	25,20
06-Mar-00	16,83	25,00
09-Mar-00	16,65	25,00
10-Mar-00	16,35	25,01
11-Mar-00	15,88	25,00
13-Mar-00	15,30	25,30
14-Mar-00	14,90	24,48
15-Mar-00	15,13	23,85
16-Mar-00	14,10	23,10
18-Mar-00	14,02	22,90
20-Mar-00	14,10	22,80
21-Mar-00	14,70	22,60
22-Mar-00	14,62	23,28
23-Mar-00	13,65	22,26
24-Mar-00	13,20	21,45
27-Mar-00	14,10	23,85
28-Mar-00	13,85	22,40
30-Mar-00	14,00	22,55
31-Mar-00	14,00	22,55
01-Abr-00	13,16	20,81
03-Abr-00	13,40	20,40
04-Abr-00	12,50	20,17

05-Abr-00	12,06	18,90
06-Abr-00	12,30	18,05
08-Abr-00	13,05	20,00
10-Abr-00	13,26	21,30
11-Abr-00	13,11	22,19
12-Abr-00	12,01	20,99
13-Abr-00	12,29	20,50
14-Abr-00	11,90	20,80
18-Abr-00	11,10	18,95
19-Abr-00	11,21	19,00
20-Abr-00	11,75	18,45
25-Abr-00	11,60	17,81
27-Abr-00	11,66	18,25
28-Abr-00	12,02	17,80
29-Abr-00	12,01	17,25
03-Mai-00	12,27	17,40
04-Mai-00	12,60	18,14
05-Mai-00	12,98	18,50
06-Mai-00	12,21	18,41
08-Mai-00	12,43	18,85
09-Mai-00	11,92	18,51
11-Mai-00	11,50	18,37
12-Mai-00	10,75	17,81
13-Mai-00	10,98	17,62
15-Mai-00	11,13	17,52
16-Mai-00	10,65	17,40
18-Mai-00	11,32	18,00
19-Mai-00	10,99	17,40
21-Mai-00	11,32	17,25
22-Mai-00	10,89	16,50
23-Mai-00	10,80	15,81
24-Mai-00	10,80	16,89
26-Mai-00	10,31	16,15
28-Mai-00	10,56	16,25
29-Mai-00	10,80	16,20
30-Mai-00	11,20	16,20
31-Mai-00	11,61	17,22
01-Jun-00	11,70	17,40
03-Jun-00	12,15	17,90
05-Jun-00	12,50	18,55
06-Jun-00	12,01	18,00
07-Jun-00	12,20	17,94
08-Jun-00	12,04	17,56
09-Jun-00	12,05	17,15
12-Jun-00	12,43	17,28
13-Jun-00	12,43	17,42
14-Jun-00	12,51	17,89
15-Jun-00	12,78	17,85
16-Jun-00	12,69	17,99
19-Jun-00	12,41	17,70
20-Jun-00	12,15	17,41

21-Jun-00	11,90	17,06
23-Jun-00	11,90	17,10
26-Jun-00	11,35	16,85
27-Jun-00	11,09	16,56
28-Jun-00	11,42	16,47
29-Jun-00	11,94	16,17
30-Jun-00	11,66	15,50
01-Jul-00	13,10	16,96
03-Jul-00	11,76	15,90
04-Jul-00	11,90	16,25
05-Jul-00	11,90	15,95
06-Jul-00	12,03	15,99
07-Jul-00	11,95	15,81
10-Jul-00	12,33	16,00
11-Jul-00	12,42	15,90
12-Jul-00	12,70	15,95
13-Jul-00	12,85	16,10
14-Jul-00	12,97	16,23
18-Jul-00	13,50	16,76
19-Jul-00	13,25	16,60
20-Jul-00	13,10	16,40
21-Jul-00	13,63	16,83
24-Jul-00	13,31	16,41
25-Jul-00	13,80	16,25
26-Jul-00	13,12	16,08
27-Jul-00	13,00	15,80
28-Jul-00	12,99	15,96
31-Jul-00	11,75	15,90
01-Ago-00	11,75	16,00
02-Ago-00	12,15	16,02
03-Ago-00	12,35	16,19
04-Ago-00	11,81	16,03
07-Ago-00	11,90	16,12
08-Ago-00	11,65	15,90
09-Ago-00	11,38	15,98
10-Ago-00	11,89	15,81
11-Ago-00	12,07	15,75

Gráfico representativo da evolução das cotações da Portugal Telecom e Telecel

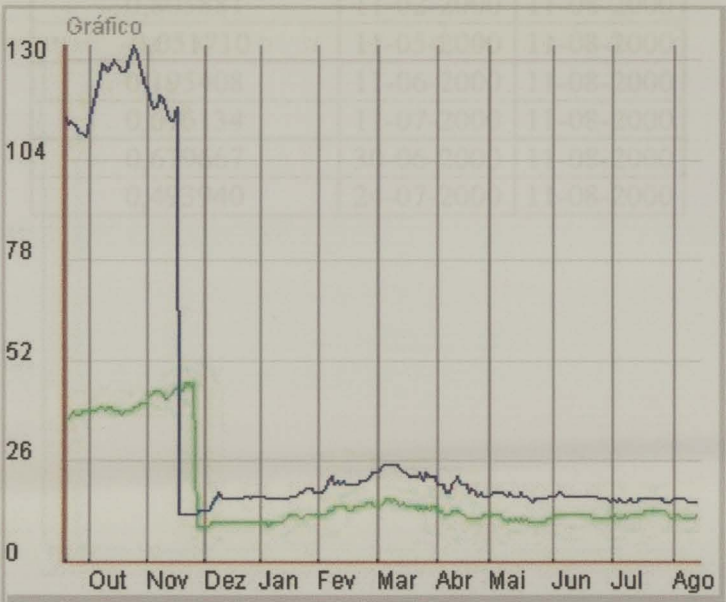


Figura A.5- Gráfico representativo da evolução das cotações da PT e Telecel

Factores de correlação considerando diferentes períodos de análise

Factor de correlação	Data Inicial	Data Final
0,871097	21-09-1999	11-08-2000
0,805543	03-01-2000	11-08-2000
0,805881	11-02-2000	11-08-2000
-0,051710	11-05-2000	11-08-2000
0,195408	11-06-2000	11-08-2000
0,676134	11-07-2000	11-08-2000
0,679667	30-06-2000	11-08-2000
0,493940	24-07-2000	11-08-2000

Anexo 3

Neste anexo estão descritas as funções implementadas em linguagem K e que tal como foi descrito anteriormente, quando chamadas directamente nas instruções implementadas em KSQL permitem que o servidor devolva o resultado.

/ funções implementadas em linguagem K

/ função que nos dá a média de uma série

avg: $\{ (+ / x) \% \# x \}$

/ função que nos dá a variância de uma série

var: $\{ avg[_sqr\ x] - _sqr\ avg[x] \}$

/ função que nos dá o desvio padrão de uma série

std: $\{ _sqrt\ var[x] \}$

/ função que nos dá a covariância de duas séries

cov: $\{ avg[x * y] - avg[x] * avg[y] \}$

/ função que nos dá a correlação entre duas séries

cor: $\{ cov[x; y] \% std[x] * std[y] \}$



FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

BIBLIOTECA



0000051694