

INSTITUTO POLITÉCNICO DE VIANA DO CASTELO
ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO

MESTRADO EM GESTÃO DAS ORGANIZAÇÕES
Ramo Gestão de Empresas

Recorrendo a uma Regressão Linear Múltipla

15 de Fevereiro de 2012

Disciplina:

Tratamento e Análise de Dados

Professor: José Miguel Veiga

Maria Cândida Cunha Costa - n° 12308

Conteúdos

Introdução.....	5
Concepção do modelo de previsão.....	6
Linearidade das variáveis independentes.....	6
.....	9
Criação e validação do modelo.....	11
1ª Regressão	11
Justificação das variáveis seleccionadas	11
Coeficientes - Variáveis a incluir no modelo.....	12
Validação do modelo: Verificação da Homocedasticidade.....	13
Validação do modelo: Teste à co-variância nula.....	14
Validação do modelo: Teste à normalidade dos resíduos	14
Validação do modelo: Teste à multicolinearidade.....	15
Reformulação do modelo.....	16
2ª e 3ª Regressões	16
5ª Regressão.....	17
6ª Regressão.....	17
Validação do modelo: Homocedasticidade.....	18
Validação do modelo: Teste à normalidade dos resíduos	19
Análise de outliers.....	19
Análise de casos influentes	20
Casos influentes: alavancagem:	20
Casos influentes: valor estimado ajustado	21
Distância de Cook.....	21
Casos influentes: DfBeta	22

Casos influentes: DfFit.....	22
Resumo: Oultiers e casos influentes:	23
Modelo Final	23
Estimação de um novo modelo refinado	23
Modelo de previsão	24
Conclusões	25
Anexos.....	Erro! Marcador não definido.

Introdução

O objectivo do trabalho será desenvolver um modelo matemático que permita prever a evolução do índice PSI 20 - Portuguese Stock Index, tendo em conta:

- As flutuações dos mercados internacionais, personificada através da introdução de variáveis como a evolução dos principais Índices mundiais;
- O desempenho das empresas cotadas no PSI 20;

Esta referência no mercado financeiro nacional reflecte o desempenho médio das vinte empresas com maior dimensão e liquidez existentes no Mercado de Cotações Oficiais. Com a sua génese foram dadas respostas a duas questões colocadas aos mercados financeiros nacionais:

- ***Servir de indicador da evolução do mercado accionista português;***
- ***Fornecer um suporte à negociação de contratos de futuros e opções.***

De forma a ser possível a criação de uma previsão, foram definidas um conjunto de variáveis entre as quais será retirada informação para que seja possível extrapolar um modelo de previsão. As diversas variáveis de entrada podem ser segmentadas em dois conjuntos diferentes:

1. Um conjunto de dados que deverá permitir medir a influência do mercado interno na evolução do Índice. Para tal foram seleccionadas todas as empresas que no último trimestre de 2010 estavam inseridas no PSI 20::

Altri, Banco Comercial Português, Banco Espírito Santo, Banco Português de Investimento, Brisa, Cimpor, EDP, EDP Renováveis, Galp, Inapa, Jerónimo Martins, Mota Engil, Portucel, Portugal Telecom, REN, Semapa, Sonae Indústria, Sonae, Sonaecom e ZON

2. O segundo grupo pretende reflectir a influência dos mercados internacionais através da inclusão das maiores praças mundiais, bem como aquelas que por tradição têm fortes laços comerciais com a economia nacional.

Ibovespa (Brasil), Nikkei (Japão), BSE (India), FTSE (Inglaterra), CAC (França), DAX (Alemanha), NASDAQ (USA-electronica), DOW (USA) e Hanseng (China)

Identificada a base de dados a analisar e feito o seu enquadramento, resta definir o limite de observações: o mês de Outubro de 2010 e as quatro primeiras semanas de Janeiro de 2011, sendo, portanto o principal objectivo criar um modelo de previsão a curto prazo.

Foram utilizados métodos baseados em regressões lineares, utilizando como plataforma de desenvolvimento o *software SPSS*.

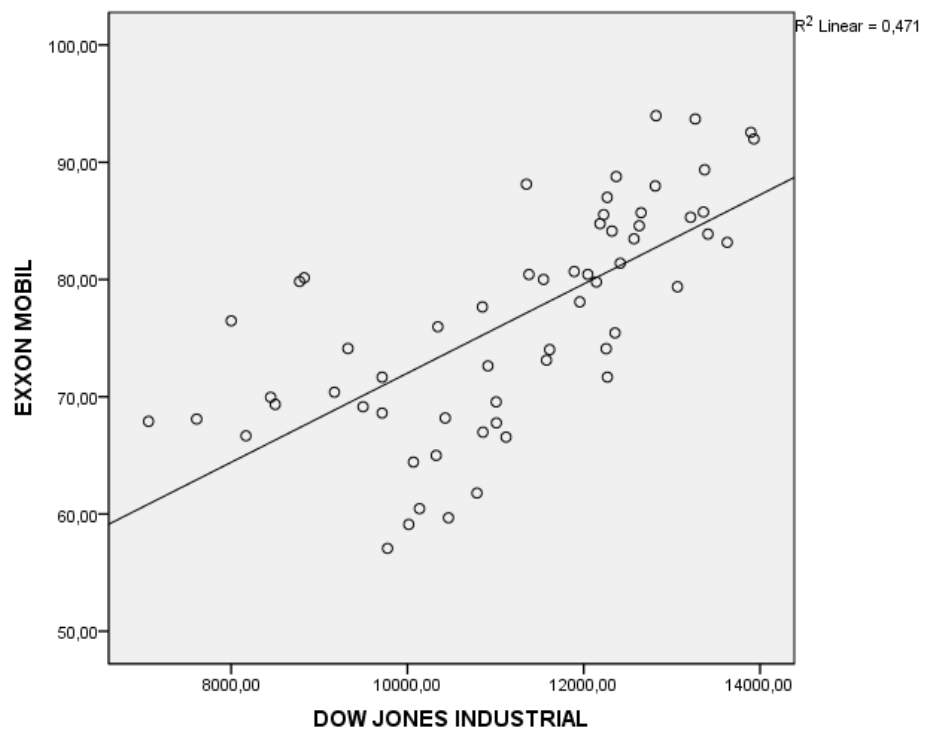
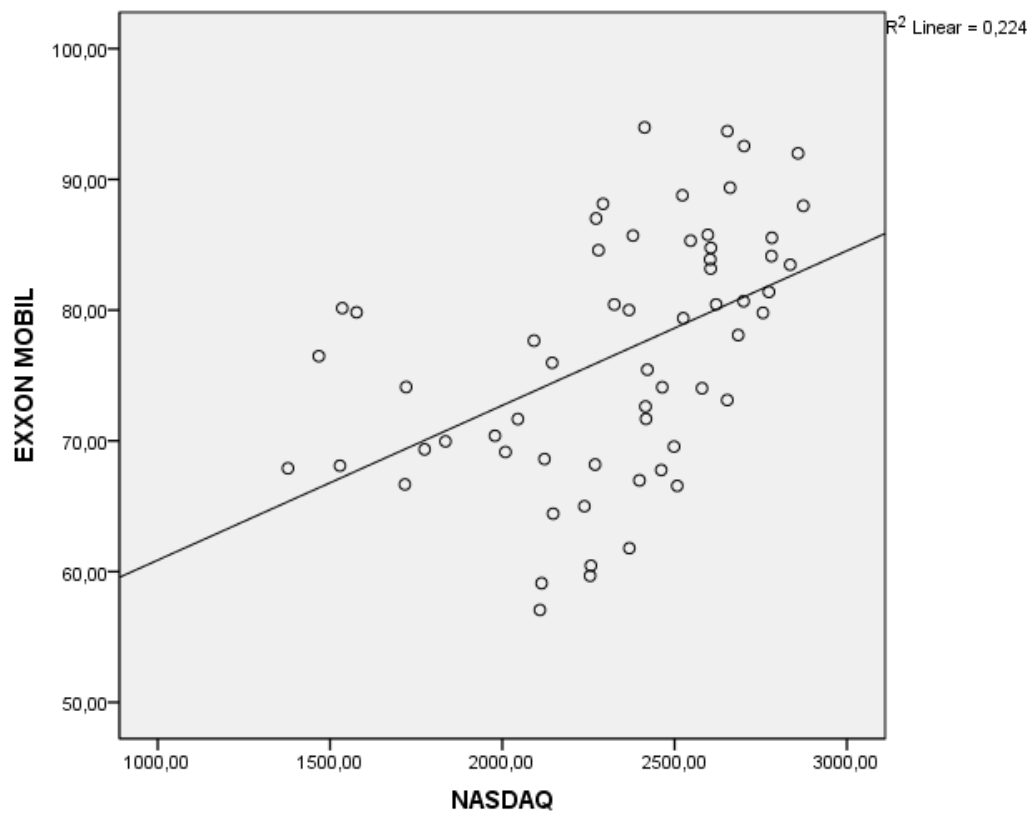
Concepção do modelo de previsão

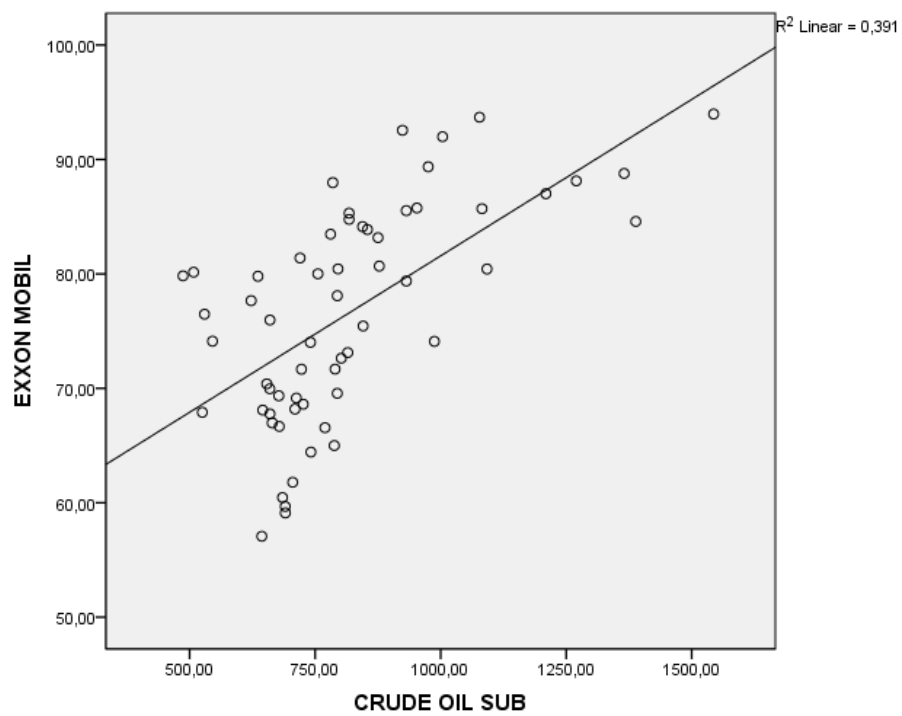
Perante este cenário, procura-se mensurar a influência de determinadas variáveis na cotação do índice nacional, obtendo-se assim um modelo que possibilita prever a cotação do índice do PSI20 num espaço temporal de uma semana. As variáveis a introduzir no modelo definem-se como sendo as cotações de fecho dos índices previamente referidos.

Na perspectiva de tornar mais poderoso o modelo em causa, foi decidido introduzir as cotações do dia anterior como inputs do modelo. Ou seja, para fazer previsão da cotação de fecho do PSI 20 para o dia seguinte ($n+1$), será apenas necessário estar na posse das cotações de fecho das variáveis de entrada no dia actual (n):

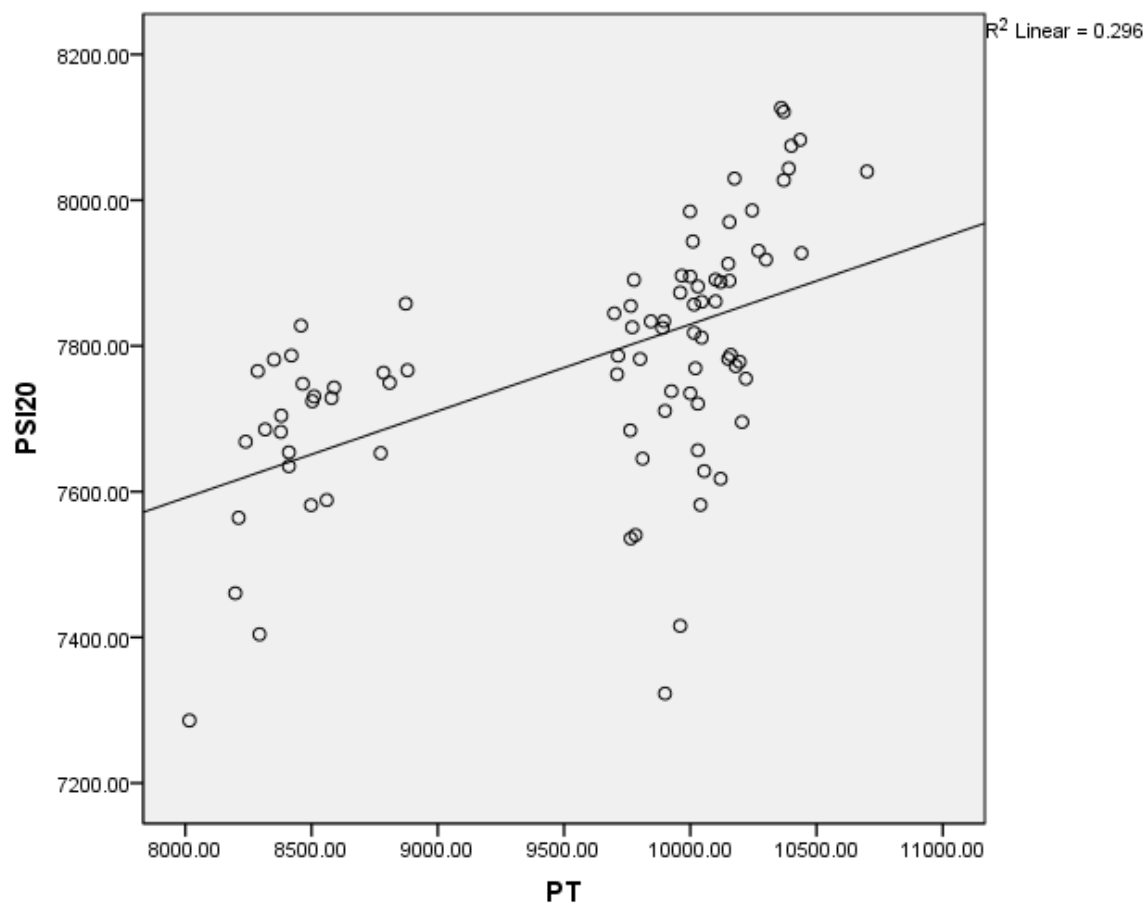
Linearidade das variáveis independentes

O primeiro passo a efectuar é verificar se as variáveis independentes são linearmente independentes. Para tal, para cada uma das vinte e nove variáveis foram criados gráficos de dispersão onde foram ajustadas linhas de tendência entre a variável independente e dependente. Nos gráficos seguintes encontram-se expostos os diagramas de dispersão das variáveis mais influentes na elaboração do modelo.

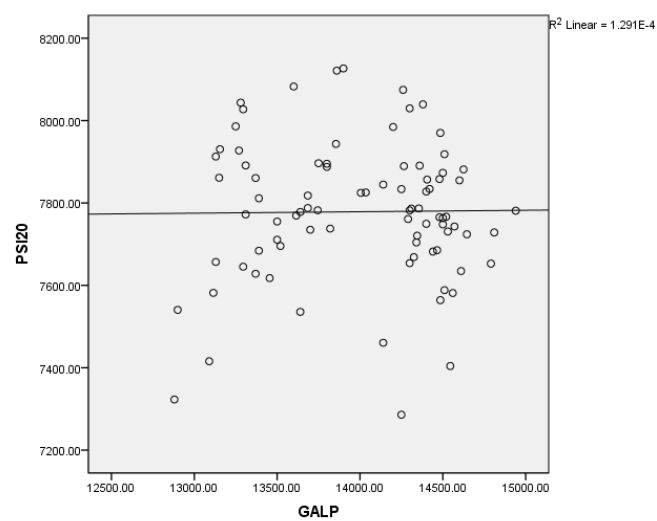




Como é possível observar nos gráficos, não existem sinais de não linearidade das variáveis pelo que se pode afirmar que são variáveis linearmente independentes, com uma intensidade relativamente forte e positiva.



As variáveis BES e PT não aparentam apresentar padrões não lineares, tem uma relação positiva e com menor intensidade do que as anteriores.



A Galp apesar de não apresentar nenhum padrão indesejado, demonstra ter uma relação muito reduzida com a variável dependente.

As restantes variáveis foram testadas, mas a sua exposição neste capítulo não traria qualquer tipo de valor acrescentado.

Criação e validação do modelo

Após a fixação de requisitos e defenição e objectivos, é necessário criar o modelo. Seleccionadas as variáveis independentes e dependentes, o passo seguinte é fazer uma primeira regressão.

1ª Regressão

Justificação das variáveis seleccionadas

A análise da tabela anova permite testar as seguintes hipóteses:

$$\begin{cases} H_0 & \forall \beta_i = 0 \\ H_1 & \beta_i \neq 0 \end{cases}$$

H_0 : Todos os coeficientes β são nulos, esta situação implica que as variáveis escolhidas não têm influência na variável dependente

H_1 : Pelo menos uma das variaveis independentes não nula, tendo assim influencia no modelo

Consultando um excerto da tabela Anova da regressão

ANOVA ^c						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2482,906	1	2482,906	51,661	,000 ^a
	Residual	2787,581	58	48,062		
	Total	5270,486	59			
2	Regression	2798,807	2	1399,403	32,272	,000 ^b
	Residual	2471,680	57	43,363		
	Total	5270,486	59			

a. Predictors: (Constant), DOW JONES INDUSTRIAL

b. Predictors: (Constant), DOW JONES INDUSTRIAL, CRUDE OIL SUB

c. Dependent Variable: EXXON MOBIL

A estatística de teste segue uma distribuição F com um valor de 51,661 e com uma significância de 0,000. Assim, para um intervalo de confiança de 95% vemos que o modelo 1 rejeita a hipótese nula, pois:

Sig<0,05 → $\alpha=0,05$; Significando esta situação que, pelo menos uma variável escolhida influencia a EXXON MOBIL.

Coefficientes - Variáveis a incluir no modelo

Após a análise da tabela Anova sabe-se que, pelo menos, uma variável tem influência no modelo. Para identificar quais as variáveis que se vão manter no modelo temos de testar a hipótese que nos identifica qual das variáveis não influenciam.

$$\begin{cases} H_0 & \beta_i = 0 \\ H_1 & \beta_i \neq 0 \end{cases}$$

Esta estatística de teste segue uma distribuição T Student e testa a nulidade β através de um teste T. Para todas as variáveis cuja significância seja inferior a determinado valor, é possível afirmar que o coeficiente não é nulo. Desta forma calcula-se as variáveis mais adequadas a incluir no modelo.

No caso em estudo as variáveis para as quais Sig < 0,05 são:

Coefficients^a

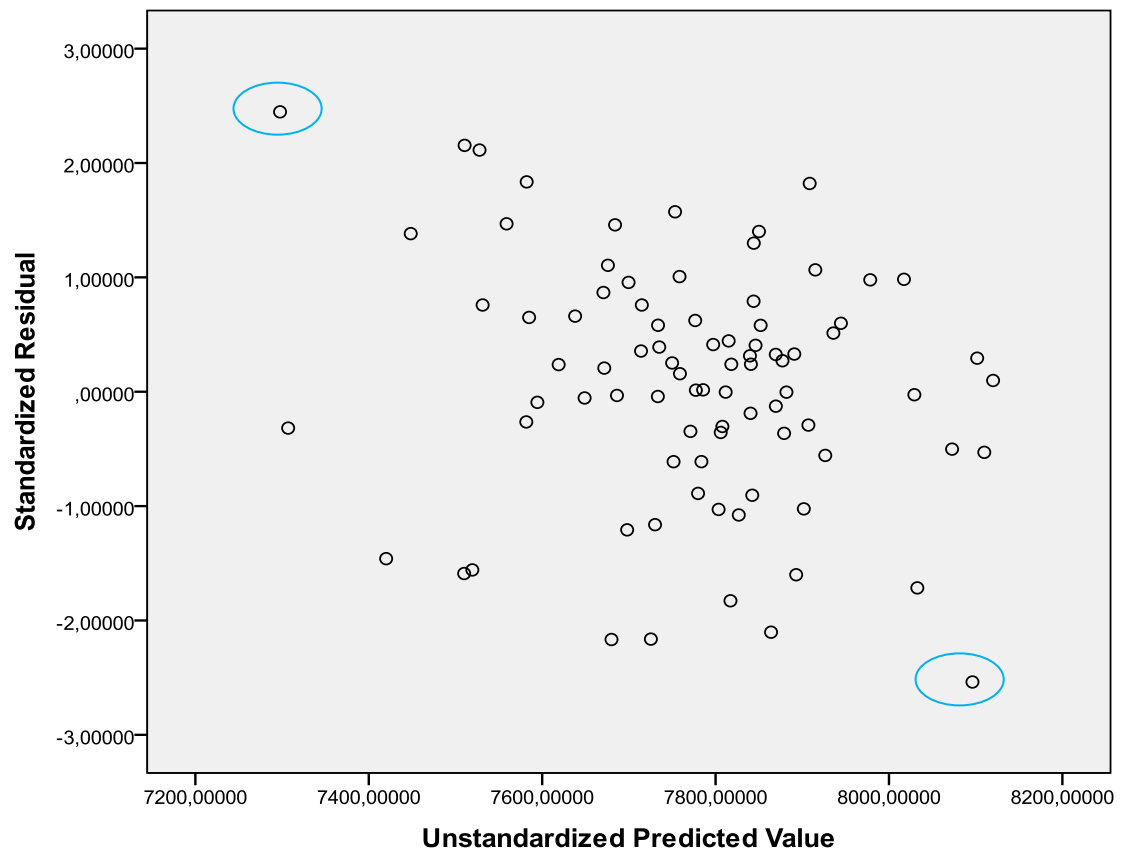
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
7 (Constant)	1664,812	521,029		3,195	,002					
SONAEOM	,942	,277	,453	3,404	,001	,811	,424	,164	,131	7,633
BRISA	,382	,079	,380	4,843	,000	,627	,554	,233	,377	2,649
BCP	4,524	,899	,732	5,035	,000	,684	,569	,243	,110	9,088
GALP	,136	,032	,405	4,280	,000	-,014	,507	,206	,259	3,855
BES	-,501	,123	-,895	-4,073	,000	,578	-,488	-,196	,048	20,778
JMARTINS	-,113	,030	-,523	-3,737	,000	-,144	-,457	-,180	,119	8,421
PT	,095	,036	,373	2,618	,011	,585	,338	,126	,114	8,751

Aplicando uma regressão linear a todas as variáveis, é possível obter sete modelos. O modelo utilizado é o sétimo, pois é o que possui maior coeficiente de correlação múltipla. Este modelo baseia-se na cotação de sete empresas: SonaeCom, Brisa, BCP, Galp, BES, Jmartinns, PT.

Calculado um primeiro modelo é agora necessário proceder à sua validação.

Validação do modelo: Verificação da Homocedasticidade

Este teste permite verificar se a covariância é constante. Para tal analisa-se a evolução dos resultados da previsão versus resíduos padronizados (segundo uma distribuição T de Student).



Não se observa nenhum padrão relativamente ao eixo horizontal pelo que é possível afirmar que a **homocedasticidade não é violada**.

É também possível identificar observações que aparentam ser *outliers*.

Validação do modelo: Teste à co-variância nula

Na tabela de resumo do modelo acima descrita, observamos um valor para o teste de *Durbin Watson* de 1,711 pelo que é lícito afirmar que não há autocorrelação dos resíduos.

Validação do modelo: Teste à normalidade dos resíduos

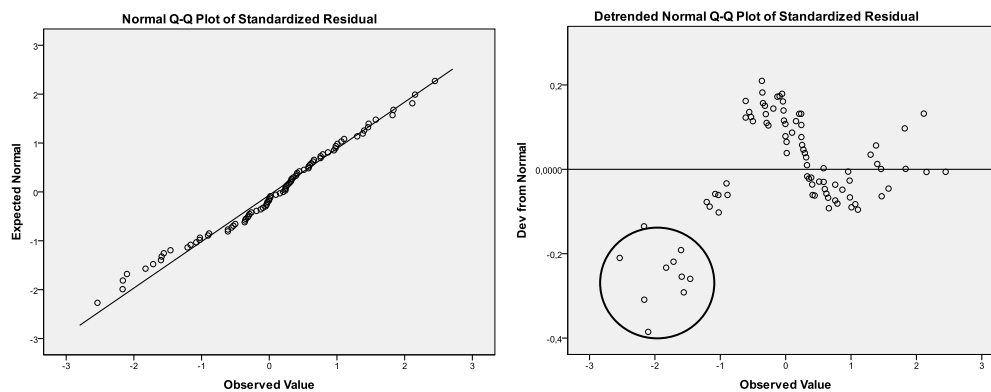
Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	,082	85	,200*	,984	85	,359

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Em ambos os testes as significâncias são superiores a $\alpha=0,05$ pelo que se deverá assumir que os resíduos seguem uma distribuição normal.



Analisando os gráficos, é possível confirmar uma disposição de observações em torno das linhas o que confirma o resultado do teste *Kolmogorov-Smirnov*.

Validação do modelo: Teste à multicolinearidade

A possível existência de multicolinearidade entre variáveis é verificada pelos três seguintes indicadores:

1. Correlações de Pearson

Verificação de quais os coeficientes de correlação de *Pearson* superiores a 0,9. Caso isto se verifique a variável será retirada do modelo. Esta situação permite indicar uma forte possibilidade de multicolinearidade.

Na tabela das correlações podemos observar correlações demasiado altas, sendo de referir as correlações entre BCP/BES/ PT

Correlations								
	PSI20	JMARTINS	BRISA	GALP	SONAECON	BES	PT	BCP
PSI20	1	-,106	,639**	,011	,804**	,564**	,544**	,632**
JMARTINS	-,106	1	,357**	,548**	-,309**	-,704**	-,324**	-,505**
BRISA	,639**	,357**	1	,368**	,412**	,008	-,007	,095
GALP	,011	,548**	,368**	1	-,120	-,521**	-,619**	-,570**
SONAECON	,804**	-,309**	,412**	-,120	1	,781**	,692**	,685**
BES	,564**	-,704**	,008	-,521**	,781**	1	,782**	,892**
PT	,544**	-,324**	-,007	-,619**	,692**	,782**	1	,843**
BCP	,632**	-,505**	,095	-,570**	,685**	,892**	,843**	1

** . Correlation is significant at the 0.01 level (2-tailed).

2. Tolerância e VIF:

Para que não existam problemas de multicolinearidade: $VIF = \frac{1}{\text{tolerância}}$ $VIF < 10$

Model		Correlations			Collinearity Statistics	
		Zero-order	Partial	Part	Tolerance	VIF
7	SONAECON	,811	,424	,164	,131	7,633
	BRISA	,627	,554	,233	,377	2,649
	BCP	,684	,569	,243	,110	9,088
	GALP	-,014	,507	,206	,259	3,855
	BES	,578	-,488	-,196	,048	20,778
	JMARTINS	-,144	-,457	-,180	,119	8,421
	PT	,585	,338	,126	,114	8,751

a. Dependent Variable: PSI20

Neste ponto é possível concluir que a condição $VIF < 10$ não é cumprida pela variável “BES”

3. Condition index e proporção da variância:

Condition Index terá um valor tal que:

$CI < 15$ Não existe multicolinearidade
 $15 < CI < 30$ Possível existência de multicolinearidade $CI = 1$
 $CI > 30$ Multicolinearidade severa

Collinearity Diagnostics ^a											
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions							
				(Constant)	SONAECON	BRISA	BCP	GALP	BES	J MARTINS	PT
7	1	7,978	1,000	,00	,00	,00	,00	,00	,00	,00	,00
	2	,017	21,621	,00	,00	,00	,00	,00	,01	,01	,00
	3	,003	55,345	,00	,00	,01	,00	,03	,01	,03	,12
	4	,001	79,397	,05	,21	,01	,02	,00	,00	,01	,00
	5	,001	107,795	,00	,01	,37	,02	,14	,01	,01	,03
	6	,000	169,579	,19	,12	,02	,03	,01	,47	,62	,25
	7	,000	228,217	,28	,08	,06	,81	,16	,47	,12	,01
	8	,000	250,620	,47	,58	,53	,11	,65	,02	,19	,60

Analisando os três indicadores de multicolinearidade verifica-se que a variável BES não cumpre os requisitos pelo que deverá ser excluída do modelo.

Reformulação do modelo

Construiu-se então uma nova regressão com todas as variáveis de entrada anteriores excluindo apenas o BES.

2ª e 3ª Regressões

A segunda e terceira iteração devolveram modelos que incluem os índices das bolsas brasileiras e indiana. Estas variáveis foram excluídas por já não se incluírem nos objectivos deste trabalho. Isto é não faz muito sentido fazer previsão das cotações do índice português em função das cotações de índices internacionais de países emergentes como são os casos do Brasil e Índia.

5ª Regressão

Removidas essas variáveis foi criado um novo modelo em que se inseria uma nova variável a **Altri**.

Em anexo encontra-se uma tabela que resume os resultados que foram obtidas nas diversas regressões e sucessivas validações aos respectivos modelos.

De seguida pode-se observar o “model summary” da regressão linear criada.

Model Summary ^f										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.806 ^a	.650	.644	103.35251	.650	116.940	1	63	.000	1.639
2	.858 ^b	.737	.728	90.32544	.087	20.483	1	62	.000	
3	.893 ^c	.798	.788	79.81302	.061	18.408	1	61	.000	
4	.906 ^d	.821	.810	75.63304	.024	7.929	1	60	.007	
5	.913 ^e	.833	.819	73.76821	.012	4.072	1	59	.048	

e. Predictors: (Constant), SONAECON, BRISA, BCP, GALP, ALTRI

Depois de testar todas as hipóteses da mesma forma que foi feito para a 1ª regressão, verificou-se que este modelo não passou na condição de multicolinearidade, pois possui uma elevada correlação com a variável BCP. Sendo assim, também esta variável foi removida do modelo.

6ª Regressão

Feita uma nova regressão chegou-se a um modelo com as seguintes variáveis: SONAECON, BRISA, BCP, GALP. De notar que a remoção da variável levou a uma diminuição da qualidade do modelo (R=0,913→R=0,905)

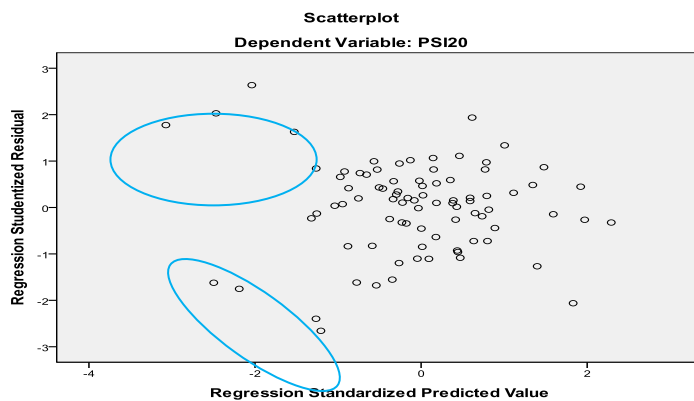
Model Summary ^e										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.808 ^a	.653	.647	102.89742	.653	122.079	1	65	.000	1.590
2	.861 ^b	.742	.734	89.41240	.089	22.085	1	64	.000	
3	.892 ^c	.796	.786	80.14652	.054	16.654	1	63	.000	
4	.905 ^d	.819	.807	76.05263	.023	7.965	1	62	.006	

d. Predictors: (Constant), SONAECOM, BRISA, BCP, GALP

e. Dependent Variable: PSI20

Este modelo passou em todos os testes incluindo o teste a multicolinearidade pelo que continuará a ser alvo dos testes de validação.

Validação do modelo: Homocedasticidade



Mais uma vez não se observa nenhum padrão, pelo que se conclui que a hipótese de **homocedasticidade** não é violada. Também neste modelo parecem surgir *ouliers*.

Validação do modelo: Teste à normalidade dos resíduos

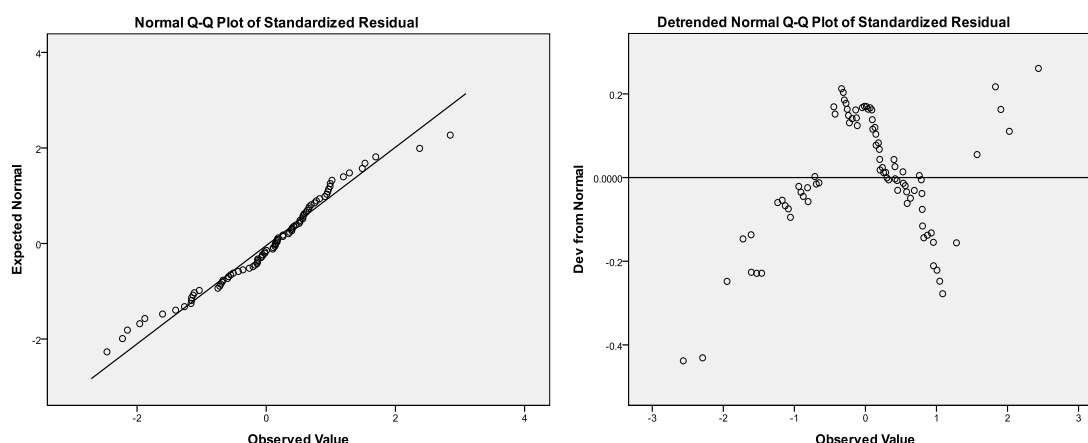
Pretende-se que os erros sigam uma distribuição normal, para comprovar essa hipótese efectua-se os seguintes testes à normalidade.

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.084	85	.200*	.983	85	.330

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

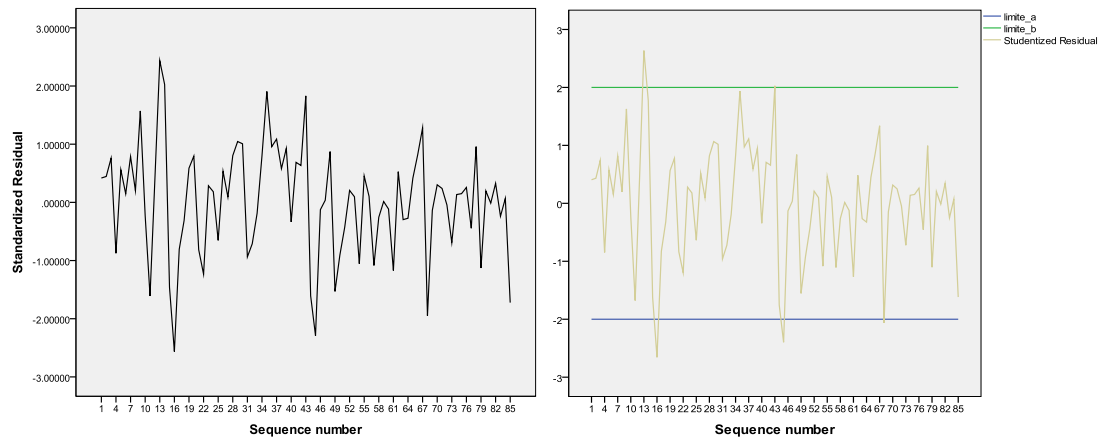
Em ambos os testes as significâncias são superiores a $\alpha=0,05$ pelo que se deverá assumir que os resíduos seguem uma distribuição normal.



A tendência dos resíduos segue as linhas de referência pelo que se confirma que os resíduos seguem uma distribuição normal.

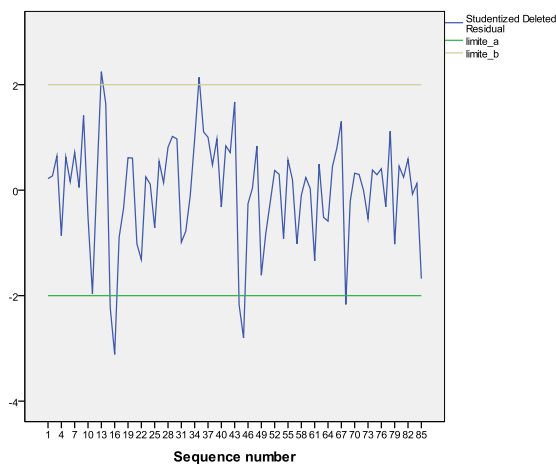
Análise de outliers

Tendo o modelo passado a todos os testes de validações próprios de uma regressão linear, é necessário proceder a um refinamento de modelo. Para tal devem ser identificados todas as observações que induzem erro no modelo. Deve-se proceder à identificação de *outliers* e casos influentes:



Observação	ZRE_1		SRE_1	
	$-3 < ZRE1 < 3$		$-2 < ZRE1 < 2$	
13	2,43515	-	2,635676	outlier
16	-2,56372	-	-2,65701	outlier
43	1,830568	-	2,028216	outlier
45	-2,29113	-	-2,39738	outlier
68	-1,94744	-	-2,06241	outlier

Podemos então identificar os seguintes *outliers*: 13, 16, 43, 45 e 68.



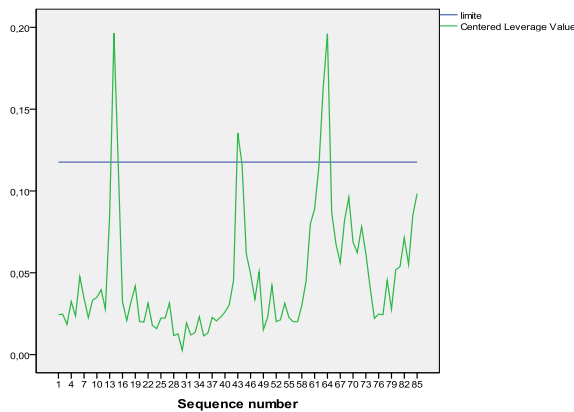
Identificação dos valores que influenciam o modelo estimado.

Obs	SDR_1	$-2 < ZRE1 < 2$
13	2,774378	Influente
16	-2,79971	Influente
43	2,08205	Influente
45	-2,49649	Influente
68	-2,11971	Influente

Análise de casos influentes

Identificados os *outliers* é necessário identificar os casos que influenciam a qualidade do modelo. Para tal são utilizados os seguintes parâmetros de teste.

Casos influentes: alavancagem:



Verificação a influência de uma observação na qualidade de ajustamento:

$$Lev_1 < \frac{2(k+1)}{n} < 0,118 \quad \begin{matrix} k = 4 \\ n = 85 \end{matrix}$$

Todos os valores que ultrapassarem o limite Lev1 influenciam a qualidade do ajustamento

Casos influentes: valor estimado ajustado

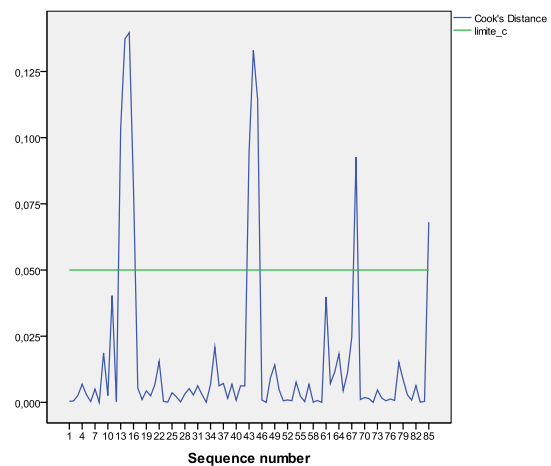


Pode-se observar um bom ajustamento entre as previsões ajustadas e previsões não tratadas.

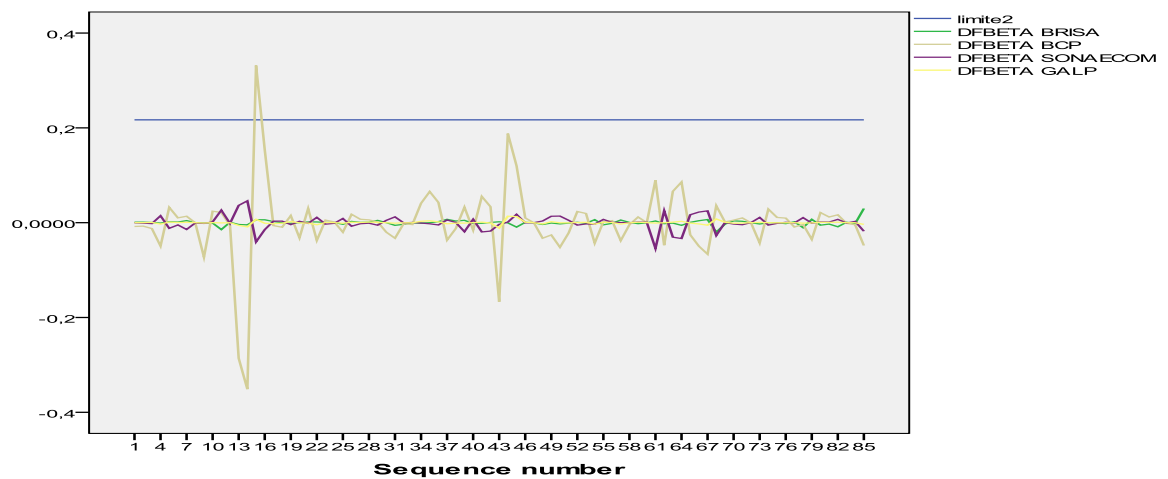
Distância de Cook

A distância de Cook identifica as observações que influenciam os coeficientes de regressão.

$$Dist. Cook > \frac{4}{n - k - 1}$$

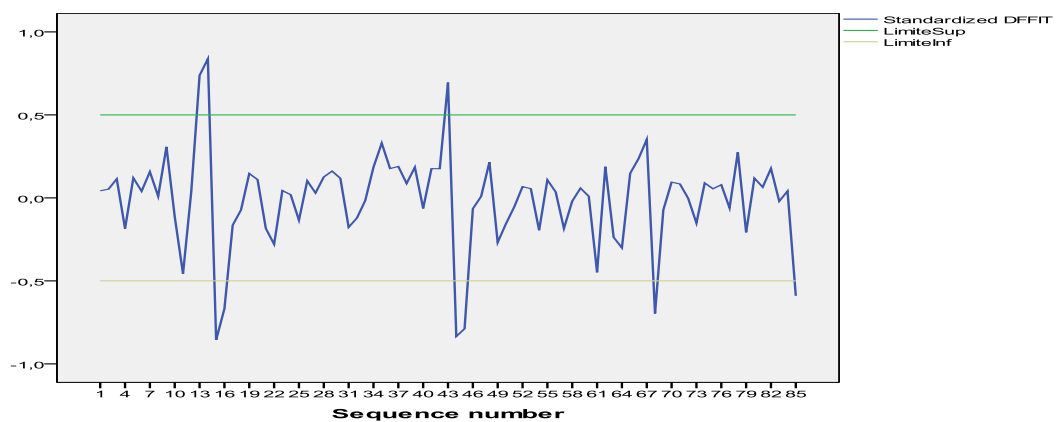


Casos influentes: DfBeta



Os casos identificados neste teste permitem identificar as observações que influenciam a estimativa do respectivo parâmetro em casa.

Casos influentes: DfFit



Esta medida, tal como a distância de *Cook*, permite medir o efeito de cada observação no ajustamento global do modelo.

Resumo: Oultiers e casos influentes:

Na tabela seguinte estão expostos todos os *outliers* e casos influentes existentes no modelo.

Obs	LEV_1	COO_1	Brisa SDB8_1	Sonae SDB11_1	Galp SDB14_1	BCP SDB17_1	SDF_1
9						-0,221	
11			-0,283				-0,501
13	0,131	0,238		0,411	-0,411	-0,923	1,149
14	0,217	0,204		0,599	-0,734	-1,290	1,319
15	0,174	0,123		-0,275	0,310	0,640	-0,795
16		0,105				0,391	-0,762
22					-0,233		
43	0,170	0,187			-0,739	-0,504	0,993
44	0,146	0,118			0,586	0,378	-0,780
45		0,109			0,576	0,265	-0,769
61	0,131	0,055		-0,431		0,252	-0,525
62	0,142			0,241			
63	0,188						
64	0,276						
65	0,121						
67				0,234	-0,278		
68		0,103	-0,339	-0,259	0,371		-0,739
85		0,075	0,677				-0,690

Modelo Final

Estimação de um novo modelo refinado

Conhecidas as observações que induzem erro no modelo, retiram-se as mesmas observações e calcula-se um novo modelo.

Model Summary ^d										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
3	,901 ^c	,813	,801	52,41791	,034	9,201	1	50	,004	1,746

c. Predictors: (Constant), SONAECOM, BRISA, BCP

d. Dependent Variable: PSI20

A primeira observação a fazer é exclusão da Galp do modelo. Tal facto dever-se-á, apesar dos indicadores “R” terem uma evolução negativa mínima, ao erro padrão da estimação pois este diminui de 76,05 para 53,41.

Este modelo permite afirmar que 81,3% da variação do PSI 20 pode ser explicada pela variação das cotações da SONAE.COM, Brisa e BCP.

	Modelo inicial	Modelo refinado	Efeito
R	0,905	0,901	Ligeira diminuição dos indicadores da qualidade do ajustamento, o que traduz uma evolução negativa do modelo.
R Square	0,819	0,813	
Adjusted R Square	0,807	0,801	
Std. Error of the Estimate	76,05	52,41	Diminui a dispersão dos resíduos em volta da recta, valorizando assim o modelo.

Modelo de previsão

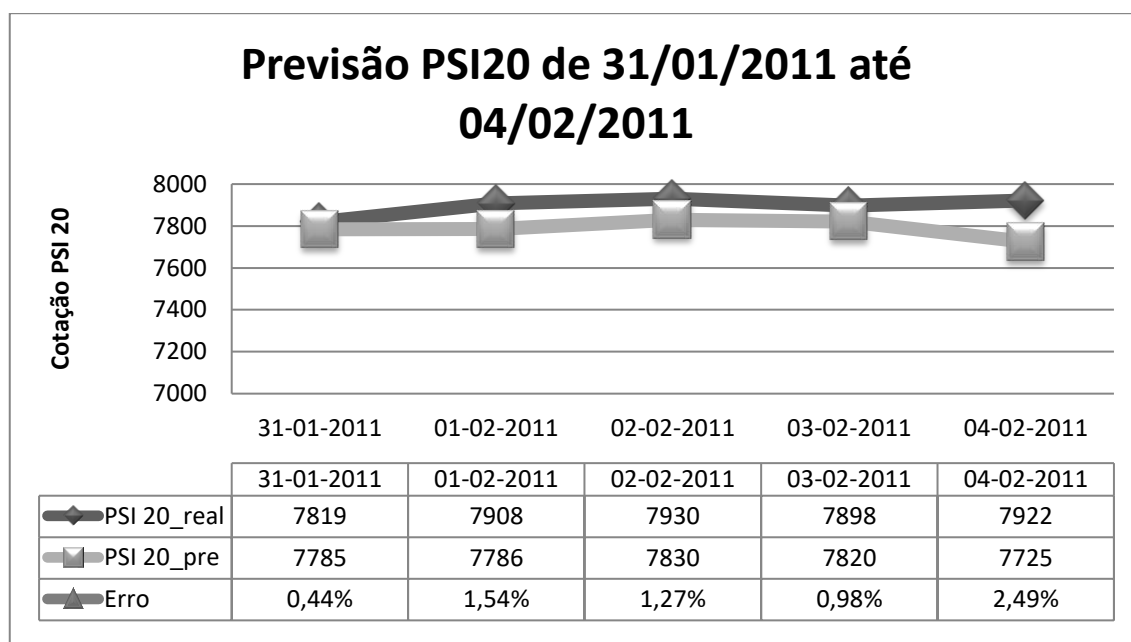
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
3 (Constant)	3453,911	318,032		10,860	,000					
SONAE.COM	1,016	,174	,482	5,842	,000	,713	,637	,358	,550	1,819
BRISA	,401	,045	,564	8,932	,000	,603	,784	,547	,939	1,065
BCP	1,312	,433	,249	3,033	,004	,515	,394	,186	,555	1,803

a. Dependent Variable: PSI20

Consultando a tabela dos coeficientes da regressão é possível escrever a equação da recta de previsão:

$$\text{PSI20} = 3453,911 + 1016 * \text{SONAE.COM} + 401 * \text{BRISA} + 1312 * \text{BCP}$$

Utilizando os dados dos primeiros dias de 2011 para as variáveis escolhidas podemos fazer uma previsão. Utilizando os valores reais para BCP, BRISA e SONAE é possível verificar a correspondência entre previsão e valor real.



Conclusões

O objectivo inicial deste trabalho era criar um modelo de previsão da evolução da cotação base do Índice PSI 20 para o dia seguinte, tendo em conta:

A primeira conclusão a retirar é que a influência dos mercados estrangeiros não é tão determinante no desempenho do PSI 20 como seria de esperar. Seria de prever que os índices das bolsas espanhola, principalmente e da bolsa alemã influenciassem as cotações da bolsa nacional, já que a economia portuguesa é fortemente influenciada pela economia espanhola. Apenas num momento variáveis de um índice que espelhe o comportamento dos mercados internacionais entraram no modelo. O índice Indiano foi removido por se considerar que não seria de grande racionalidade incluir esta variável no modelo. O brasileiro apesar de em teoria fazer sentido, foi removido, pois foi constatado o baixo impacto dos índices estrangeiros nos modelos ensaiados e foi decidido basear o modelo de previsão no desempenho das empresas nacionais.

Uma segunda conclusão que se pode retirar, é que cada uma das variáveis finais representa um sector de actividade de grande importância na nossa economia. A Banca está

representada pelo BCP, o sector de energia pela Galp, novas tecnologias pela SONAE.COM. Se considerarmos as primeiras variáveis escolhidas podemos ver que também o sector da distribuição e comércio representado pela Jerónimo Martins ou o sector de comunicação e novas tecnologias pela PT.

Por fim é possível verificar que erro associado a uma previsão baseada na equação:

$$\mathbf{PSI20 = 3453,911 + 1016 * SONAE.COM + 401 * BRISA + 1312 * BCP}$$

Apresenta um valor médio 2% para uma extrapolação para os primeiros 5 dias. Esta situação confirma que mesmo com modelos com grande precisão matemática, o erro associado à previsão nunca é eliminado.

