



Universidade de Brasília – UnB  
Departamento de Economia  
Ciências Econômicas

Guilherme da Cruz Souza

ARBITRAGEM ESTATÍSTICA UTILIZANDO UM MODELO DE  
COINTEGRAÇÃO

BRASÍLIA - DF

2017



GUILHERME DA CRUZ SOUZA

ARBITRAGEM ESTATÍSTICA UTILIZANDO UM MODELO DE  
COINTEGRAÇÃO

Monografia de conclusão de curso  
apresentada ao curso de Ciências  
Econômicas da Universidade de  
Brasília

Orientador: Prof. Dr. Daniel Oliveira Cajueiro

BRASÍLIA

2017



## **AGRADECIMENTOS**

Agradeço ao professor Cajueiro pelo suporte e comentários construtivos, ao professor José Guilherme pelo conhecimento em teoria das finanças, à minha família que sempre me apoiou. Agradeço também aos meus amigos que me ajudaram a me adaptar nesses anos morando longe da minha família e à minha antiga empresa pelos conhecimentos práticos em finanças e na vida.



## RESUMO

O presente trabalho busca testar um modelo de cointegração para identificar pares de ações do IBrX 100 e realizar operações de pair trading. É feito o teste de Engle Granger em todos os pares possíveis, em seguida eles são reanqueados pelo método da distância de *pair trading*. São montadas três carteiras diferentes a partir daí. Foram testados 2 modelos, um com parâmetros estáticos e outro com parâmetros dinâmicos, tendo esse segundo demonstrado resultados melhores. Foi feito o *backtest* e o modelo se provou promissor.

Palavras-chave: Arbitragem, cointegração.





## **ABSTRACT**

The present work tests a cointegration model to identify stock pairs from IBrX 100 and make pair trading operations. The Engle Granger test is applied on all possible pairs, next they are ranked based on the pair trading distance method. Three different portfolios are build. Two different models are tested, one with static parameters and another one with dynamic parameters, with the last one showing better results. The backtest was made and the model has proven to be promising.

Key-words: Arbitrage, cointegration.



## Sumário

Introdução.....	13
Revisão de literatura .....	15
Conceitos.....	19
3.1) Cointegração .....	19
3.2) Processo estacionário .....	21
3.3) Método da distância .....	22
3.4) Arbitragem estatística.....	22
Dados.....	24
Métodos .....	25
Resultados.....	31
Conclusão .....	43
Referências.....	44



## Introdução

---

O objetivo do trabalho será testar um modelo de cointegração para achar pares de ações na Bovespa. Depois de selecionados os pares, será testada uma estratégia *long short*, onde a ação subvalorizada é comprada e a ação sobrevalorizada é vendida. A ideia do modelo é que alguns pares de ações teriam um *spread* estacionário, de forma que ele retornaria à média. Isso ocorre pois o preço relativo entre os dois ativos não se alteraria em um prazo curto a menos que houvesse alterações intrínsecas em uma das ações. Por exemplo, pares de ações de um mesmo setor. As duas ações são de uma mesma área, então a princípio não se espera que uma ação se valorize à uma taxa muito superior à outra. Quando uma delas se valoriza muito (por causa de um grande player do mercado que fez uma compra grande por exemplo) e o *spread* entre as duas aumenta, em um segundo momento haveria uma oportunidade de arbitragem afim que os preços relativos voltassem ao seu nível padrão.

A estratégia de *pair trading* vem sendo usada desde a década de 80, e tem a característica de ser neutra em relação ao mercado, servindo como um possível mecanismo de diversificação. Em 1987 Nunzio Tartaglia trabalhando no Morgan Stanley formou um grupo para desenvolver estratégias de arbitragem quantitativas. Essa estratégia se baseava em análise quantitativa dos dados históricos e os *trades* eram realizados automaticamente.

Um dos fundos quantitativos mais famosos do mundo hoje é o *Medallion Fund*. É um fundo quantitativo onde um terço de seus funcionários possuem PhD's e suas estratégias são baseadas em análises de dados. Entre os anos de 2009 e 2015 o retorno médio do fundo foi superior a 30% ao ano, e a última vez que o fundo teve um retorno anual negativo foi em 1989. O setor de fundos quantitativos no Brasil ainda é muito inexplorado, sendo possível perceber isso ao analisar a quantidade de recursos alocados no mesmo. Em 2016 o patrimônio investido em fundos ultrapassava os R\$ 3,4 trilhões (o décimo maior no mundo) de acordo com o anuário da FGV de 2017. No mesmo ano, um artigo da Valor Econômico indicou que o patrimônio alocado em fundos quantitativos havia batido a marca de R\$ 1 Bilhão. Em termos relativos isso é

menos de 0,1% do mercado de fundos. Nos Estados Unidos, com um patrimônio de US\$ 3,1 Trilhão (R\$10,2 trilhão com o câmbio à 3,30), os fundos quantitativos possuem US\$ 930 bilhões, ou 30% do mercado de fundos, de acordo com o site *Business Insider*. Esses números indicam que pode haver espaço a ser explorado por estratégias quantitativas no Brasil.

Soma-se a esse cenário de terreno inexplorado o atual momento da economia brasileira, onde as taxas de juros vem caindo de forma que a rentabilidade dos títulos de renda fixa, tanto privados quanto públicos, deixa de ser tão competitiva. Assim o custo de oportunidade para investir na bolsa se reduziu. Ao mesmo tempo o cenário político se encontra em um momento de instabilidade e com algum grau de incerteza, o que pode tornar a estratégia mais interessante uma vez que ela se propõe a ser neutra em relação ao mercado, ou seja, os eventos políticos que geram volatilidade na bolsa não devem ter uma grande influência sobre a estratégia.

Para montar o modelo serão testados 4 anos de dados da bolsa, particularmente de ações do IBrX 100, um índice composto pelas 100 empresas de maior negociabilidade na Bovespa. É importante usar ativos que possuam liquidez pois a falta da mesma aumentaria o risco da estratégia e o seu custo. O primeiro passo será utilizar o primeiro ano dos dados para estimar e testar os pares. Serão encontrados diversos possíveis pares entre as 100 ações, assim os 12 primeiros meses servem para selecionar os 20 melhores entre eles. Com a análise feita se inicia o período de *trade* nos 6 meses seguintes. Ao fim desse período se reinicia a estimação com os 12 meses anteriores e assim por diante. Também foi testado um modelo com parâmetros dinâmicos onde não se tem a divisão dos períodos e eles são atualizados diariamente.

Para achar os pares temos que testar as ações para raiz unitária, pois para haver cointegração necessariamente ambas as ações tem que possuir raiz unitária. Em seguida fazemos o teste de cointegração entre todas as ações restantes. Todos os pares que forem cointegrados são candidatos ao *trade*, então será necessário mais critérios para afunilar e selecionar apenas os melhores.

## Revisão de literatura

---

Vidyamurthy (2004) apresenta a ideia de cointegração entre pares. Com a ideia de vender o ativo sobrevalorizado e comprar o subvalorizado, ele argumenta que apenas seria possível fazer isso sabendo o valor real do mesmo, o que seria muito difícil de se conseguir. A ideia de pair trading então seria tentar fazer o mesmo através da precificação relativa de dois objetos similares. Com isso se eliminaria a necessidade de saber o preço específico. O spread entre os dois ativos passa a ser a informação de maior importância. Quando o spread se afastar demais de sua média histórica haveria uma indicação de um ativo mal precificado. Pelo fato de se ter uma posição comprada e uma vendida de dois ativos similares, a operação é montada afim de ter um beta próximo de zero, o que colocaria a estratégia como neutra ao mercado. O critério utilizado para formar esses pares de ativos seria o de cointegração.

Em seu artigo, Caldeira (2013) analisa uma estratégia considerada neutra em relação ao mercado, o que significa um modelo de cointegração beta neutro. Ele testa os 50 ativos de maior participação do Ibovespa, formando pares e utilizando o teste Dickey-Fuller aumentado e o teste de Johansen. Após achar os pares cointegrado é feito um ranqueamento através do Índice de Sharpe do período de análise. São selecionados os 20 primeiros pares sem que haja restrição para ativos repetidos. Dados de cinco anos entre janeiro de 2005 e dezembro de 2009 são utilizados, sendo o primeiro ano utilizado como período de análise, de forma que não ocorre nenhum *trade* ali. A cada quadrimestre os parâmetros são reestimados. Não são feitas restrições com relação ao setor das empresas ao formar os pares, de forma que duas empresas de setores distintos podem ficar juntas. Os critérios de *trade* são baseados em uma banda de dois desvios padrões para abrir a operação e meio desvio padrão para finalizar a operação. Outro critério para fechar uma operação é a meia vida do par. Ele é calculado no período de análise, e se dentro do período de meia vida a operação não tiver terminado ainda, ela é finalizada. Os resultados apresentados pelo modelo são positivos, com uma

rentabilidade acumulada de 89,78% no período total e uma rentabilidade anual média de 17,37%. Além disso foi apresentada uma correlação baixa com o mercado, apresentando um coeficiente de correlação de 0,079.

Brito e Pinto (2012) apresentam um modelo de pairs trading com dados entre 2006 e 2010 com preços intra-diários (a cada 15 minutos) das ações do Ibovespa. O critério usado para selecionar os pares é a cointegração e os *trades* são automatizados. O período de análise no modelo é de 10 semanas e é utilizado o teste Dickey-Fuller aumentado para selecionar o subconjunto de pares cointegrados. Aqui as ações analisadas foram as que compunham o Ibovespa em março de 2011, o que poderia gerar um viés de sobrevivência ao tentar utilizar a estratégia em uma situação real. O período de *trade* utilizado aqui é de uma semana, depois desse período os parâmetros são recalculados com as últimas 10 semanas. Para cada ano são descartados as duas últimas semanas de *trade*. O teste de cointegração é rodado com nível de confiabilidade de 99%. O máximo de pares utilizados é de 100 e não há restrição com relação à repetição de ativos nos pares. Após os pares estarem prontos é feita a regressão entre os ativos para se obter o *hedge ratio* e os resíduos utilizados para obter a média e as bandas superiores e inferiores onde os *trades* irão ocorrer. As bandas utilizadas estão a dois desvios padrões da média. O ponto de saída do *trade* ocorre a meio desvio padrão da média. Também é usado um sistema de *stop loss* quando o resíduo atinge três desvios padrões. O custo de abrir as posições não é zero, sendo utilizada uma estratégia beta neutro de forma a reduzir o risco direcional. Os resultados obtidos não demonstram muita consistência do modelo. Dos cinco anos onde o modelo foi utilizado, dois tiveram retornos negativos, um ano retorno próximo de zero e dois anos com retorno positivo.

Yu e Lu (2017) argumentam que o modelo tradicional de cointegração com dois ativos ainda podem ser considerados arriscados demais e que utilizar cestas de ativos ajudaria a reduzir ainda mais o risco da operação. O trabalho então se propõe a identificar cestas de ativos que apresentam co-movimento. Os testes são feitos em ações da bolsa de Hong Kong, e são usadas apenas do setor financeiro e comercial, que representam 75% do peso do Hang Seng Index (HSI), o índice das 50 maiores empresas. São usados dois anos de



dados para o período de análise entre 2014 e 2016. O período de *trade* é de 25 dias, e o tempo máximo de operação é de 50 dias. Ao testar uma cesta com 3 ativos os resultados foram positivos, particularmente em períodos de baixa do mercado.

Elliot et al (2005) testa um modelo de spread com reversão à média. Assim a ideia era que dois ativos similares teriam um spread com um valor de equilíbrio. Então quando ele aumenta o modelo prevê uma convergência num segundo momento, e quando ele diminui o oposto ocorre.

Gatev et al (2006) aplica um modelo de pairs trading no período de 1962 a 1997. O método usado é replicado nesse trabalho. São utilizados 12 meses para formar os pares e os 6 meses seguintes para realizar as operações. Para selecionar os pares é utilizado o método da distância, onde os melhores pares são os que minimizam a soma do quadrado da diferença dos dois preços normalizados. Isso significa selecionar os pares que possuem preços que se movem relativamente juntos. Quando se inicia o período de *trade*, o critério para abrir uma posição é quando os preços divergem por uma distância de dois desvios padrões da média. A operação só é finalizada quando os preços se cruzam ou quando os 6 meses acabam.

Avellaneda e Lee (2008) testam uma estratégia de arbitragem estatística utilizando ativos americanos. São utilizados modelos de reversão à média e backtest dos dados entre os anos de 1997 e 2007. Nesse trabalho não são usadas apenas ações, mas também ETF's. Ao testar para cointegração uma ação e uma ETF do setor da ação escolhida, seria possível testar se ela está subvalorizada ou sobrevalorizada em relação à outras empresas do seu setor. A janela para estimar os parâmetros é de 60 dias. O ponto de entrada é 1,25 desvio padrão da média e o ponto de saída a 0,5 desvio padrão da média. É feito o rebalanceamento da carteira diariamente e é assumido um custo de 10 pontos base sobre o valor da operação. O Índice de Sharpe no período foi 1,1.

Do et al (2006) analisa diferentes abordagens de pairs trading. A estratégia explora a má precificação de ativos, e esse artigo apresenta três modelos que podem ser usados para esse fim. O método da distância, a cointegração e o método do spread estocástico. O primeiro visa identificar

pares que se movem juntos utilizando a distância, que seria o quadrado da diferença dos dois preços normalizados. O *trade* seria iniciado quando essa distância se afastasse da média definida no período de análise. Uma forma de selecionar pares aqui seria escolher os que apresentam as menores distâncias. O *trade* se iniciaria a dois desvios padrões da média e poderia haver um *stop loss*. Essa estratégia assume que a distância entre os preços se manterá ao longo do tempo. O segundo método apresentado é o de cointegração, onde duas séries temporais com raiz unitária podem ser combinadas para gerar uma série temporal estacionária. Ao utilizar a cointegração em pares de ações é possível obter uma combinação que gere um spread estacionário, o que geraria oportunidades de arbitragem quando o mesmo se afasta do seu valor de equilíbrio. O trabalho aponta possíveis falhas para esse modelo, como a ordem das variáveis ao fazer o teste de cointegração, já que essa informação muda o resultado do modelo, e se o par não for cointegrado, o teste irá gerar um resultado espúrio. O último método apresentado é o de spread estocástico. Aqui o spread é a diferença simples entre o log do preço de dois ativos. O modelo assume que esse spread é estacionário. Isso seria um problema pois restringiria muito os pares no longo prazo já que os retornos deveriam ser idênticos após um período grande.

## Conceitos

---

### 3.1) Cointegração

A formalização do método de cointegração ocorreu com Engle e Granger (1987). Em seu artigo eles demonstraram que duas séries temporais não estacionárias podem ter uma combinação linear que seja estacionária. De maneira mais formal temos que se  $\{y_t: t = 1, 2, \dots\}$  e  $\{x_t: t = 1, 2, \dots\}$  forem  $I(1)$ , de forma geral teríamos  $P_{S1,t} - \beta P_{S2,t}$  como um processo  $I(1)$ . Mas se para um dado  $\beta \neq 0$ ,  $P_{S1,t} - \beta P_{S2,t}$  for  $I(0)$ , isso significará que ele terá média e variância constante. Se esse  $\beta$  existir, dizemos que esse par de variáveis é cointegrado.

$$P_{S1,t} - \beta P_{S2,t} = \mu + \epsilon_t$$

Sendo P o preço, S1 e S2 os ativos 1 e 2 respectivamente,  $\epsilon$  o erro de cointegração e  $\mu$  a média de cointegração. Se o par de ações for cointegrado o spread entre as duas deve se manter em torno da média. Dessa forma, quando o spread se afasta da média, o modelo compra(vende) 1 unidade do ativo S1 e vende(compra)  $\beta$  unidades do ativo S2. Quando ocorre o retorno à média a operação é finalizada. Essa convergência à média esperada pode produzir um retorno positivo. Isso significa que se um par de ativos for cointegrado,  $P_{S1,t} - \beta P_{S2,t}$  será estacionário. A seguir, um exemplo com um par de ações cointegradas da Bovespa.

Gráfico 1 - Preço das ações ITUB4 e ITSA4

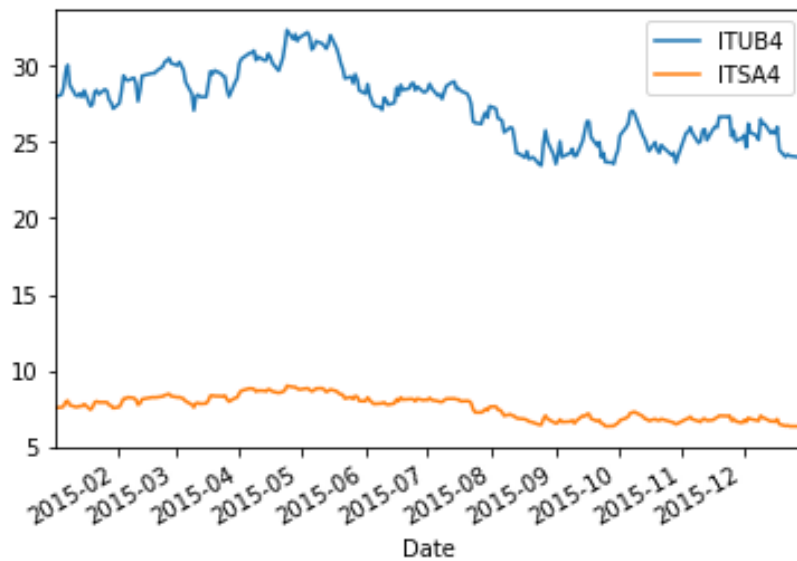
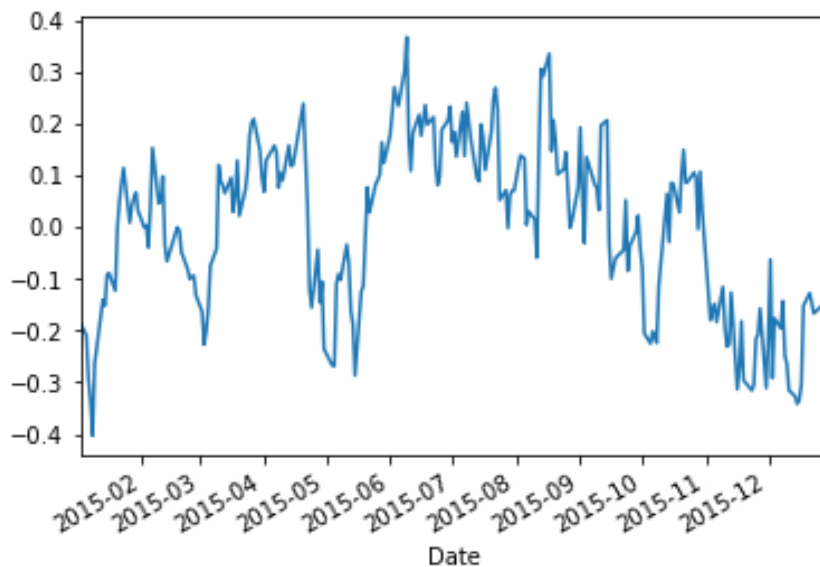


Gráfico 2 – Resíduo da regressão de ITUB4 e ITSA4



O gráfico 1 mostra o preço das ações ITSA4 e ITUB4, enquanto o gráfico 2 mostra o resíduo da regressão dos ativos que é estacionário.

Uma forma de testar se um par de ativos é cointegrado consiste em duas etapas. Primeiro é feita uma regressão para estimar a relação linear do par, em seguida se testa o resíduo para estacionariedade. Nesse trabalho foi utilizado o teste Dickey Fuller aumentado.

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + \epsilon_t$$

Sendo  $\beta_1$  o intercepto,  $\beta_2$  o coeficiente de tendência,  $\delta$  o coeficiente de raiz unitária e  $m$  o número de defasagens tomadas na série. A hipótese nula é de que  $\delta = 0$ , ou seja, o resíduo não é estacionário. Ao rodarmos a regressão podemos calcular a estatística  $T$  para testar a hipótese.

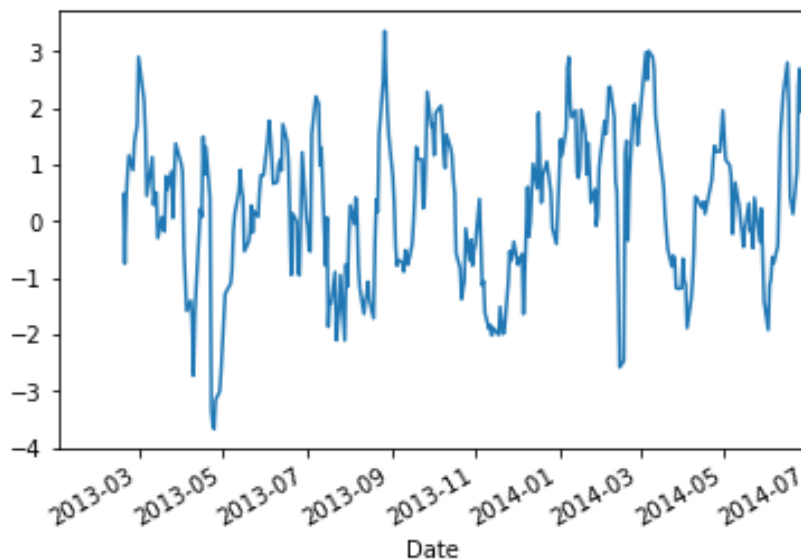
$$T = \frac{\hat{\delta}}{se(\hat{\delta})}$$

Onde  $\hat{\delta}$  é um estimador para  $\delta$ , e  $se(\hat{\delta})$  é um estimador para o desvio padrão do erro de  $\delta$ .

### 3.2) Processo estacionário

De acordo com Wooldridge (2011), um processo estacionário é aquele em que as distribuições de probabilidades são estáveis no decorrer do tempo. Pela definição formal, o processo  $\{x_t: t = 1, 2, \dots\}$  é estacionário se para todas as coleções de índices temporais  $1 \leq t_1 < t_2 < \dots < t_m$ , a distribuição conjunta de  $(x_{t_1}, x_{t_2}, \dots, x_{t_m})$  é a mesma que a distribuição conjunta  $(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_m+h})$  para todos os inteiros  $h \geq 1$ . A seguir, uma representação gráfica de uma série estacionária.

Gráfico 3 – Série estacionária gerada de maneira aleatória



### 3.3) Método da distância

Consiste em um método usado para analisar 2 ativos que tenham se movido juntos em um dado período de tempo. Em um primeiro momento se normaliza os preços e em seguida se calcula o somatório do quadrado da diferença.

$$\sum_{t=1}^N (P_{t1} - P_{t2})^2$$

Esse método será utilizado para ranquear os pares após o teste de cointegração afim de achar os pares de ativos que variam de maneira similar.

### 3.4) Arbitragem estatística

Arbitragem é uma operação que extrai lucros de uma falha de mercado. Quando dois ativos com características idênticas ou similares apresentam preços muito diferentes pode haver uma oportunidade de arbitragem. Esse *trade* consiste em uma compra e uma venda simultânea, de forma que ele se autofinancia. A arbitragem é um mecanismo que garante que os preços não vão se desviar muito no longo prazo. O primeiro uso desse método é atribuído à Nunzio Tartaglia, que na década de 80 reuniu uma equipe de matemáticos, físicos e cientistas da computação, afim de desenvolver estratégias quantitativas no banco Morgan Stanley. Ali começou o uso de pares de ações, de forma que o modelo identificava anomalias na relação entre os dois ativos e fazia o *trade* com a ideia que ela se corrigiria.

O uso de pares de ações para fazer arbitragem é chamada de pair trading. Ele consiste na ideia simples de comprar o ativo subvalorizado e vender o ativo sobrevalorizado. O modo de saber se um ativo está caro ou barato é através de uma avaliação do preço relativo. Dois ativos com características similares devem ter preços que se movem relativamente juntos. Se um dos dois se valorizar demais em relação ao outro sem que tenha ocorrido alguma mudança fundamental na empresa pode haver uma oportunidade de arbitragem. Uma característica interessante desse tipo de

operação é o fato de que a posição terá um beta próximo de neutro, o que significa que o retorno da carteira terá baixa correlação com o retorno de mercado, sendo assim considerada uma estratégia neutra.

Em seu artigo, Gatev et al. (2006) testa uma estratégia de pairs trading usando dados diários entre 1962 e 1997. É utilizado um período de 12 meses para análise dos dados seguido de um período de 6 meses de *trade*. Os pares são selecionados utilizando dois métodos, sendo o primeiro o de cointegração para determinar quais pares são elegíveis e depois minimizando a soma do quadrado do desvio padrão das duas séries de preços normalizadas para ranquear os pares.

## Dados

---

Todos os dados são extraídos do Google Finance. São usados os preços de fechamentos diários das ações do IBrX 100 de janeiro de 2013 até dezembro de 2016. O IBrX 100 é uma carteira teórica de ativos com as 100 ações de maior negociabilidade e representatividade do mercado de ações brasileiro. A carteira do índice tem vigência de 4 meses, sendo no final desse período rebalanceada.

Para rodar os dados foram utilizados pacotes de análise estatística no Python. Em seguida os resultados foram exportados para o Excel.



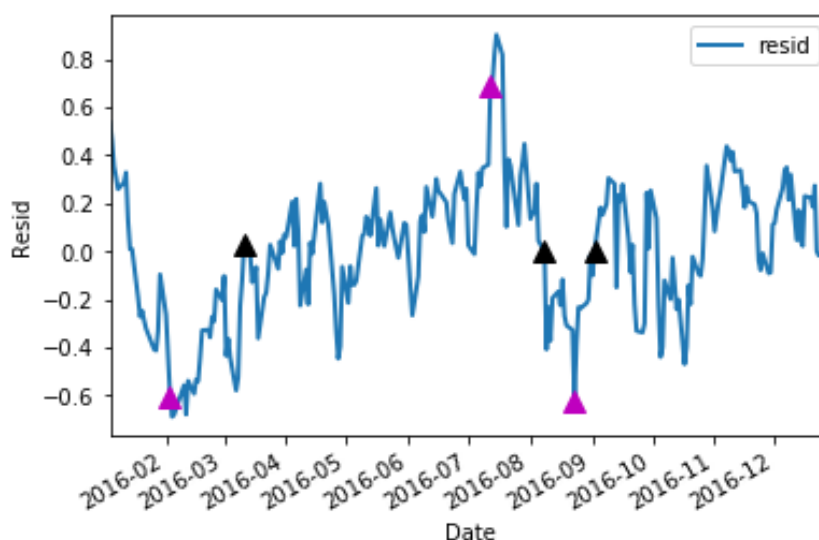
## Métodos

Tabela 1 - Divisão dos períodos analisados

	Ano 1	Ano 2	Ano 3	Ano 4			
Período 1	Análise	Trade					
Período 2		Análise	Trade				
Período 3			Análise	Trade			
Período 4				Análise	Trade		
Período 5					Análise	Trade	
Período 6						Análise	Trade

Após um período de 12 meses de análise há um período de 6 meses de *trade*, como apresentado na tabela 1. No período de análise são utilizados os dados de fechamentos diário das 100 ações do IBrX 100. Com os dados em mãos é feito o teste de cointegração em todos os pares possíveis. Em seguida esses pares são ranqueados através do método da distância. Os 15 primeiros pares são selecionados para o *trade*. Também são feitas outras duas carteiras, uma com empresas do mesmo setor e outra com ativos da mesma empresa. Com os dados do desvio padrão e *hedge ratio* coletados no período de análise se inicia o período de *trade*. É definida uma banda em volta do resíduo médio de dois desvios padrões para cima e dois para baixo. O *trade* se inicia quando o resíduo atravessa uma das duas bandas e é encerrado quando o resíduo retorna a uma distância de meio desvio padrão da média ou quando chega no *stop loss*, que fica a três desvios padrões da média. A seguir, um exemplo.

Gráfico 4 – Exemplo de *trade* baseado no resíduo de um par



A linha azul do gráfico 4 é o resíduo do período, os triângulos rosas são os pontos de entrada e os pretos são os pontos de saída.

No método descrito anteriormente os parâmetros utilizados são calculados nos 12 meses de análises e ficam fixos no período de *trade*. Assim o desvio padrão, o beta da regressão dos ativos e a média dos resíduos se mantêm por 6 meses.

A seguir, um exemplo completo utilizando o modelo de parâmetros estáticos. São usados dados do ano de 2015 para o período de estimação, entre 2 de janeiro e 30 de dezembro. O número de observações para cada ação é de 246. Os ativos utilizados são ITUB4 e ITSA4. O primeiro é do Banco Itaú e o segundo da Itaúsa, uma holding que controla, entre outros empreendimentos, o Itaú. O primeiro passo é testar se os dois ativos possuem raiz unitária. Nenhum dos dois ativos rejeita a hipótese nula de que as variáveis contêm raiz unitária. O próximo passo então é fazer o teste de cointegração. O teste rejeita a hipótese nula de que as duas ações não são cointegradas a 5%, o que indica que o resíduo é estacionário. A seguir, os gráficos 5 e 6 mostram o preço das duas ações e do seu resíduo no ano analisado, respectivamente.

Gráfico 5 - Preço das ações ITUB4 e ITSA4

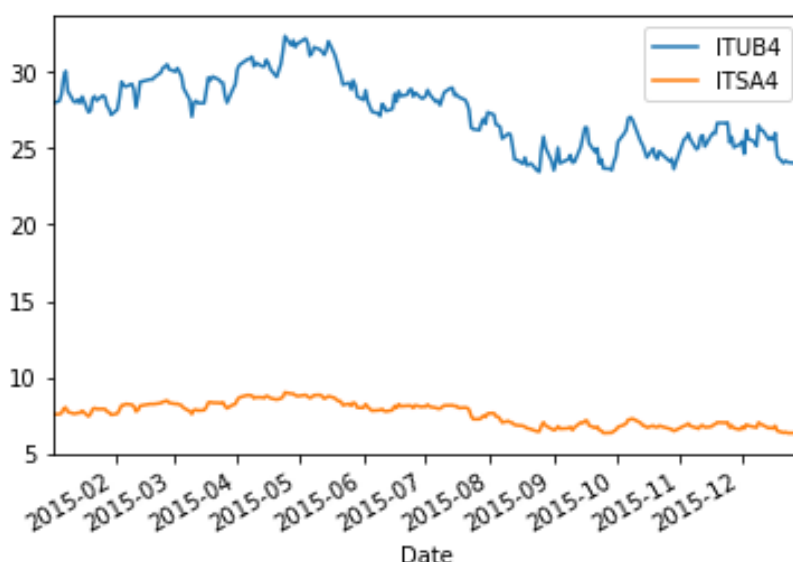
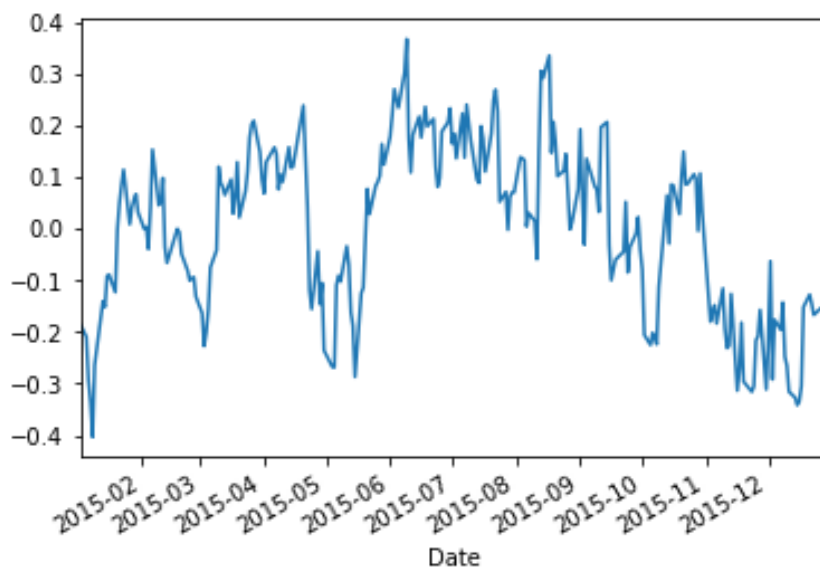
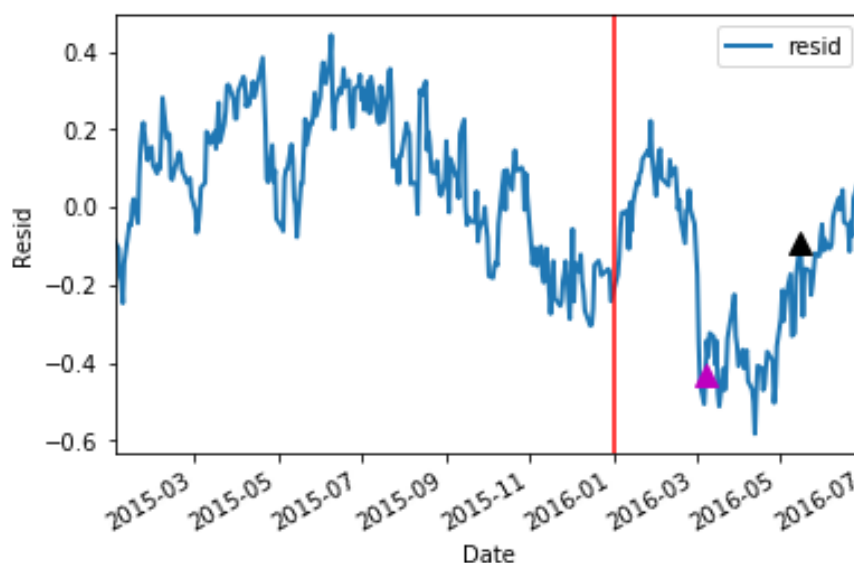


Gráfico 6 – Resíduo da regressão de ITUB4 e ITSA4



Agora que foi constatado que o par é elegível para o *trade*, se inicia a operação quando o resíduo superar dois desvios padrões de distância da média. A proporção entre os ativos é encontrada estimando a relação linear entre as duas séries de preços. O beta da regressão é o *hedge ratio*. No exemplo o *hedge ratio* é de 0.3198, o que significa que para cada ação da ITSA4 comprada será vendida 0.3198 de ITUB4. Nos gráficos 7 e 8 é possível ver a abertura e o fechamento do *trade*.

Gráfico 7 – Exemplo de *trade* com o gráfico de preços ITUB4 e ITSA4

Gráfico 8 – Exemplo de *trade* com o gráfico de resíduos de ITUB4 e ITSA4

Nesse exemplo foi feito apenas um *trade* no período analisado. Com uma rentabilidade de 1.86% e volatilidade de 0.5% em um período de 89 dias. Os ativos vendidos foram suficientes para financiar a operação, sobrando então os custos operacionais apenas. Nos gráficos 7 e 8 é possível visualizar o *trade* do ponto de vista dos preços e do resíduo.

Outro método foi utilizado afim de comparar os resultados. Foi também testado um modelo de parâmetros dinâmicos, de forma que o beta, o desvio padrão do resíduo e a média variam ao longo do tempo. Foi utilizada uma janela de 250 dias para cada um dos parâmetros, o que seria aproximadamente um ano de dados. Nesse teste a carteira de ativos se mantém a mesma do começo ao fim. Os primeiros 12 meses servem de estimação dos parâmetros e os 36 meses seguintes são um período de *trade* ininterrupto com os parâmetros sendo atualizados diariamente. Aqui também foram usados dados de fechamento diários das ações e as carteiras escolhidas foram as mesmas do primeiro período do modelo com parâmetros estáticos.

Outra diferença em relação ao modelo estático é que nesse foi utilizado o resíduo normalizado. Os pontos de entrada ficam entre 1,5 e 2,5 desvios padrões da média, o *stop loss* fica a 4 desvios padrões da média e a operação é finalizada quando o resíduo atravessa a média. Nesse meio tempo uma das empresas teve seu capital fechado e suas ações pararam de ser negociadas na bolsa. O par não é substituído e o capital direcionado à ele fica parado até o

final do período de *trade*. A seguir, um exemplo completo do segundo método. Será utilizado como exemplo o par ELET3 e ELET6. São calculados os parâmetros com os dados dos 250 dias anteriores. A seguir o gráfico 9 ilustra o preço dos ativos.

Gráfico 9 - Preço das ações ELET3 e ELET6

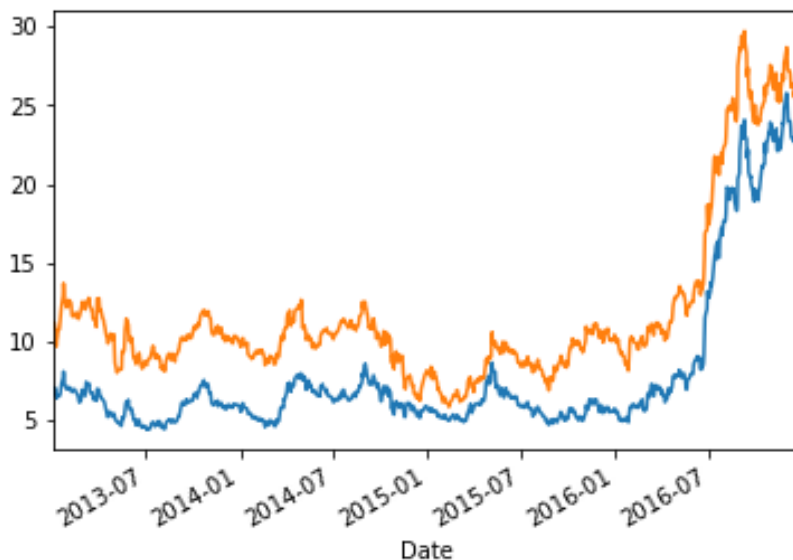
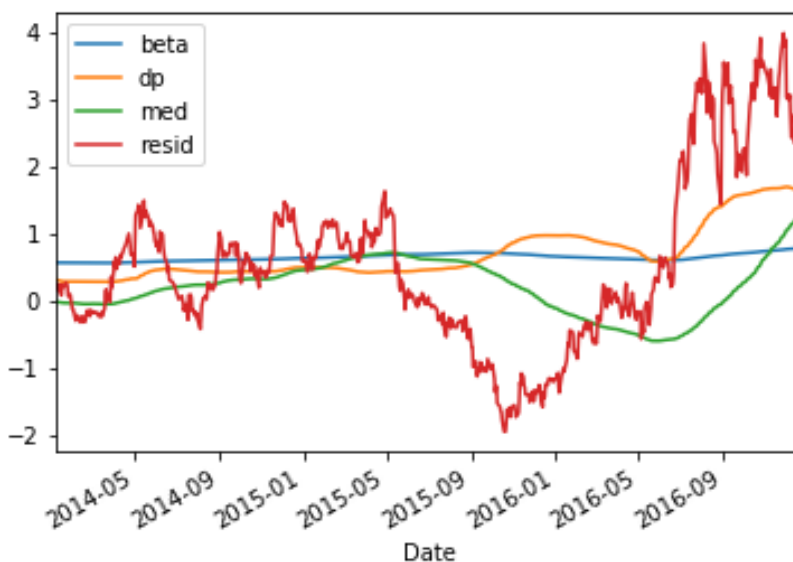
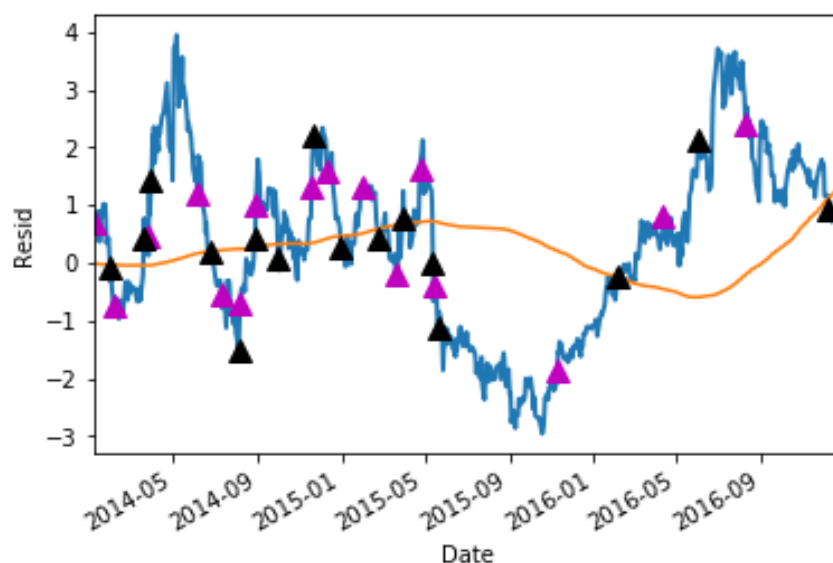


Gráfico 10 - Parâmetros calculados do par ELET3 e ELET6



Com os parâmetros em mãos é feita a normalização do resíduo e a partir daí são computados os sinais para o backtest. No gráfico 10 está exposta a variação dos parâmetros ao longo do período de análise e no gráfico 11 o *trade*. Nele se tem o resíduo em azul, a média em laranja, os pontos de entrada em rosa e os pontos de saída em preto.

Gráfico 11 – Exemplo de *trade* com o gráfico de resíduos de ELET3 e ELET6

Outra questão a ser levantada é a composição dos pares. Existem duas maneiras de balancear os pares. Uma é beta neutra, onde a composição é definida de acordo com o beta da regressão dos ativos. Nela se busca ter um par sem exposição direcional e o mais neutro possível. A outra é a carteira *cash* neutra. Nela se aloca metade em cada um dos ativos. A vantagem aqui é que a carteira terá custo zero, o que significa que o dinheiro obtido na ponta vendida é usado para financiar a ponta comprada. Assim, o dinheiro destinado ao par poderia ficar alocado em outros ativos de baixo risco. O problema aqui é que dessa maneira o par fica exposto ao risco direcional do mercado e deixa de ser beta neutro. No modelo dinâmico foram rodados tanto carteiras beta neutro quanto *cash* neutro afim de comparar se os retornos são similares e se o risco de ter uma carteira a custo zero é compensado pelo retorno gerado.

Com relação aos custos operacionais foi utilizado o valor de 0,5% do valor total do *trade* por operação. Isso inclui o custo de corretagem, de custódia e o aluguel das ações vendidas. É difícil ter um valor exato para essa variável já que esses custos podem variar dependendo da corretora, do momento do mercado e do ativo que está sendo operado. Ativos com pouca liquidez podem ter um spread bid/ask maior e um custo de aluguel mais elevado por exemplo.

## Resultados

---

No primeiro momento foram feitos testes de cointegração com os ativos do IBrX 100 e obtidos os pares que foram usados no modelo. A seguir, os pares do primeiro semestre de 2014 e que foram usados pelo modelo de cointegração dinâmica.

Tabela 2 – Pares da carteira de 15 ativos

('BVMF:GGBR4 - BVMF:GOAU4')
('BVMF:OIBR3 - BVMF:OIBR4')
('BVMF:CSMG3 - BVMF:SBSP3')
('BVMF:ENEV3 - BVMF:RADL3')
('BVMF:BRKM5 - BVMF:CIEL3')
('BVMF:BBDC4 - BVMF:BRML3')
('BVMF:LLXL3 - BVMF:PDGR3')
('BVMF:BRAP4 - BVMF:ELPL4')
('BVMF:ALLL3 - BVMF:VIVT4')
('BVMF:JBSS3 - BVMF:TIMP3')
('BVMF:ITUB4 - BVMF:BBAS3')
('BVMF:LIGT3 - BVMF:QUAL3')
('BVMF:EQTL3 - BVMF:PSSA3')
('BVMF:DTEX3 - BVMF:PCAR4')
('BVMF:ODPV3 - BVMF:MPLU3')

Os pares acima foram selecionados primeiro pelo critério de cointegração e depois ranqueados pelo método da distância. Aqui não há critério de restrição por setor ou qualquer outro parâmetro fundamentalista.

A segunda carteira foi selecionada através do critério de cointegração e depois por separações setoriais. Todos os pares são de empresas que atuam no mesmo setor ou que são ligadas de alguma forma (ações ON e PN de uma mesma empresa). Os pares também foram depois ranqueados pelo método da distância.

Tabela 3 – Pares da carteira setorial

('BVMF:GGBR4 - BVMF:GOAU4')
('BVMF:OIBR3 - BVMF:OIBR4')
('BVMF:CSMG3 - BVMF:SBSP3')
('BVMF:BISA3 - BVMF:BRPR3')
('BVMF:MULT3 - BVMF:BRML3')

A última carteira não passou por testes de cointegração e é composta apenas por ações da mesma empresa (ON e PN).

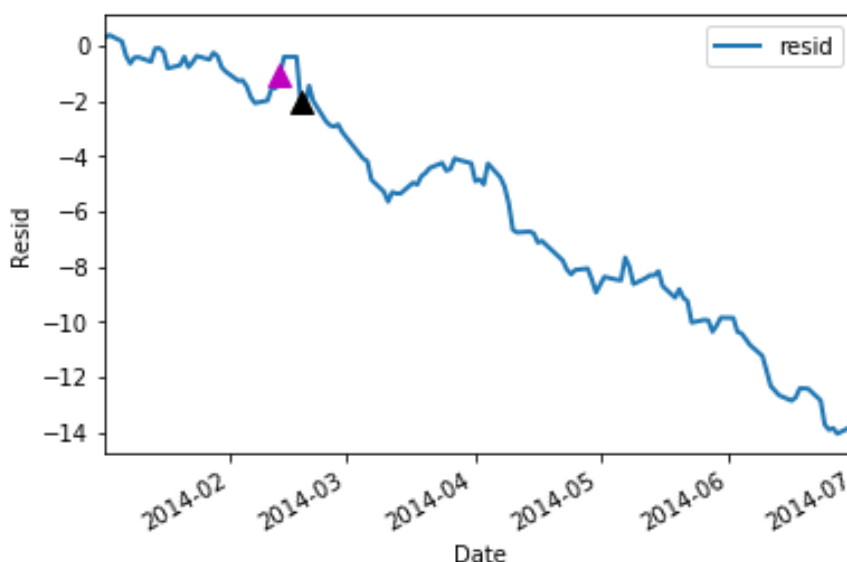
Tabela 4 – Pares da carteira ON e PN

('BVMF:OIBR3 - BVMF:OIBR4')
('BVMF:VALE3 - BVMF:VALE5')
('BVMF:PETR3 - BVMF:PETR4')
('BVMF:ELET3 - BVMF:ELET6')
('BVMF:BBDC3 - BVMF:BBDC4')

A primeira carteira possui 15 pares sem repetição de ativos, o que representa 30% dos ativos analisados. As outras duas carteiras possuem 5 pares que foram selecionados com critérios mais rígidos, enquanto a primeira possui pares de empresas aparentemente sem muita relação, como JBSS3 e TIMP3, uma do ramo de proteína animal e outra do setor de telecomunicações.

No primeiro momento foi rodado o modelo de parâmetros estáticos. Foi observado que de maneira muito comum os pares perdiam a relação de cointegração ao longo do período de *trade*, com o resíduo adquirindo uma tendência clara. Com isso muitos pares sequer foram operados. Um exemplo, a seguir, de um par que perdeu a relação.

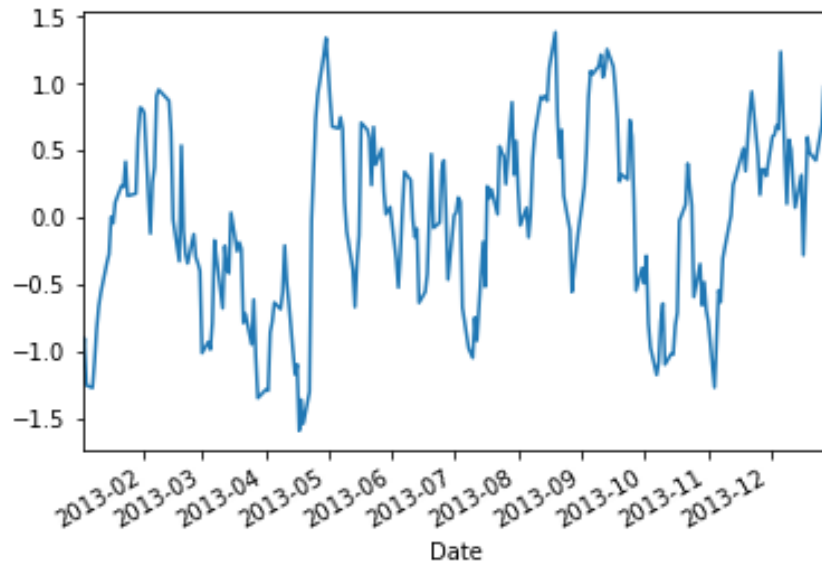
Gráfico 12 – Trade do par BRKM5 e CIEL3





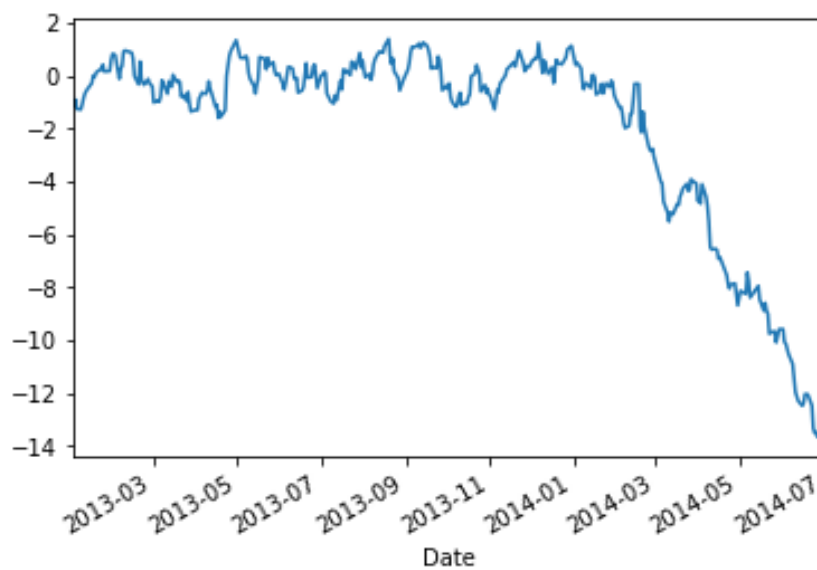
O par do gráfico 12 é composto por BRKM5 (Braskem) e CIEL3 (Cielo). O primeiro do setor químico e o segundo do setor financeiro. A seguir, o resíduo desse mesmo par no período de análise.

Gráfico 13 – Resíduo da regressão de BRKM5 e CIEL3



Nos primeiros testes com o modelo de parâmetros estáticos foi possível perceber que essa quebra de relação não é rara, e isso se refletiu no retorno das carteiras. A seguir, o gráfico dos dois períodos juntos.

Gráfico 14 – Resíduo do período analisado de BRKM5 e CIEL3



Como o modelo estático falha em captar as mudanças na relação dos ativos, o resultado de duas das três carteiras foi negativo mesmo sem considerar custos. A única carteira com resultado positivo foi a de pares do mesmo setor, com uma rentabilidade bruta de 6.35%. A seguir, os resultados das carteiras com parâmetros dinâmicos.

### Carteira com 15 ativos e *cash* neutro

A seguir, o gráfico 15 com a evolução do patrimônio e o gráfico 16 comparando com uma carteira de mercado que acompanha o Ibovespa

Gráfico 15 – Evolução patrimonial da carteira de 15 ativos *cash* neutro

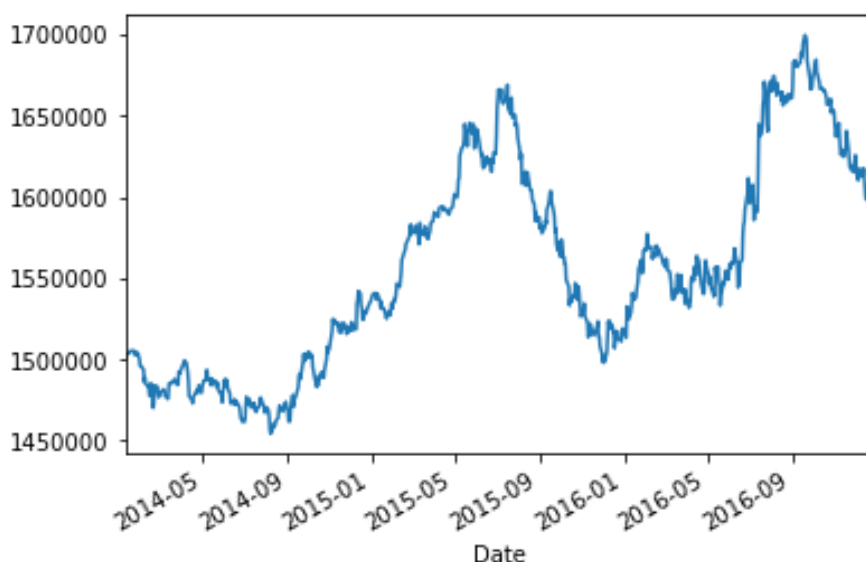
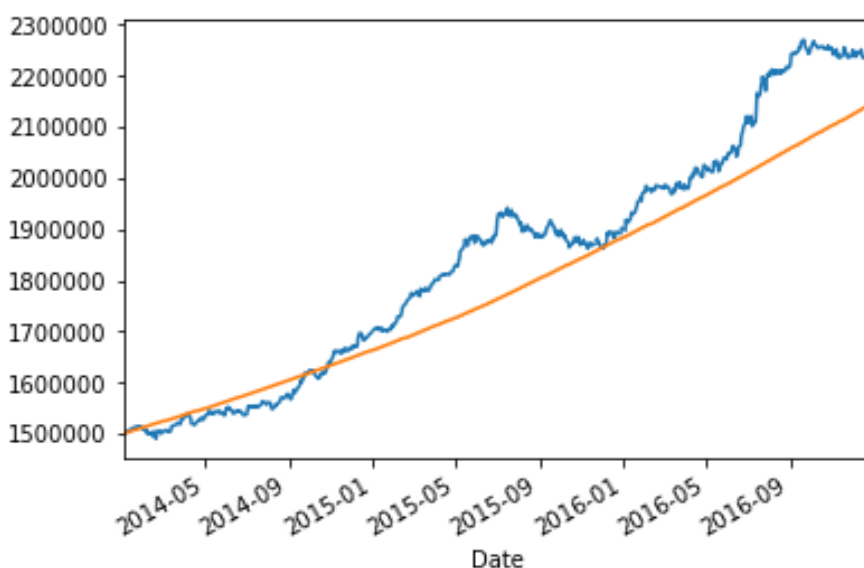


Gráfico 16 – Comparativo de uma carteira de mercado e da carteira de 15 ativos *cash* neutro



Pelo fato de a estratégia usada ser *cash* neutra é possível alocar o capital usado como garantia do *trade* em um ativo de baixo risco ao longo da operação. Com isso o retorno é maximizado. Aqui foi utilizado como base a Selic. Em uma operação real seria o equivalente a alocar em um fundo de liquidez diária, um título público, como uma LFT, ou qualquer outro ativo de baixo risco e alta liquidez. Abaixo o gráfico da evolução da carteira representada em azul. Em laranja está uma carteira comparativa onde só houve alocação em títulos aplicados à taxa Selic.

Gráfico 17 - Carteira de baixo risco e carteira combinada 15 ativos



A seguir, uma tabela com o resumo dos resultados da carteira.

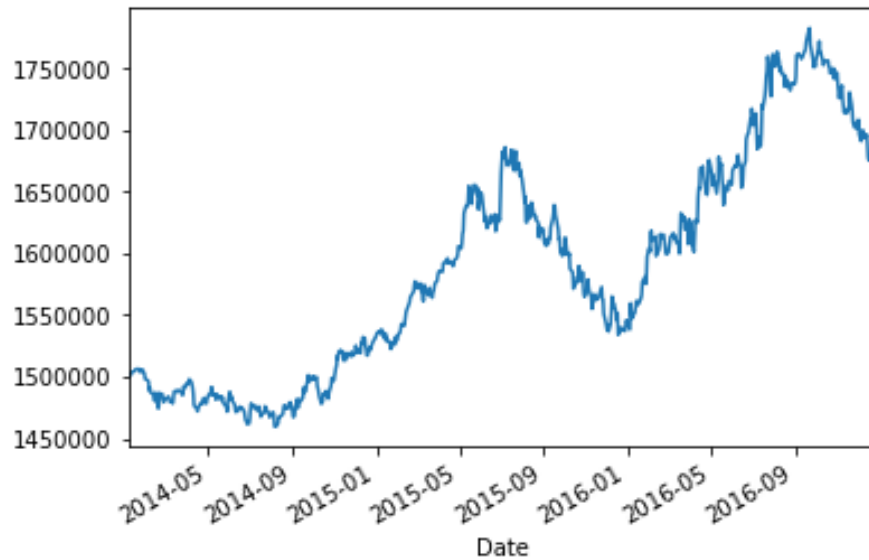
Tabela 5 – Resultados da carteira de 15 ativos *cash* neutro

Ret Máximo de um par	160.90%
Ret Mínimo de um par	-57.90%
Volatilidade anualizada	5.32%
Rentabilidade Líquida	7.30%
Pares com retorno positivo	8
Pares com retorno negativo	7
Correlação Ibovespa	0.1575

## Carteira com 15 ativos e beta neutro

A seguir, o gráfico da evolução patrimonial.

Gráfico 18 - Evolução patrimonial da carteira de 15 ativos beta neutro



Uma comparação entre a estratégia e o Ibovespa utilizando preços normalizados.

Gráfico 19 – Comparativo de uma carteira de mercado e da carteira de 15 ativos beta neutro

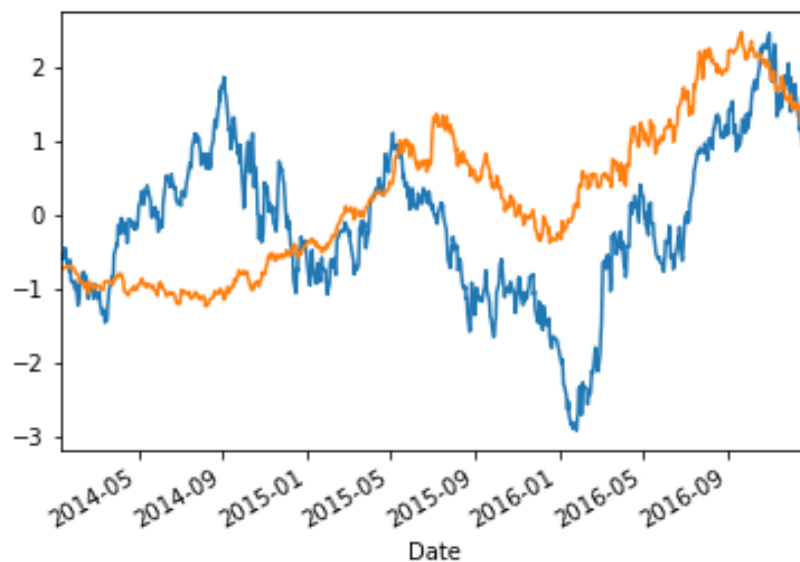


Tabela 6 – Resultados da carteira de 15 ativos beta neutro

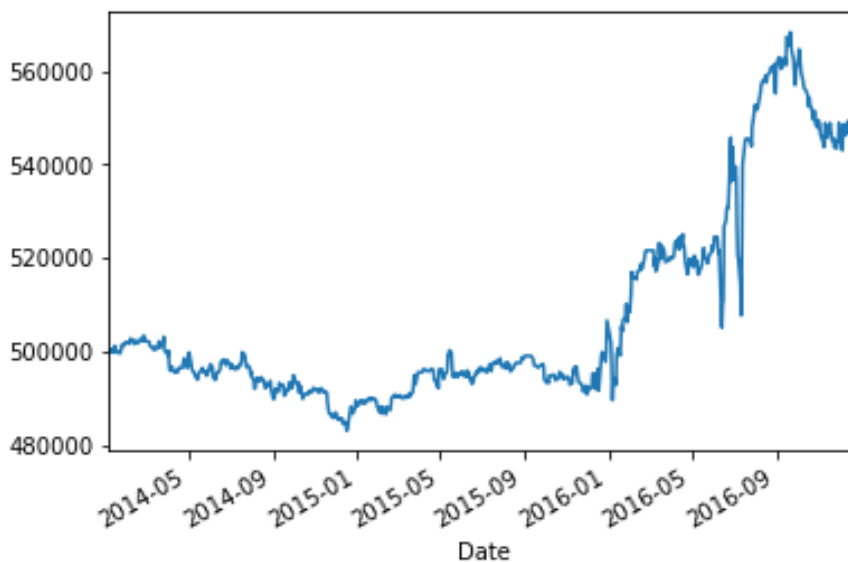
Ret Máximo de um par	201.33%
Ret Mínimo de um par	-44.80%
Volatilidade anualizada	5.93%
Rentabilidade Líquida	12.65%
Pares com retorno positivo	8
Pares com retorno negativo	7
Correlação Ibovespa	0.29

Uma observação em relação aos dados da carteira de 15 ativos é que a beta neutro apresentou uma correlação com o Ibovespa maior do que a *cash* neutro, o que vai contra o esperado a princípio. Foram também as carteiras que apresentaram a maior correlação com o Ibovespa entre todas as que foram testadas.

### Carteira ON e PN beta neutro

A seguir, o gráfico da evolução patrimonial.

Gráfico 20 - Evolução patrimonial da carteira ON e PN beta neutro



Uma comparação entre a estratégia e o Ibovespa utilizando preços normalizados.

Gráfico 21 - Comparativo de uma carteira de mercado e da carteira ON e PN beta neutro

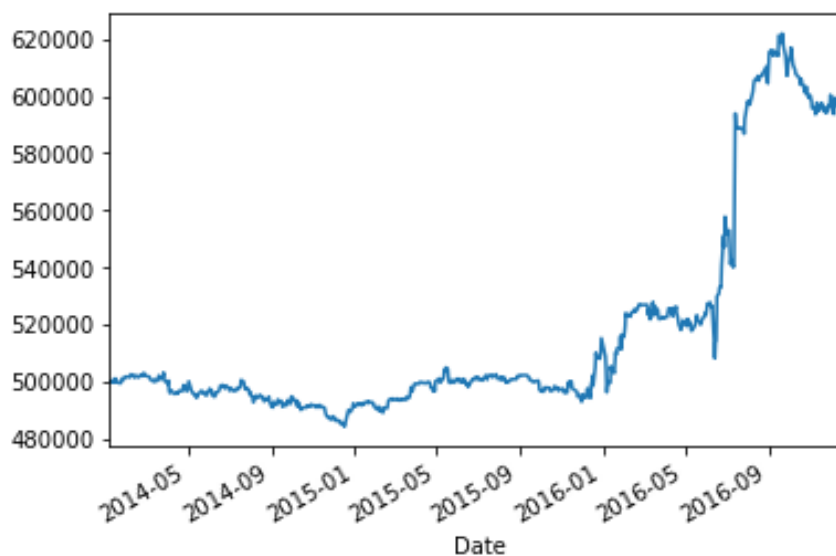


Tabela 7 – Resultados da carteira ON e PN beta neutro

Ret Máximo de um par	49.34%
Ret Mínimo de um par	-11.67%
Volatilidade anualizada	6.01%
Rentabilidade Líquida	9.39%
Pares com retorno positivo	3
Pares com retorno negativo	2
Correlação Ibovespa	-0.02

### Carteira ON e PN *cash* neutro

A seguir, o gráfico da evolução patrimonial.

Gráfico 22 - Evolução patrimonial da carteira ON e PN *cash* neutro

Uma comparação entre a estratégia e o Ibovespa utilizando preços normalizados.

Gráfico 23 - Comparativo de uma carteira de mercado e da carteira ON e PN *cash* neutro



Gráfico 24 - Carteira de baixo risco e carteira combinada setorial

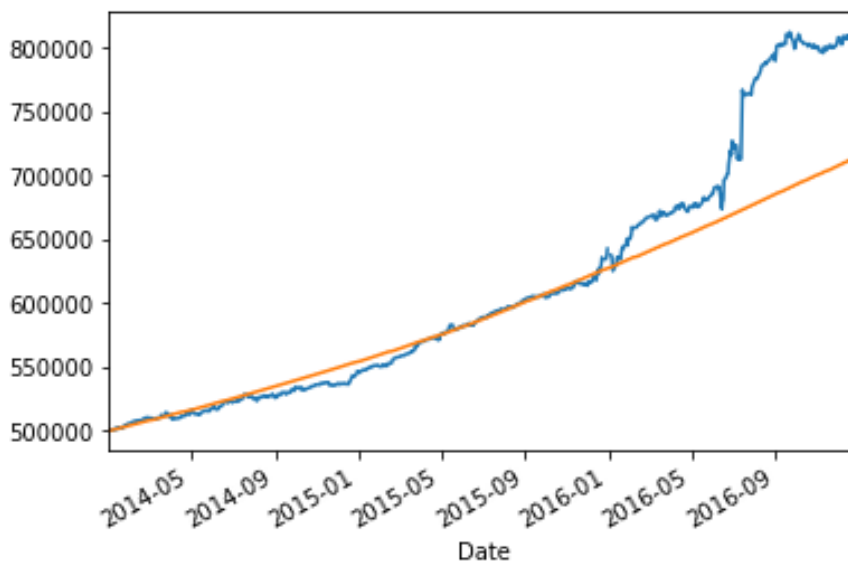


Tabela 8 – Resultados da carteira ON e PN *cash* neutro

Ret Máximo de um par	97.30%
Ret Mínimo de um par	-12.14%
Volatilidade anualizada	6.62%
Rentabilidade Líquida	19%
Pares com retorno positivo	3
Pares com retorno negativo	2
Correlação Ibovespa	0.02

As carteiras formadas por pares da mesma empresa foram as que apresentaram a menor correlação com o Ibovespa, mas só apresentaram bons resultados no último ano analisado, tendo ganhos quase nulos até 2016.

### Carteira setorial *cash* neutro

Gráfico 25 - Evolução patrimonial da carteira setorial *cash* neutro

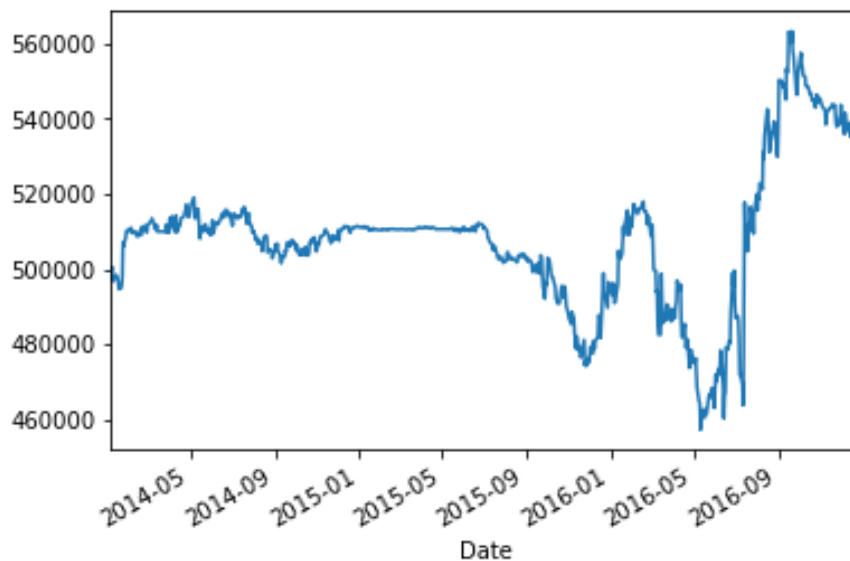


Gráfico 26 - Comparativo de uma carteira de mercado e da carteira setorial *cash* neutro

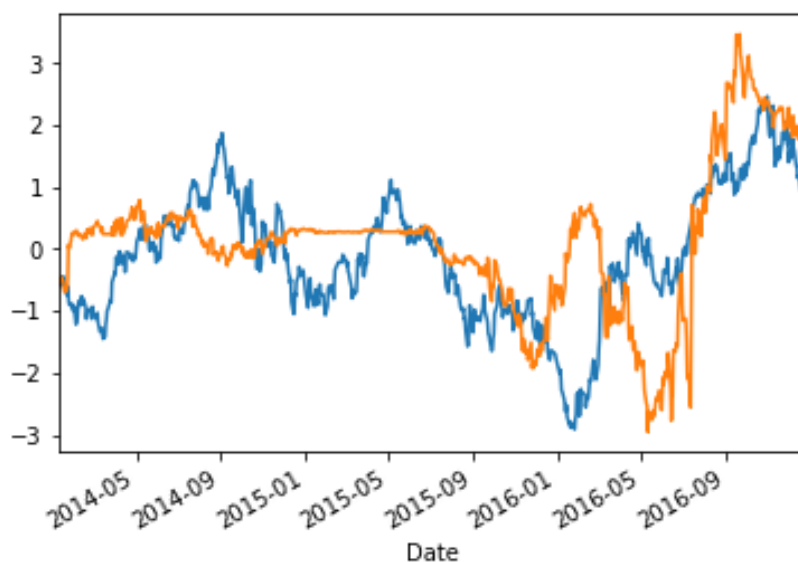
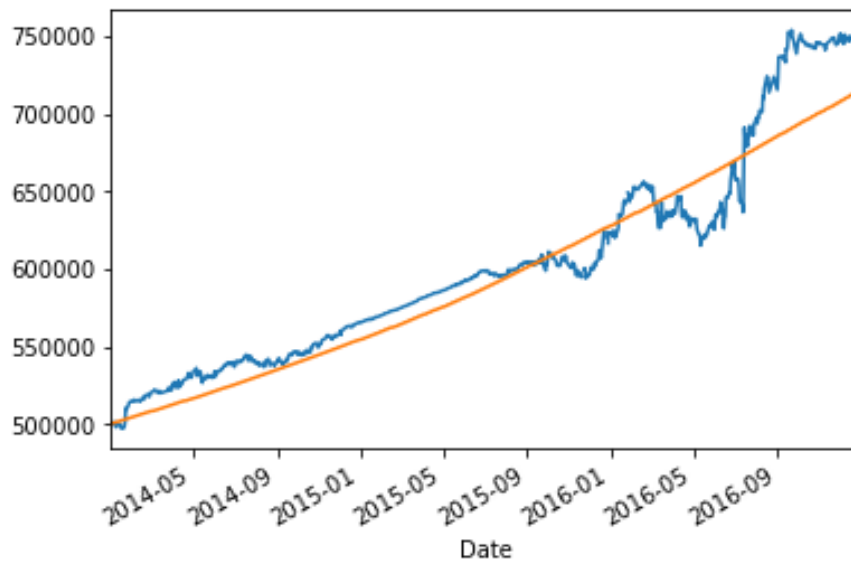




Gráfico 27 - Carteira de baixo risco e carteira combinada ON e PN

Tabela 9 – Resultados da carteira setorial *cash* neutro

Ret Máximo de um par	97.30%
Ret Mínimo de um par	-57.93%
Volatilidade anualizada	0.094307
Rentabilidade Líquida	6.87%
Pares com retorno positivo	3
Pares com retorno negativo	2
Correlação Ibovespa	-0.056

### Carteira setorial beta neutro

Gráfico 28 - Evolução patrimonial da carteira setorial beta neutro

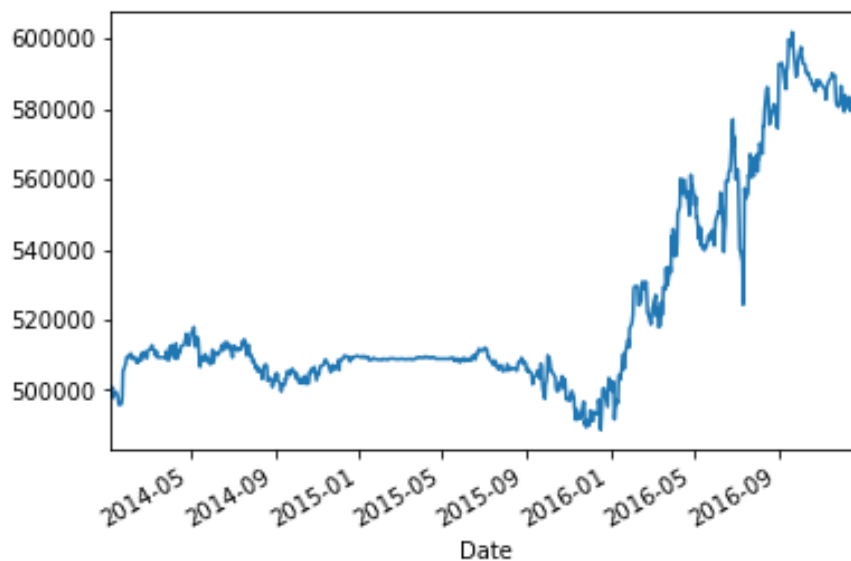


Gráfico 29 - Comparativo de uma carteira de mercado e da carteira setorial beta neutro



Tabela 10 – Resultados da carteira setorial beta neutro

Ret Máximo de um par	49.30%
Ret Mínimo de um par	-17.30%
Volatilidade anualizada	0.078031
Rentabilidade Líquida	15.60%
Pares com retorno positivo	4
Pares com retorno negativo	1
Correlação Ibovespa	0.11

As carteiras restritas por setor apresentaram correlação baixa com o Ibovespa e apresentaram um bom resultado acumulado ao longo do período de análise.

## Conclusão

---

Este trabalho testou diferentes estratégias e parâmetros em operações baseadas no método de cointegração. Foram selecionados pares entre os ativos do IBrX 100 após aplicar testes de Dickey Fuller aumentados nos resíduos das regressões. Num segundo momento foi utilizado o método da distância para ranquear os pares. Foram montadas três carteiras diferentes e cada uma foi testada com dois critérios, um com o método de beta neutro e outro com o método de *cash* neutro. Em duas das três carteiras o primeiro método gerou retornos maiores sem grandes mudanças na volatilidade da carteira. O segundo método porém permite a aplicação dos recursos em ativos de baixo risco em paralelo à operação, o que tem potencial para gerar resultados melhores no longo prazo.

Foram também observados problemas em relação ao modelo. O primeiro problema é a questão de usar o preço de fechamento dos ativos, o que não é realista já que ao longo do dia esses valores flutuam de forma que poderiam haver oportunidades de entrada e saída que o presente trabalho não considerou. Outro problema aparece ao tentar modelar os custos desse tipo de operação, já que ele seria variável dependendo do ativo e do momento do mercado. O trabalho considerou o mesmo custo para qualquer par e independente do tempo da operação, hipóteses que podem ser questionadas. Outra questão está relacionada aos ativos das carteiras consideradas. Um par em específico teve retornos bem acima da média e foi utilizado em todas as carteiras testadas. Isso pode ter sido reflexo de um evento extremo que gerou esses retornos e não necessariamente o modelo funcionaria tão bem se o período analisado fosse maior. Os parâmetros utilizados também podem ser questionados, já que no geral em trabalhos envolvendo cointegração de ativos financeiros não se tem uma regra em relação ao período analisado ou aos pontos de entrada e de saída das operações, tendo o trabalho usado os parâmetros que melhor se encaixaram para o período analisado.

## Referências

---

AVELLANEDA, M.; LEEY, J.H.. Statistical arbitrage in the US equities market. **Quantitative Finance**, v. 10, p. 761–782, 2010.

BRITO, S. C. S. **Comportamento de pares de ações no mercado brasileiro sob a ótica da cointegração, para preços intra-diários**. Dissertação (Mestrado Profissional em Finanças e Economia) - FGV - Fundação Getúlio Vargas, São Paulo, 2011.

CALDEIRA, J. F. Arbitragem Estatística, Estratégia Long-Short Pair Trading, Abordagem com Cointegração Aplicada ao Mercado de Ações Brasileiro. **Revista EconomiA**, v. 14, n. 1B, p. 521-546, 2013.

DO, B.; FAFF, R.; HAMZA, K. A new approach to modeling and estimation for pairs trading. **Proceedings of 2006 Financial Management Association European Conference**. 2006.

ELLIOT, R. J.; HOEK, J. V., MALCOLM, W. P. Pairs Trading. **Quantitative Finance**, v. 5, n. 3, p. 271–276, 2005.

ENGLE, R.; GRANGER, C. Co-integration and error correction: Representation, estimation and testing. **Econometrica**, v. 2, p.251–76, 1987.

GATEV, E.; GOETZMANN, W. N. ROUWENHORST, K. G. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. **Review of Financial Studies**, Oxford University Press for Society for Financial Studies, v. 19(3), p. 797-827, 2006.

PAPANTONIS, I. Cointegration-based trading: evidence on index tracking & market neutral strategies. **Managerial Finance**, v. 42(5), p. 449-471, 2016.

PUSPANINGRUM, H. **Pairs trading using cointegration approach**, Doctor of Philosophy thesis, School of Mathematics and Applied Statistics, University of Wollongong, 2012.

TEIXEIRA, A. A. E. **Pair trading in Bovespa with a quantitative approach: cointegration, OrnsteinUhlenbeck equation and Kelly criterion**. Dissertação

(Mestre em Modelagem Matemática da Informação) – FGV - Fundação Getúlio Vargas, Rio de Janeiro, 2014.

UNGEVER, C. PAIRS TRADING TO THE COMMODITIES FUTURES MARKET USING COINTEGRATION METHOD. **International Journal of Commerce and Finance**, v. 1(1), p. 25-38, 2015.

VIDYAMURTHY, G. **Pairs Trading, Quantitative Methods and Analysis**. John Wiley & Sons, Canada, 2004.

WOOLDRIDGE, J. M. **Introdução à econometria: Uma abordagem moderna**. Cengage Learning, 2013.

YU, P. L. H. LU, R. Cointegrated market-neutral strategy for basket trading. **International Review of Economics and Finance**, v. 49, p. 112–124, 2017.