



Trabalho de Conclusão de Curso

**Estratégias de Pairs Trading utilizando
Cointegração**

Vitor Coutinho Borges

21 de abril de 2023

Vitor Coutinho Borges

Estratégias de Pairs Trading utilizando Cointegração

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador(a): Prof. Dr. Flavio Ziegelmann

Porto Alegre
Abril de 2023

Vitor Coutinho Borges

Estratégias de Pairs Trading utilizando Cointegração

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientador(a) e pela Banca Examinadora.

Orientador(a): _____
Prof. Dr. Flavio Ziegelmann,

Banca Examinadora:

Prof. Dr. Eduardo de Oliveira Horta,

Porto Alegre
Abril de 2023

“A economia é muito importante para ser deixada apenas aos economistas.”
(Joseph E. Stiglitz)

Agradecimentos

Agradeço a Sra Lylcia de Lorena da Silva Coutinho Blascke, minha mãe, por ter me mostrado o valor do estudo e me proporcionar todo o suporte possível em minha jornada acadêmica.

Agradeço também a Sra Eunice da Silva Coutinho, minha tia e madrinha, um dos seres mais estudiosos que tive o prazer de conhecer, e que sempre me dá forças para seguir avançando em busca de meus objetivos.

Resumo

Este trabalho teve como objetivo investigar a aplicação de uma estratégia de Pairs-Trading utilizando a teoria de cointegração para identificar pares de séries temporais que apresentavam uma relação de longo prazo. Foram utilizados os preços de fechamento das ações de empresas que fazem parte do índice IBOVESPA em março de 2023, no horizonte de maio de 2011 até março de 2023. Os resultados mostraram que a estratégia foi capaz de gerar rentabilidades positivas, com rentabilidade média de 4,21% em cada janela anual, e descorrelacionadas com o mercado, em um total de 194 operações. A rentabilidade média de cada operação foi de 4,06%, já descontados os custos de operação, indicando a viabilidade da abordagem. O estudo também exemplifica cada etapa da aplicação da estratégia, incluindo a seleção de pares de ações, a modelagem da relação de cointegração e a execução da estratégia de negociação. No geral, o estudo demonstrou a importância da teoria de cointegração na aplicação da estratégia de pairs trading e como isso pode levar a resultados positivos na negociação de ativos financeiros.

Palavras-Chave: Séries Temporais, Estacionariedade, Cointegração, Pairs-Trading.

Sumário

1	Introdução	10
2	Metodologia	13
2.1	Estacionariedade	13
2.2	Cointegração via Engle-Granger	14
2.3	Arbitragem Estatística e Pairs Trading	14
2.4	Custos de Operação	16
2.5	Rentabilidade Líquida	17
3	Experimento	18
3.1	Base de Dados	19
3.2	Exemplo Simples	19
3.2.1	Testes de estacionariedade	20
3.2.2	Testes de Cointegração	21
3.2.3	Backtesting	22
3.2.4	Testes fora da Amostra	23
3.3	Resultados Gerais	24
3.3.1	Janela 2012-2013	25
3.3.2	Janela 2013-2014	26
3.3.3	Janela 2014-2015	27
3.3.4	Janela 2015-2016	28
3.3.5	Janela 2016-2017	29
3.3.6	Janela 2017-2018	30
3.3.7	Janela 2018-2019	31
3.3.8	Janela 2019-2020	31
3.3.9	Janela 2020-2021	32
3.3.10	Janela 2021-2023	33
4	Conclusão	35
	Referências Bibliográficas	35

Lista de Figuras

Figura 3.1: Gráfico de Preços de MRVE3 e PETR3	20
Figura 3.2: Resíduos Normalizados	23
Figura 3.3: Resíduos Normalizados no Período de Teste	24
Figura 3.4: Gráfico de dispersão entre a Rentabilidade de cada Trade e do IBOV durante o período	25
Figura 3.5: Rentabilidade dos Pares da Janela 2012-2013	26
Figura 3.6: Rentabilidade dos Pares da Janela 2013-2014	27
Figura 3.7: Rentabilidade dos Pares da Janela 2014-2015	28
Figura 3.8: Rentabilidade dos Pares da Janela 2015-2016	29
Figura 3.9: Rentabilidade dos Pares da Janela 2016-2017	30
Figura 3.10: Rentabilidade dos Pares da Janela 2017-2018	30
Figura 3.11: Rentabilidade dos Pares da Janela 2018-2019	31
Figura 3.12: Rentabilidade dos Pares da Janela 2019-2020	32
Figura 3.13: Rentabilidade dos Pares da Janela 2020-2021	33
Figura 3.14: Rentabilidade dos Pares da Janela 2021-2023	34

Lista de Tabelas

Tabela 3.1: Coeficientes da regressão de MRVE3 e PETR3 com seus lags	21
Tabela 3.2: Avaliação do Modelo	21
Tabela 3.3: Regressão Linear entre MRVE3 e PETR3	21
Tabela 3.4: Regressão Linear para checar estacionariedade nos resíduos	22
Tabela 3.5: Regressão Linear para checar estacionariedade nos resíduos	23
Tabela 3.6: Trades no período de teste	24
Tabela 3.7: Rentabilidade Média por Janela	24
Tabela 3.8: Índices de Sharpe no período de treino	25
Tabela 3.9: Índices de Sharpe no período de treino	26
Tabela 3.10: Índices de Sharpe no período de treino	27
Tabela 3.11: Índices de Sharpe no período de treino	28
Tabela 3.12: Índices de Sharpe no período de treino	29
Tabela 3.13: Índices de Sharpe no período de treino	30
Tabela 3.14: Índices de Sharpe no período de treino	31
Tabela 3.15: Índices de Sharpe no período de treino	32
Tabela 3.16: Índices de Sharpe no período de treino	33
Tabela 3.17: Índices de Sharpe no período de treino	34

1 Introdução

A arbitragem estatística é um conjunto de estratégias de investimento market-neutral, que se baseiam em regras estatísticas em vez de fundamentos e visam retornos não correlacionados com as tendências do mercado. Essas estratégias envolvem a compra e venda simultânea de diferentes ativos financeiros para obter lucros com as ineficiências do mercado, identificadas por meio de análises estatísticas sofisticadas utilizando modelos matemáticos e estatísticos. A arbitragem estatística é frequentemente utilizada por gestores de hedge funds que buscam obter retornos independentemente das flutuações do mercado. No entanto, a implementação adequada requer conhecimentos sofisticados em matemática, estatística e finanças, bem como uma compreensão detalhada das nuances do mercado financeiro. Como qualquer estratégia de investimento, há riscos envolvidos, tornando importante uma abordagem cuidadosa e bem informada por parte dos gestores.

Pairs trading é uma estratégia de investimento que utiliza arbitragem estatística para lucrar com a convergência de preços entre dois ativos com comportamentos históricos semelhantes. Investidores assumem posições compradas e vendidas em cada ativo quando o spread entre eles se amplia. Essa estratégia é considerada segura por reduzir a exposição ao mercado geral, sendo objeto de estudos para aprimorar seus métodos e aumentar a efetividade, com destaque para métodos baseados na distância e cointegração entre os pares de ativos. Em 2006, Evan Gatev conduziu um estudo relevante sobre o método da distância em pares trading, utilizando dados diários de preços do mercado norte-americano entre 1962 e 1998. Os pares foram formados com base na menor soma do quadrado das diferenças das séries normalizadas e selecionados com base em um período de 12 meses e ativos negociados em todos os dias de pregão. Para abrir uma posição, a diferença entre os pares deveria ultrapassar dois desvios-padrão. Gatev demonstrou que os pares de menor distância apresentaram resultados mais satisfatórios. Embora tenha sido amplamente estudado nos últimos anos, alguns trabalhos, como [Silva et al. \(2023\)](#) e [Xie e Wu \(2013\)](#) apontam que uma única medida de distância pode falhar em capturar a dinâmica do spread entre um par de preços de ações, uma vez que a linearidade entre os preços de ações é extremamente rara.

Com o objetivo de identificar pares de ativos de forma mais eficaz, surgiu o método da cointegração. A teoria da cointegração pode explicar a relação de longo prazo entre preços de ativos, caso essa relação seja linear. Esse método foi primeiramente apresentado por [Lucas e Franses \(1997\)](#) e [Alexander \(1999\)](#) e amplamente discutido em vários trabalhos, como os de [Caldeira e Portugal \(2010\)](#), [Caldeira \(2011\)](#), [Herlemont \(2000\)](#) e [Caldeira e Moura \(2013\)](#). A ideia central do método

é encontrar uma combinação linear de duas séries de preços que seja estacionária. Para verificar essa propriedade, uma abordagem muito utilizada é a proposta por [Engle e Granger \(1987\)](#), que é baseada na minimização da variância e é desejável para a minimização do risco de mercado em estratégias, já que grande parte da metodologia é baseada em correlação. É importante lembrar que correlação e cointegração são conceitos diferentes: duas séries podem apresentar cointegração mas não terem correlação linear em um certo período de tempo. De acordo com [Caldeira e Moura \(2013\)](#), a correlação reflete co-movimentos nos retornos, os quais são suscetíveis a grande instabilidade ao longo do tempo, trata-se de uma medida intrinsecamente de curto prazo e estratégias baseadas em correlação frequentemente exigem rebalanceamentos frequentes. Por outro lado, a cointegração é uma medida de co-movimentos de longo prazo nos preços e podem ocorrer mesmo em períodos em que a correlação parece baixa. O método da cointegração busca relacionar os preços de suas ações em nível e tornou-se uma ferramenta poderosa que permite a aplicação de modelos de estimação simples com variáveis não estacionárias.

Os custos de operação são um aspecto crucial para a rentabilidade do pairs-trading, pois as posições compradas são financiadas pelas posições vendidas. É importante considerar os custos de aluguel das ações vendidas e outros custos operacionais. Além disso, a correlação dos retornos com os índices de mercado é um fator a ser observado, uma vez que o objetivo da estratégia é obter lucros independentemente da direção do mercado. [Caldeira e Moura \(2013\)](#) obteve resultados positivos para todos os anos em que testou a estratégia, incluindo o ano de 2001, que foi fortemente afetado pelo desastre de 11 de setembro. A estratégia também apresentou baixa correlação com o Ibovespa, cumprindo assim o objetivo de neutralidade em relação ao mercado, tanto para dados anuais quanto para o período fechado.

Neste estudo, desenvolvemos uma estratégia de pairs-trading baseada em cointegração, abrangendo desde conceitos fundamentais até a aplicação prática da estratégia em testes fora da amostra. A estratégia será aplicada em um período de 12 meses para a formação de pares, seguido por 12 meses para a verificação dos resultados. Apenas os pares cujo spread ultrapassar 2 desvios padrão serão considerados para a estratégia, utilizando ações que fazem parte do índice Ibovespa em março de 2023 e que já estavam em operação em maio de 2011, com dados obtidos pela Bloomberg. Após a identificação de mais de 20 pares cointegrados, a seleção dos pares será realizada com base no índice de Sharpe durante o período de treino, e a rentabilidade líquida da estratégia será avaliada em cada período de trading. Este estudo tem como objetivo consolidar conceitos importantes para o desenvolvimento de estratégias de pairs-trading usando cointegração de séries temporais, incluindo testes de Engle e Granger para verificar a cointegração dos pares, e testar a lucratividade dessas regras comerciais.

Com o objetivo de desenvolver uma estratégia de pairs-trading utilizando cointegração com esse pares de ativos provindos do índice Ibovespa, foram sugeridas as seguintes questões de pesquisa:

1. Desenvolver um algoritmo capaz de identificar pares de ativos que sejam cointegrados e aptos para serem utilizados em uma estratégia de pairs-trading.
2. Avaliar a rentabilidade líquida de tal estratégia, além de sua descorrelação com o mercado.

Foram obtidos resultados satisfatórios, com rentabilidades positivas e descorrelacionadas com o mercado em todas as janelas de operação. Ao total, 194 operações foram realizadas, com uma rentabilidade média de 4,06%.

2 Metodologia

2.1 Estacionariedade

Uma série temporal y_t será fracamente estacionária se

1. $Ey_t = \mu$, onde μ é uma constante, para $\forall t$.
2. $E|y_t^2| < \infty$, $\forall t$.
3. $Cov(y_t, y_s) = \gamma(m)$, onde γ é uma função que depende apenas da diferença $m = t - s$, $\forall t, s$.

O teste de Dickey-Fuller Aumentado (ADF, na sigla em inglês) é baseado no seguinte modelo de regressão:

$$\Delta y_t = \alpha + \beta y_{t-1} + \gamma_1 \Delta y_{t-1} + \dots + \gamma_{p-1} \Delta y_{t-p+1} + \epsilon_t ,$$

onde Δy_t é a diferença entre y_t e y_{t-1} e ϵ_t é o erro do modelo. As hipóteses do teste ADF podem ser definidas como

1. Hipótese nula (H0): A série temporal possui raiz unitária, ou seja, não é estacionária. $H_0 : y_t = \phi y_{t-1} + \epsilon_t$, onde $\phi = 1$.
2. Hipótese alternativa (H1): A série temporal não possui raiz unitária, ou seja, é estacionária. $H_1 : y_t = \phi y_{t-1} + \epsilon_t$, onde $|\phi| < 1$.

A estatística de teste é definida como

$$t_{ADF} = \frac{\hat{\beta}}{SE(\hat{\beta})} ,$$

onde $\hat{\beta}$ é o estimador de mínimos quadrados do coeficiente de y_{t-1} e $SE(\hat{\beta})$ é o erro padrão desse estimador. Esse valor é então comparado a um valor crítico para se determinar se a hipótese nula deve ser rejeitada ou não. Se o valor de t_{ADF} for menor em módulo que o valor crítico, a hipótese nula não é rejeitada e conclui-se que a série pode possuir raiz unitária. Se o valor de t_{ADF} for maior em módulo que o valor crítico, a hipótese nula é rejeitada e conclui-se que a série não deve possuir raiz unitária.

2.2 Cointegração via Engle-Granger

A cointegração é um importante conceito que descreve relações de longo prazo entre séries temporais que não são $I(0)$. O fenômeno da cointegração ocorre quando duas séries temporais de mesma ordem de integração (maior que zero), y_{1t} e y_{2t} , possuem uma relação de equilíbrio de longo prazo, mesmo que individualmente não apresentem estacionariedade. A relação de cointegração pode ser modelada como

$$y_{2t} = \beta_0 + \beta_1 y_{1t} + u_t, \quad (2.1)$$

onde β_0 e β_1 formam o vetor de cointegração e u_t é uma série estacionária.

Logo, para identificar a cointegração, é necessário realizar testes estatísticos que verifiquem se as séries temporais possuem uma combinação linear que resulte em uma série estacionária. Esse processo envolve a aplicação de um modelo de correção de erros, que incorpora os erros no equilíbrio das séries e fornece informações sobre os comportamentos de curto e longo prazo dessas séries.

A aplicação da metodologia de Engle-Granger para a verificação de cointegração entre um par de séries envolve a aplicação de uma regressão linear entre as duas séries temporais y_{1t} e y_{2t} da forma

$$y_{2t} = \beta_0 + \beta_1 y_{1t} + \epsilon_t. \quad (2.2)$$

Se ambas as séries temporais não são estacionárias, a série de resíduos ϵ_t pode ser estacionária, o que indica que as séries temporais são cointegradas. A estacionariedade dos resíduos pode ser testada utilizando um teste estatístico, como o teste de Dickey-Fuller Aumentado (ADF). Formalmente, o procedimento de Engle-Granger pode ser realizado da seguinte maneira:

1. Verificar se as séries temporais y_{1t} e y_{2t} são integradas de ordem k , ($k > 0$), ou seja, se ambas as séries são não estacionárias. Isso pode ser feito utilizando um teste de raiz unitária, como o teste ADF.
2. Regredir y_{2t} em y_{1t} e obter os resíduos ϵ_t .
3. Verificar se a série dos resíduos ϵ_t é estacionária. Isso pode ser feito utilizando um teste estatístico, como o teste ADF.
4. Se a série dos resíduos ϵ_t for estacionária, então as séries temporais y_{1t} e y_{2t} são cointegradas, o que indica que existe uma relação de longo prazo entre elas.

2.3 Arbitragem Estatística e Pairs Trading

A arbitragem estatística é uma técnica utilizada por traders para explorar desequilíbrios no mercado financeiro, o conceito se baseia na teoria de mercados eficientes, onde o desempenho do mercado não é proporcional a todas as informações disponíveis, baseando-se no preço relativo, conceito que diz que ativos similares devem ter preços similares, pois eles podem ser considerados substitutos um do outro. Para que uma estratégia de pairs trading seja bem-sucedida, é necessário identificar desequilíbrios nos preços relativos que tenham a característica de retornar à média a longo prazo, logo, é necessário algum mecanismo que força tal reversão à média.

Estratégias de arbitragem estatística não são isentas de riscos, para reduzir o risco, é possível utilizar estratégias com a característica de serem *market-neutral*, em que o trader mantém uma posição comprada e vendida ao mesmo tempo, eliminando assim o risco direcional do mercado e descorrelacionando o retorno da estratégia com o do índice de referência. Isso significa que, independentemente da direção que o mercado tomar, a estratégia poderá permanecer lucrativa. Em resumo, a arbitragem estatística é uma abordagem puramente estatística para explorar ineficiências do mercado definidas pelo desvio dos equilíbrios de longo prazo entre preços de ações.

Conforme dito, a técnica de pairs trading é baseada na ideia de que, em um mercado eficiente, os preços de dois ativos correlacionados devem se mover juntos, de modo que, quando há uma divergência significativa entre esses preços, espera-se que essa diferença seja corrigida em algum momento futuro, a cointegração entra em jogo nesse cenário porque é capaz de identificar a relação de longo prazo entre os ativos, permitindo que se estime o spread esperado. Assim, quando ocorre uma divergência entre os preços atuais e o spread esperado, pode-se realizar uma operação de compra e venda simultâneas dos ativos, apostando que essa diferença será corrigida no futuro. Tal abordagem tem sido amplamente utilizada em diversos estudos sobre pairs trading, como os de [Lucas e Franses \(1997\)](#), [Alexander \(1999\)](#), e [Vidyamurthy \(2004\)](#). Esses estudos demonstraram que a cointegração é capaz de capturar a relação de longo prazo entre os ativos e, portanto, pode ser usada para construir estratégias lucrativas de pairs trading. Essa abordagem tem se mostrado útil em outros contextos, como na modelagem de séries financeiras e na análise de políticas macroeconômicas. O trabalho de [Vidyamurthy \(2004\)](#) é uma das principais referências para o estudo de estratégias de pairs trading utilizando cointegração, neste artigo, o autor apresenta uma visão geral das estratégias de pairs trading e aborda de forma detalhada a aplicação da cointegração na modelagem de pares de ativos financeiros. O autor explica que a estratégia de pairs trading baseia-se na premissa de que pares de ativos financeiros com alta correlação apresentam um comportamento similar em relação aos movimentos de preço, e que desvios temporários dessa relação de preço são corrigidos no longo prazo.

Em [Caldeira e Moura \(2013\)](#) os autores propõem uma metodologia de seleção de pares de ativos com base na análise de cointegração entre suas séries temporais de preços, os resultados indicam que a estratégia de pairs trading com base em cointegração pode gerar lucros consistentes ao longo do tempo, apresentando desempenho superior em comparação com outras abordagens de seleção de pares de ativos, como a correlação e o índice beta. O estudo também aborda a importância da escolha adequada do horizonte de tempo e da janela de cointegração para obter resultados confiáveis e significativos. Neste trabalho, onde os autores utilizaram uma estratégia de pairs trading utilizando dados do mercado brasileiro, o spread entre os dois ativos foi calculado utilizando os logaritmo dos preços de ativos, representados por P_1 e P_2 :

$$e_t = \log(P_{1t}) - \log(P_{2t}) , \quad (2.3)$$

onde P_{1t} e P_{2t} representam os preços de fechamento dos ativos A e B em um determinado tempo t . Após isso, foi calculado um escore Z com uma abordagem utilizando o desvio padrão para identificar bons pontos de entrada. Sejam μ_e a média dos spreads calculados entre os ativos e σ_e o seu desvio padrão, logo o escore Z será da forma:

$$Z_t = \frac{e_t - \mu_e}{\sigma_e}, \quad (2.4)$$

para justificar a utilização do spread normalizado, Caldeira menciona que

“A média e o desvio padrão são calculados com base no período dentro da amostra e são usados para normalizar spread tanto no período dentro da amostra quanto fora da amostra. Nesta etapa uma alternativa é usar algum critério de suavização exponencial para a média e desvio padrão, de tal forma que as duas medidas possam variar ao longo do tempo.”

Uma posição era aberta sempre que Z_t ultrapassava +2 ou -2, utilizando os limites de -0.5 e 0.75 para encerrar as posições em caso de lucro. Como medida de risco, as posições também eram encerradas sempre que apresentavam 7% de perda. Além disso, os autores também incluíram como critério de seleção o índice de Sharpe dos pares apresentado no período de formação da estratégia, foram obtidos em média, 90 pares de ações para cada período estudado e o retorno da estratégia foi de 16,38%, com baixa correlação com o índice Ibovespa.

2.4 Custos de Operação

Para obter uma medição mais precisa do retorno da estratégia, é crucial considerar os custos de operação envolvidos. Esses custos incluem taxas de bolsa de valores, corretagem e taxas de aluguel de ações. As taxas de bolsa de valores são geralmente de cerca de 0,345% do principal de compra ou venda, mas podem ter um impacto significativo nos resultados da estratégia ao longo do tempo. A corretagem varia de acordo com o volume transacionado, enquanto as taxas de aluguel de ações são cobradas quando são abertas posições vendidas e podem representar uma parcela significativa dos custos totais, dependendo do tempo que a posição vendida é mantida. Portanto, é importante levar em conta esses custos ao calcular o retorno da estratégia.

Para este trabalho, os custos serão estimados em 0,5% para cada ponta da estratégia. Ao abrir uma operação de pairs trading, será cobrado 0,5% sobre o principal do valor de venda e 0,5% sobre o valor de compra e ao fechar a estratégia, esses valores serão cobrados novamente para vender a parte comprada e recomprar a parte vendida da estratégia. Essa estimativa pode variar de acordo com o tipo de operação e o volume transacionado, portanto é fundamental estar sempre atento a esses custos ao calcular o retorno da estratégia.

Em (Caldeira e Moura, 2013) os custos de operação são definidos por:

$$C_o = 2 \frac{(1-C)}{(1+C)},$$

onde C_o representa o custo total de operação e C o custo percentual de entrada e saída (aqui fixado em 0,5%).

2.5 Rentabilidade Líquida

A rentabilidade líquida de um investimento é um indicador importante que mostra a rentabilidade descontando os custos e impostos, fornecendo uma visão mais precisa do retorno. É fundamental considerar esses custos ao avaliar a rentabilidade de um investimento, pois podem ter um impacto significativo na rentabilidade ao longo do tempo. A rentabilidade líquida pode ser calculada de várias maneiras, dependendo do tipo de investimento e dos custos envolvidos, como subtraindo a taxa de administração do fundo em um fundo de investimento ou impostos e taxas de corretagem em ações.

O cálculo da rentabilidade (R_t) da estratégia será feita através do cálculo de rentabilidade de cada ponta da operação:

$$R_t = n * \left(\frac{P_t^l}{P_0^l} - \frac{P_t^s}{P_0^s} \right),$$

sendo P_0^l e P_0^s os preços das ações compradas e vendidas no momento de abertura da posição e P_t^l P_t^s os preços no momento do fechamento da operação, após isso, serão descontados os custos da operação.

3 Experimento

Os métodos utilizados serão baseados nos trabalhos (Caldeira e Moura, 2013) e (Vidyamurthy, 2004), com período de formação de pares de 12 meses e a verificação de quais séries são estacionárias de primeira ordem será feita através de testes de raiz unitária ADF e após isso, combinar as séries estacionárias de mesma ordem para checar cointegração entre essas através de uma metodologia Engle Granger, além disso, serão escolhidos os pares com os melhores 20 índices de Sharpe durante o período de treino para serem testadas no período de trading, que ocorrerá nos próximos 12 meses. Após a finalização de um período de teste, as janelas de treino e teste se moverão 12 meses a frente, e os coeficientes serão re-estimados nos 12 meses subsequentes ao primeiro período de treino e será realizado um período de trade nos 12 meses subsequentes ao novo período de treino.

A regra de trading que será utilizada baseia-se no Z-Score definido previamente. Quando o spread entre as ações ultrapassar 2 desvios padrão, uma posição será aberta e fechada quando o spread se movimentar 0,5 desvios padrão na direção contrária ou após 50 dias úteis, a fim de limitar possíveis perdas. Segundo (Caldeira e Moura, 2013), um período de 50 dias é suficiente para que um spread entre séries temporais volte ao seu equilíbrio, em média. A reversão à média deve ocorrer em 10 dias, não havendo necessidade de esperar por mais tempo.

Para cada par de ativos seguiremos os seguintes passos a fim de encontrar evidências de cointegração:

1. Será realizado um teste de Dickey-Fuller nas séries temporais dos pares individualmente, a fim de testar estacionariedade nas séries. Para que possamos continuar com o estudo do par, é necessário que ambas as séries sejam não estacionárias, pois desejamos encontrar uma relação estável de longo prazo entre as séries. Para isso, foi ajustado um código em R que faça esse teste para cada ativo.
2. Após a verificação de ausência de estacionariedade e certificar-se que as séries são $I(1)$, é ajustado um modelo de regressão linear entre os preços de fechamento dos dois ativos e é avaliada a significância do modelo por meio dos valores dos coeficientes, dos resíduos e da análise de variância (ANOVA) dos resíduos. É criado um dataset para avaliar os resíduos do modelo, onde é realizado um novo teste de estacionariedade utilizando um nível de significância de 95%, para verificar se há cointegração entre os ativos.

A partir da verificação da cointegração entre o par, é necessário identificar pontos de entrada para operações short (venda a descoberto) e long (compra).

Primeiramente, é calculado o desvio padrão dos resíduos do modelo preditivo. A partir da volatilidade dos resíduos, são definidos os valores para os gatilhos de entrada e para a perda máxima permitida em cada operação (stop loss), também é contabilizado o número de dias em que cada operação permanece aberta e definido um limite de 50 dias. São simuladas entradas de R\$ 2000 em cada operação e calcula o rendimento obtido em cada uma. O cálculo é realizado a partir da quantidade de cada ativo financeiro comprado ou vendido correspondente ao valor do capital investido e do saldo entre o valor de venda e compra dos ativos financeiros, dependendo da direção da operação. O retorno é calculado como a diferença entre o saldo de venda e compra dividido pelo capital total investido, descontados os custos de operação.

3.1 Base de Dados

A base de dados fornecida pela Bloomberg contém informações sobre o preço de fechamento diário das empresas que compõem o IBOVESPA, um índice de referência para o desempenho do mercado de ações brasileiro. Os dados se estendem desde 3 de maio de 2011, e ao longo do tempo, o IBOVESPA sofre mudanças em sua composição devido a diversos fatores, como fusões, aquisições ou desempenho financeiro. Assim, algumas empresas que atualmente fazem parte do índice não estavam presentes na data inicial da base, e algumas empresas que estavam presentes foram excluídas do estudo. Ao todo, a base de dados inclui 65 empresas que faziam parte do IBOVESPA em março de 2023.

3.2 Exemplo Simples

Este capítulo se concentra no desenvolvimento de um modelo de pairs trading por cointegração a partir dos dados históricos de preços de fechamento diário das empresas que faziam parte do IBOVESPA em março de 2023. Será possível identificar pares de ações que apresentam relações de longo prazo e estáveis, indicando a possibilidade de desenvolver uma estratégia de negociação baseada em cointegração.

Para exemplificar cada etapa do processo de desenvolvimento de estratégias, escolhemos o par MRVE3 e PETR3, ações referentes as empresas MRV e Petrobrás, no primeiro período de trade, ocorrido no ano de 2013.

No gráfico 3.1, podemos observar os preços de fechamento dos ativos durante todo o período de estudo. É curioso observar que as séries aparentam movimentar-se em conjunto no início da amostra, especialmente até o ano de 2014, enquanto no final da amostra, uma relação de cointegração não é perceptível em uma análise gráfica. Tal fato evidencia a necessidade de uma reestimação periódica dos coeficientes e relação de cointegração, especialmente quando estamos trabalhando com séries temporais muito longas, como é o caso desse estudo. É importante também notar que as séries temporais dessas ações aparentam não ser estacionárias, pois parecem possuir alguma tendência.

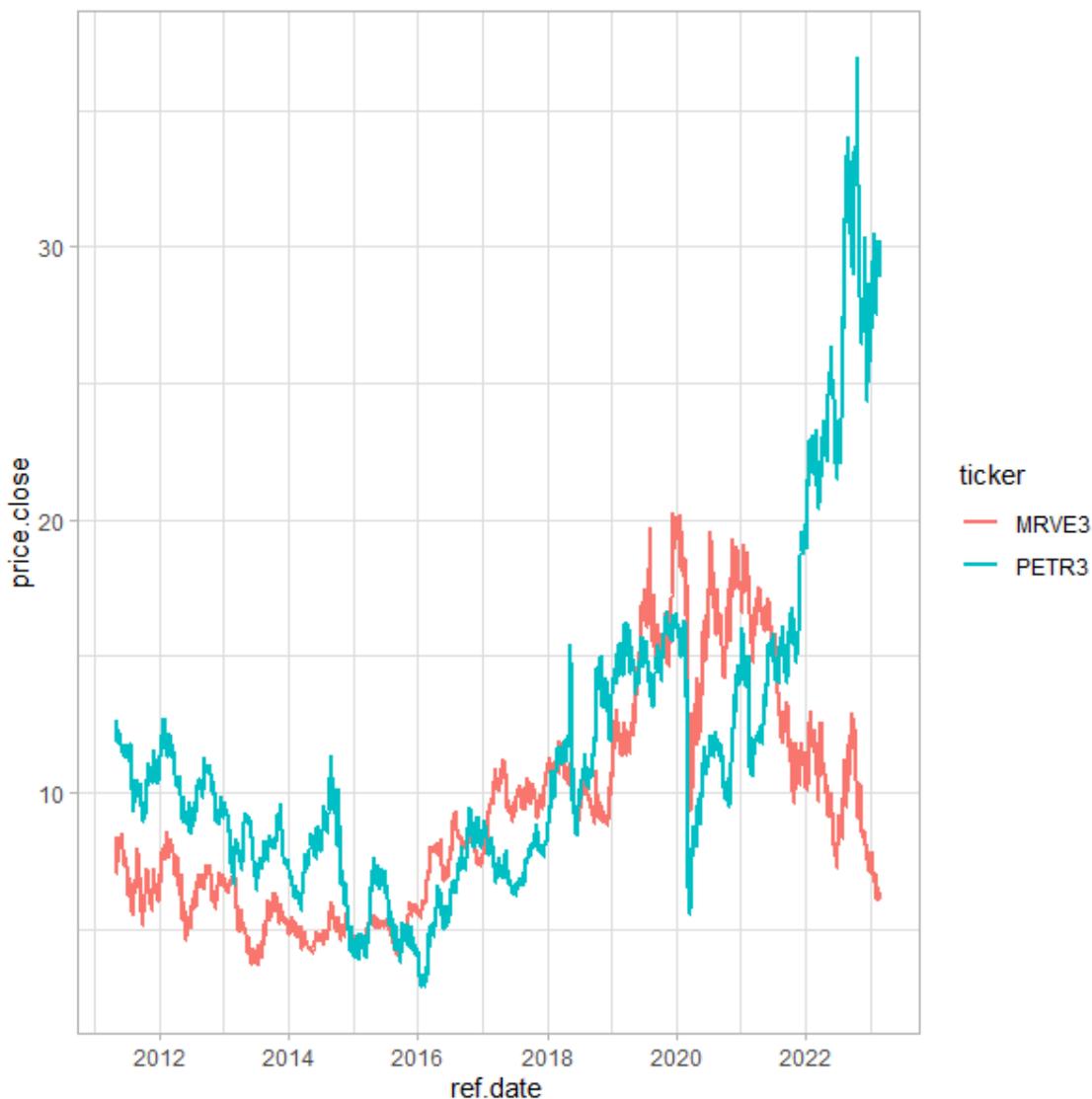


Figura 3.1: Gráfico de Preços de MRVE3 e PETR3

Caso seja indentificada a presença de cointegração entre MRVE3 e PETR3 nessa janela, serão fortes candidatas a fazerem parte de uma estratégia de negociação. Com base na análise das séries temporais, pode-se identificar pontos de entrada e saída que maximizem o retorno da estratégia, levando em consideração as particularidades de cada ação e as características do mercado.

3.2.1 Testes de estacionariedade

Apesar de análises gráficas serem úteis para visualizar padrões e tendências nas séries temporais, é importante ter uma metodologia sistemática e estatística para confirmar a presença ou ausência de estacionariedade. É aqui que entra o teste ADF (Augmented Dickey-Fuller), uma técnica amplamente utilizada para testar a estacionariedade em séries temporais.

Neste estudo, utilizamos o valor crítico de 2.96, correspondente a um nível de confiança de 95%. Os coeficientes estimados das regressões para cada um dos ativos foram calculados e comparados com o valor de referência. Como as estatísticas

dos testes para ambos os coeficientes não são maiores em módulo do que o valor de referência, a hipótese nula de que as séries são não estacionárias não foi rejeitada. Isso significa que não há evidências suficientes para afirmar que as séries são estacionárias.

A tabela 3.1 mostra os coeficientes estimados das regressões para cada um dos ativos. A estatísticas dos testes para ambos coeficientes não são maiores em módulo do que o valor de referência, logo a hipótese nula de que as séries são não estacionárias não é rejeitada e podemos continuar com os testes para cointegração. Nesta fase do código, caso alguma das séries seja estacionária, ela é eliminada do estudo e não é testada com nenhuma outra série.

Tabela 3.1: Coeficientes da regressão de MRVE3 e PETR3 com seus lags

ticker	β	statistic
MRVE3	-0,516	-2,72
PETR3	0,563	-2,45

A tabela 3.2 mostra os parâmetros de avaliação do modelo, podemos perceber que apesar do baixo r^2 , o p-valor para o modelo é significativo.

Tabela 3.2: Avaliação do Modelo

ticker	r^2	statistic	p-value
MRVE3	0,0177	7,39	0,00683
PETR3	0,0144	6,00	0,0147

3.2.2 Testes de Cointegração

Após verificar que as séries de preços do par são não estacionárias, com ordem de integração igual a 1, prosseguiremos com o estudo checando a presença de cointegração entre as séries. Conforme dito anteriormente, essa etapa será feita através de um teste ADF dos resíduos de uma regressão linear ajustada entre as séries do par. Os coeficientes da regressão linear entre os ativos e dos resíduos com seu respectivo lag estão apresentados abaixo nas tabelas 3.3 e 3.4, respectivamente.

Tabela 3.3: Regressão Linear entre MRVE3 e PETR3

Parâmetro	β
Intercepto	0,2534
PETR3	0,6176

Agora, iremos ajustar uma regressão linear na série dos resíduos da última regressão e seu respectivo lag, para checar a estacionariedade dos resíduos.

A estatística do teste da regressão nos resíduos foi igual a -3.597, maior em módulo do que o valor de referência 2,96. Isso indica que há evidências de que a série de resíduos é estacionária e que o par MRVE3 e PETR3 é cointegrado.

Tabela 3.4: Regressão Linear para checar estacionariedade nos resíduos

Parâmetro	β
Intercepto	-0,000866
Lag	-0,469901

3.2.3 Backtesting

Após a verificação da presença de cointegração das séries no período de treino, iremos realizar um backtest da estratégia no período dentro da amostra, a fim de calcular o índice de Sharpe do par nesse período, conforme dito anteriormente, apenas os pares cointegrados com os 20 maiores Índices de Sharpe no período de treino irão ter a estratégia aplicada no período fora da amostra.

Para este estudo, não foram descontados os rendimentos de uma taxa livre de risco, logo o índice de Sharpe se refere apenas à rentabilidade líquida do investimento e sua volatilidade, sendo calculado por:

$$IS = \frac{\mu_{R_{a_1, a_2}}}{\sigma_{R_{a_1, a_2}}},$$

onde $\mu_{R_{a_1, a_2}}$ significa a média da rentabilidade líquida das operações referentes ao par dos ativos a_1, a_2 , e $\sigma_{R_{a_1, a_2}}$ o desvio padrão das rentabilidades líquidas dessas operações.

Primeiramente, calculamos a estatística $Z_{stat} = \frac{\mu_{res}}{\sigma_{res}^2}$ para normalizar os resíduos da regressão, onde μ_{res} é a média amostral dos resíduos e σ_{res}^2 o erro padrão. Assim, a série temporal Z_{stat} seguirá uma distribuição assintoticamente $N(0,1)$, e os pontos de entrada +2 e -2 representarão as caudas dessa distribuição normal, mostrando que quando os desvios atingirem tais pontos, estarão suficientemente distantes de sua média histórica. O gráfico 3.2 mostra a série formada pela estatística Z, destacando os pontos de entrada definidos. É possível observar 5 pontos de entrada, sendo 3 *long* e 2 *short*. A tabela 3.5 mostra um resumo de cada trade. Ainda houve mais um ponto de entrada short no dia 06/12/2012, porém não houve tempo hábil dentro da amostra para finalizar esse trade, logo, este foi excluído das análises. Podemos observar que todos os trades tiveram rentabilidade positiva, e o maior trade teve rentabilidade de 8,8%. A rentabilidade do trade é afetada pela distância em que foi aberta a posição. A volatilidade dos retornos foi medida através do índice de Sharpe e foi igual a 4,88.

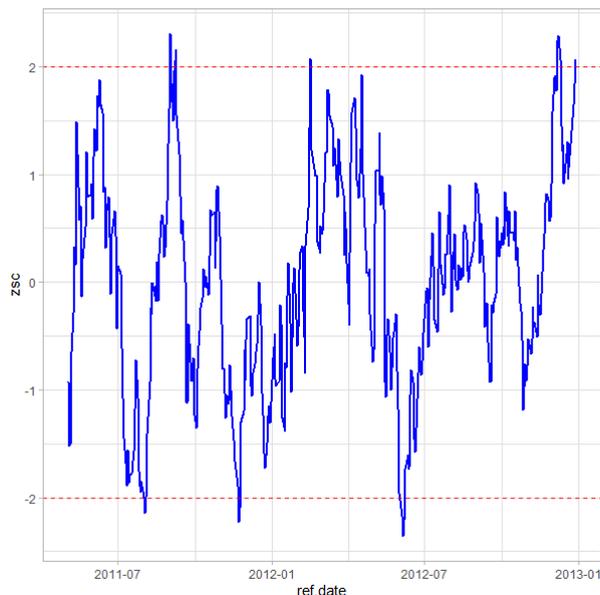


Figura 3.2: Resíduos Normalizados

Tabela 3.5: Regressão Linear para checar estacionariedade nos resíduos

Operação	Duração	retorno
long	7	0,0735
short	13	0,0549
long	10	0,0883
short	8	0,0622
long	25	0,0881

3.2.4 Testes fora da Amostra

Agora, a estratégia será aplicada em um período fora da amostra, correspondente aos 12 meses subsequentes ao período de treino, ou seja, de 01/01/2013 a 31/12/2013. Após a aplicação nesse período, os coeficientes serão reestimados e a cointegração entre os pares checada novamente na janela de 01/01/2013 a 31/12/2013. Na fase de teste fora da amostra, não iremos checar a cointegração entre os pares nem reestimar os coeficientes de sua relação, apenas iremos aplicar o modelo obtido no período de treino.

O gráfico 3.3 mostra a série formada pela estatística Z, destacando os pontos de entrada definidos. Durante o período de trading, houve apenas 1 ponto de entrada, para uma operação short. Podemos observar que, após cruzar o ponto de entrada em +2, a série de resíduos continua em uma tendência positiva, atingindo o ponto de "stop-loss", onde a operação foi encerrada com prejuízo. Tal fato evidencia que em operações de arbitragem estatística existem uma game de riscos que podem influenciar na rentabilidade.

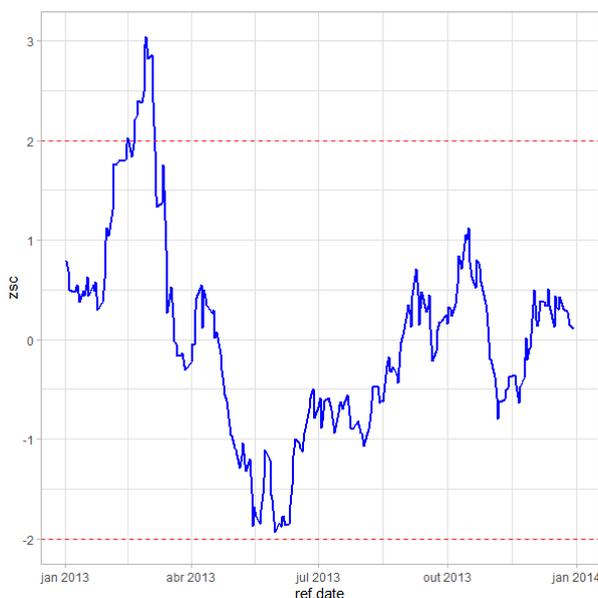


Figura 3.3: Resíduos Normalizados no Período de Teste

Tabela 3.6: Trades no período de teste

Operação	Duração	retorno
short	12	-0,0730

3.3 Resultados Gerais

Neste capítulo, iremos mostrar os resultados da estratégia em cada período de trading, bem como os pares selecionados pelo índice de sharpe para cada janela. Como um todo, a estratégia teve um retorno médio de 4,07% em 194 operações de 105 pares diferentes. A janela com maior rentabilidade média foi a de 2015-2016 (treino em 2015 e teste em 2016), com rentabilidade média de 11,83%. Os resultados por janela são mostrados na tabela 3.7.

Tabela 3.7: Rentabilidade Média por Janela

Janela	Rentabilidade Média
2012_2013	2,41 %
2013_2014	4,18 %
2014_2015	3,93 %
2015_2016	10,65 %
2016_2017	4,28 %
2017_2018	4,22 %
2018_2019	3,9 %
2019_2020	3,3 %
2020_2021	3,12 %
2021_2023	3,44 %

O gráfico 3.4 mostra que os resultados dos trades são descorrelacionados com o IBOVESPA, aderindo a propriedade "market-neutral" que uma estratégia de pairs-trading deve ter.

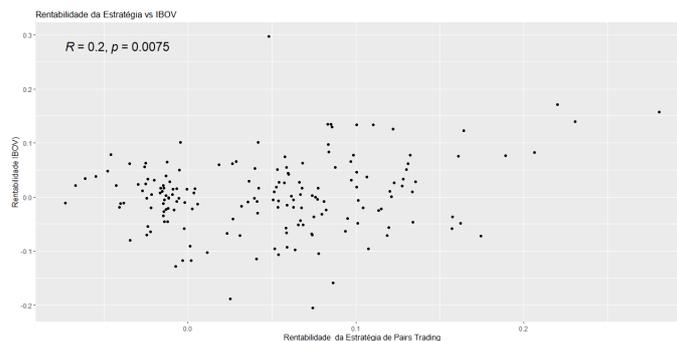


Figura 3.4: Gráfico de dispersão entre a Rentabilidade de cada Trade e do IBOV durante o período

3.3.1 Janela 2012-2013

Para a janela 2012-2013, apenas 13 dos 20 pares cointegrados com maior índice de sharpe no período de treino tiveram operações abertas no período de teste. Os índices de Sharpe calculados para o período de treino são mostrados na tabela 3.8.

A rentabilidade média das ações varia bastante entre os diferentes papéis. Alguns pares apresentaram rentabilidades bastante elevadas, como é o caso do par FLRY3 e NTCO3 e BEEF3 e NTCO3, enquanto outros apresentaram rentabilidades negativas, como MRVE3 e PETR3 e B3SA3 e WEGE3, é possível observar que alguns pares possuem empresas de mais de um setor, como é o caso da ENBR3 e EZTC3 e ENBR3 e RENT3.

Tabela 3.8: Índices de Sharpe no período de treino

CIEL3_GOAU4	79,25%
ENBR3_RENT3	68,21%
B3SA3_NTCO3	31,48%
CSAN3_RENT3	31,32%
FLRY3_NTCO3	20%
B3SA3_MULT3	18,64%
ENBR3_EZTC3	12,06%
ABEV3_EGIE3	9,87%
BEEF3_NTCO3	9,32%
CIEL3_GGBR4	9,29%
ABEV3_RENT3	9,18%
LREN3_RENT3	7,65%
B3SA3_WEGE3	6,75%
GOAU4_TOTS3	4,99%
MRVE3_PETR3	4,88%
MULT3_NTCO3	4,85%
CPFE3_JBSS3	4,8%
ECOR3_EZTC3	4,7%
FLRY3_SLCE3	4,65%
GGBR4_TOTS3	4,44%

Observando a tabela, podemos notar que as duas primeiras posições CIEL3 e GOAU4 e ENBR3 e RENT3 apresentam um índice de Sharpe significativamente maior em comparação aos demais pares de ativos, tais pares também apresentam uma rentabilidade média alta (não necessariamente a mais alta), o que sugere uma boa relação risco-retorno. Durante o período de teste, nem todos os pares tiveram operações abertas, pois o resíduo da regressão entre esses não atingiu os pontos de entrada, dos 20 pares selecionados, apenas 14 tiveram operações abertas no período de teste. O gráfico 3.5 mostra a rentabilidade média por par de ativos durante a janela de trade. Podemos perceber que dos 5 pares com menor índice de Sharpe (considerando o período de treino), 3 tiveram rentabilidade negativa, enquanto nos 5 pares com maior índice de Sharpe, todos tiveram rentabilidades positivas.

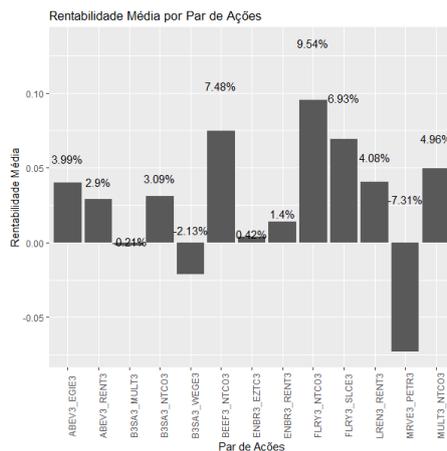


Figura 3.5: Rentabilidade dos Pares da Janela 2012-2013

3.3.2 Janela 2013-2014

Para a janela 2013-2014, dos 20 pares com maior índice de Sharpe no período de treino, 12 pares abriram operação no ano de 2014. Pode-se observar que os pares de ativos com os maiores índices de Sharpe são CCRO3 e TOTS3 e B3SA3 e ECOR3, pertencentes aos setores de Logística/Transporte e Serviços Financeiros/Saneamento, respectivamente. Esses pares apresentaram um desempenho excepcional em relação a volatilidade de seus retornos no período de treino, já os pares de ativos com os menores índices de Sharpe são BRKM5 e YDUQ3 e NTCO3 e WEGE3, ambos não abriram operações durante o período de treino. Os índices de Sharpe para essa janela são mostrados na tabela 3.9.

Tabela 3.9: Índices de Sharpe no período de treino

CCRO3_TOTS3	137,37%
B3SA3_ECOR3	122,54%
CIEL3_EGIE3	94,06%
BBDC4_CSAN3	39,91%
ABEV3_GOAU4	22,87%
SANB11_VALE3	16,08%
EMBR3_TOTS3	13,56%
CIEL3_UGPA3	12,94%
B3SA3_VIVT3	11,25%
SANB11_SMTO3	8,75%
SMTO3_VALE3	7,03%
CYRE3_NTCO3	6,54%
BBAS3_SANB11	6,32%
CSAN3_ECOR3	5,87%
BBDC3_SANB11	5,59%
HYPE3_SMTO3	5,22%
EGIE3_TIMS3	5,21%
PETR3_SANB11	5,11%
BRKM5_YDUQ3	4,96%
NTCO3_WEGE3	4,73%

A rentabilidade média varia consideravelmente entre os pares de ativos, variando de 0,89% a 8,75%, onde os pares de ativos com as maiores taxas de rentabilidade média são CIEL3_EGIE3 e B3SA3_VIVT3, pertencentes aos setores de Varejo/Energia Elétrica e Serviços Financeiros/Telecomunicações, respectivamente. O gráfico 3.6 mostra a rentabilidade média por par de ativos.

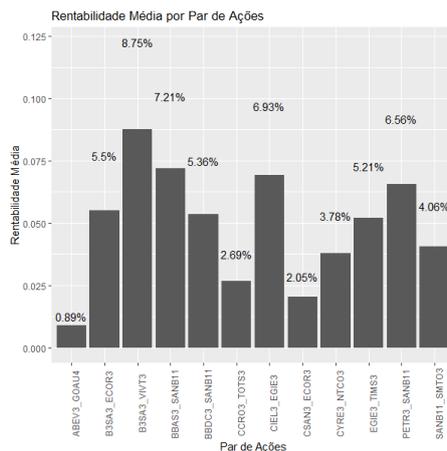


Figura 3.6: Rentabilidade dos Pares da Janela 2013-2014

3.3.3 Janela 2014-2015

Para a janela 2014-2015, tivemos operação para 13 pares, dos 20 que entraram na estratégia. O par que apresentou maior índice de Sharpe no período de treino foi composto pelos ativos CPFE3 e TAEE11, ambos pertencentes ao setor de energia elétrica, esse par, porém foi um dos poucos pares que apresentou rentabilidade negativa no período de teste, assim como o outro par formado por ações do mesmo setor, BBDC4 e ITUB4, referentes as ações preferenciais do Banco Bradesco e Banco Itaú. A tabela 3.10 mostra os pares que foram selecionados para o período de teste, baseados em seu desempenho durante o período de treino.

Tabela 3.10: Índices de Sharpe no período de treino

CPFE3_TAEE11	56,05%
EMBR3_ENGI11	53,38%
ENBR3_SANB11	24,78%
BBDC4_ITSA4	24,15%
MULT3_RENT3	24,05%
GOLL4_RADL3	18,92%
CCRO3_MULT3	16,28%
ITSA4_TAEE11	16,23%
SMT03_TAEE11	14,87%
EMBR3_RADL3	14,51%
ENGI11_JBSS3	11,58%
BBDC3_ITUB4	7,56%
ENBR3_SMT03	6,73%
CYRE3_SBSP3	5,92%
ENBR3_MRFG3	5,91%
B3SA3_ENBR3	5,39%
CPLE6_GOLL4	5,3%
SANB11_TAEE11	4,94%
GOLL4_TAEE11	4,9%
DXCO3_EZTC3	4,86%

A seguir, será mostrada a rentabilidade média por par de ativo, observe o par CPFE e TAEE11, CPFL Energia é uma empresa de distribuição de energia elétrica e Taesa é uma empresa de transmissão de energia elétrica, esse par foi um dos poucos onde houve rentabilidade negativa, o que pode indicar alguma correlação do setor, o que é indesejado para nossa análise, já o par de maior rentabilidade CPLE6 e GOLL4, também inclui uma empresa do ramo de energia elétrica, esse par teve rentabilidade líquida média de 17,91%. O gráfico 3.7 mostra a rentabilidade média de todos os pares que tiveram operações abertas no período de teste.

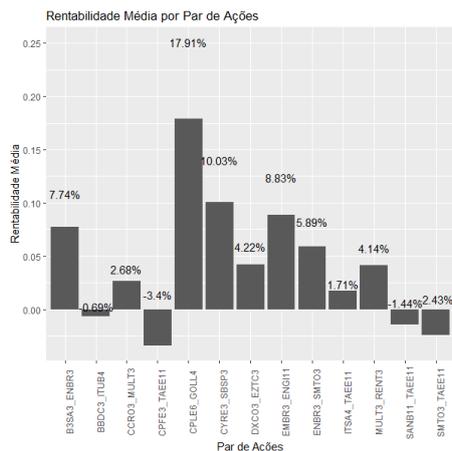


Figura 3.7: Rentabilidade dos Pares da Janela 2014-2015

3.3.4 Janela 2015-2016

Para a janela 2015-2016, tivemos operação para apenas 9 pares, dos 20 que entraram na estratégia. Podemos observar que as empresas do setor de construção, como EZTC3, estão presentes em dois dos pares com maiores índices de Sharpe: EZTC3 e USIM5 e MULT3 e VIVT3, sugerindo que o setor de construção pode ter apresentado boas oportunidades no período de treino, a mesma interpretação pode ser aplicada aos ativos como ITUB4 e ITSA4 (setor bancário), que foram cruzadas com os ativos YDUQ3 (setor educacional) e MULT3 (construção), que apresentaram índices de Sharpe relativamente altos. A tabela 3.11 mostra os pares selecionados para o período de teste.

Tabela 3.11: Índices de Sharpe no período de treino

MULT3_YDUQ3	198,61%
MULT3_VIVT3	159,69%
EZTC3_USIM5	21,37%
MULT3_RENT3	18,35%
PETR3_TOTS3	17,73%
ITSA4_MULT3	14,06%
RENT3_YDUQ3	13,32%
RENT3_VIVT3	11,97%
BBAS3_ITUB4	10,73%
ITSA4_ITUB4	10,42%
ALPA4_TOTS3	8,52%
EZTC3_PRIO3	7,44%
CMIG4_EZTC3	6,57%
ENGI11_RADL3	6,33%
ITUB4_YDUQ3	6,21%
FLRY3_UGPA3	5,78%
CCRO3_NTCO3	5,57%
CPFE3_ITUB4	5,49%
BBDC3_CYRE3	5,06%

Podemos observar que os ativos de construção e energia tiveram rentabilidades médias acima de 9%, com destaque para CMIG4 e EZTC3 e EZTC3 e USIM5 com rentabilidades superiores a 28%. Esses setores podem ter sido favorecidos por investimentos em infraestrutura e construção civil. Os ativos bancários ITSA4 e MULT3, ITUB4 e YDUQ3 e CPFE3 e ITUB4 tiveram rentabilidades médias abaixo de 6%, sugerindo um desempenho menos favorável, no entanto, ITSA4 e MULT3 apresentou um índice de Sharpe de 14,06%, indicando uma boa relação entre risco e retorno. O gráfico 3.8 mostra a rentabilidade média dos pares que abriram operação no período de teste.

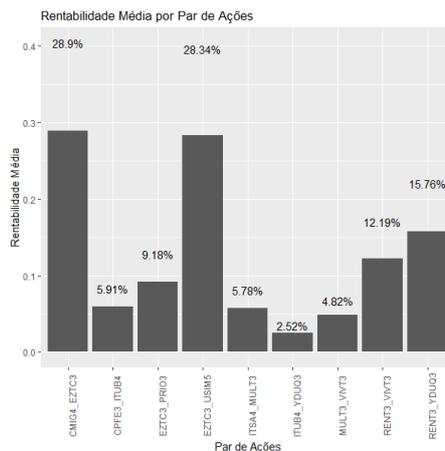


Figura 3.8: Rentabilidade dos Pares da Janela 2015-2016

3.3.5 Janela 2016-2017

Para a janela 2016-2017, tivemos operação para 11 pares, dos 20 que entraram na estratégia. Podemos observar que os pares com maior índice de Sharpe incluem o ativo CYRE3, referente a ações da empresa Cyrela, do ramo da construção civil. Esse ativo teve uma rentabilidade de aproximadamente 52% no ano de 2016, para o qual foi rodado o período de treino. A tabela 3.11 mostra os Índices de Sharpe dos pares selecionados para o período de trade.

Tabela 3.12: Índices de Sharpe no período de treino

CYRE3_EZTC3	705,52%
CYRE3_EMBR3	365,72%
CYRE3_RENT3	50,01%
BPAN4_SLCE3	49,26%
CPLE6_LREN3	43,51%
CYRE3_VVVT3	41,44%
BPAN4_COGN3	34,67%
COGN3_TAEE11	28,85%
MULT3_UGPA3	20,73%
TAEE11_VVVT3	18,28%
EQTL3_YDUQ3	16,97%
CYRE3_ECOR3	13,77%
CPFE3_VVIA3	13,54%
RADL3_TIMS3	13,42%
CSAN3_SLCE3	8,94%
GOAU4_SANB11	7,52%
ITSA4_PETR4	7,5%
CPLE6_TAEE11	7,08%

É possível observar que as ações do setor de energia elétrica tiveram uma rentabilidade média bastante positiva, com destaque para a combinação de ações da CPLE6 e TAEE11, que tiveram uma rentabilidade média de 6,75% e um índice Sharpe de 7,08%, também é interessante notar que as ações do setor de construção civil, representadas por pares como CYRE3 e EMBR3, CYRE3 e EZTC3 e CYRE3 e VVVT3, apresentaram uma rentabilidade média razoavelmente alta e um índice Sharpe muito elevado, mantendo o bom desempenho no período de teste. O gráfico 3.9 mostra a rentabilidade média de todos os pares com operações abertas no período de teste.

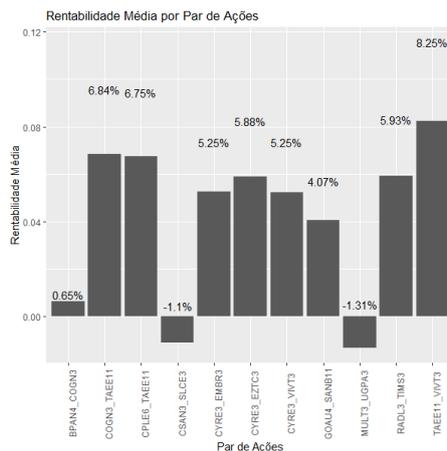


Figura 3.9: Rentabilidade dos Pares da Janela 2016-2017

3.3.6 Janela 2017-2018

Para a janela 2017-2018, tivemos operação para apenas 10 pares, dos 20 que entraram na estratégia. Os pares e seu respectivo Sharpe são mostrados na tabela 3.13. Nota-se que os pares que envolvem ações do setor imobiliário apresentaram bons índices de Sharpe, sendo EZTC3 e RENT3 o par com o melhor desempenho, seguida por ENGI11 e WEGE3 e RENT3 e VIIA3. O gráfico 3.10 mostra a rentabilidade por par de ativos.

Tabela 3.13: Índices de Sharpe no período de treino

EZTC3_RENT3	37,26%
ENGI11_WEGE3	34,87%
RENT3_VIIA3	30,39%
ENEV3_LREN3	22,41%
GOLL4_VVT3	21,69%
ECOR3_RENT3	17,75%
B3SA3_EZTC3	15,74%
CYRE3_SMT03	13,77%
ECOR3_LREN3	13,73%
ECOR3_RADL3	12,09%
CMIG4_ECOR3	11,17%
ENGI11_SLCE3	9,99%
DXCO3_VVT3	9,24%
TOTS3_VIIA3	8,27%
ENEV3_PETR3	8,27%
ENGI11_LREN3	7,68%
COGN3_MULT3	7,36%
TMS3_WEGE3	6,77%

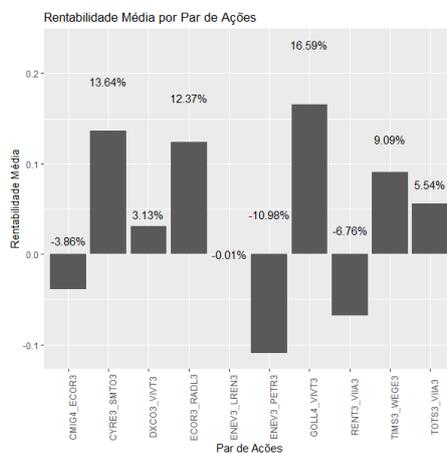


Figura 3.10: Rentabilidade dos Pares da Janela 2017-2018

3.3.7 Janela 2018-2019

Para a janela 2018-2019, tivemos operação para apenas 8 pares, dos 20 que entraram na estratégia. Observando a tabela 3.14 podemos notar que há uma grande variação nos valores de índice de Sharpe entre os pares de ações. O par CYRE3 e MULT3 apresenta o maior índice de Sharpe, com 81,09%, indicando que o retorno desse par no período de treino foi bastante expressivo em relação a volatilidade de seus retornos. Por outro lado, temos o par B3SA3 e LREN3, com índice de Sharpe de apenas 4,77%, indicando que seu retorno não foi tão expressivo em relação a volatilidade.

O gráfico 3.11 mostra a rentabilidade por par de ativos, os destaques são BEEF3 e VIIA3 com uma rentabilidade líquida de 17,51% e CSNA3 e ECOR3 com 20,65%. Já ABEV3 e VIIA3 e COGN3 e VIIA3 apresentam rentabilidades líquidas negativas de -3,39% e -3,55%, respectivamente. É interessante notar que temos a presença de VIIA3 tanto nos pares com pior rendimento na janela quanto em um dos pares que obteve melhor rendimento no período.

Tabela 3.14: Índices de Sharpe no período de treino

CYRE3_MULT3	81,09%
ABEV3_VIIA3	55,27%
CPLE6_ENGI11	24,01%
B3SA3_ITUB4	15,56%
JBSS3_LREN3	12,73%
RENT3_VIVT3	8,92%
BRAP4_PRIO3	8,08%
BEEF3_VIIA3	7,72%
COGN3_VIIA3	6,9%
CSNA3_ECOR3	6,77%
CYRE3_EQTL3	6,61%
EQTL3_ITSA4	6,5%
DXCO3_MULT3	5,2%
B3SA3_LREN3	4,77%

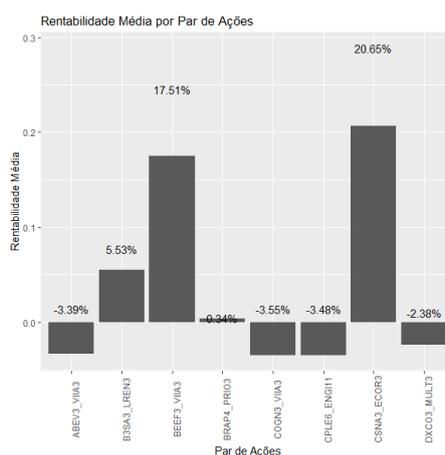


Figura 3.11: Rentabilidade dos Pares da Janela 2018-2019

3.3.8 Janela 2019-2020

Para a janela 2019-2020, tivemos operação para apenas 9 pares, dos 20 que entraram na estratégia. Os pares e seu respectivo Sharpe são mostrados na tabela 3.15 e o gráfico 3.12 mostra a rentabilidade média dos ativos que tiveram operações abertas no período de trade. Entre os pares que abriram operações, os destaques

são EGIE3 e TAEE11 com um índice de Sharpe muito alto de 353,26% e uma rentabilidade média de 10,78%, seguido de perto por ENBR3 e SBSP3 com um índice de Sharpe de 155,03% e uma rentabilidade média de 8,28%. Esses dois pares possuem ações referentes aos setores de energia elétrica e saneamento básico, respectivamente. Por outro lado, CMIG4 e CSNA3 apresenta um baixo índice de Sharpe de 11,25% e uma rentabilidade média negativa de -6,58%. Esse par pertence aos setores de energia elétrica e siderurgia, respectivamente.

Tabela 3.15: Índices de Sharpe no período de treino

EGIE3_TAEE11	353,26%
ENBR3_SBSP3	155,03%
CMIG4_CSNA3	11,25%
EQTL3_VIVT3	9,58%
ENEV3_WEGE3	7,32%
TAEE11_VIA3	6,99%
NTCO3_RENT3	6,83%
ENGI11_LREN3	6,77%
CSAN3_YDUQ3	6,53%
ITSA4_YDUQ3	6,52%
EGIE3_EQTL3	6,49%
SBSP3_VIA3	6,06%
ITUB4_RADL3	5,69%
CCRO3_EZTC3	4,67%
ITUB4_TAEE11	4,51%
ITSA4_TAEE11	4,49%
CCRO3_ITSA4	3,87%
CSAN3_ENEV3	3,49%
GGBR4_GOAU4	3,41%

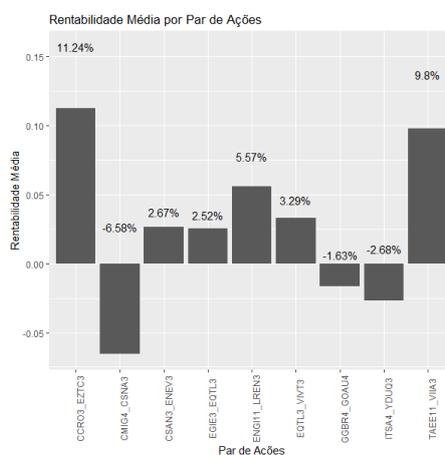


Figura 3.12: Rentabilidade dos Pares da Janela 2019-2020

3.3.9 Janela 2020-2021

Para a janela 2020-2021, tivemos operação para apenas 7 pares, dos 20 que entraram na estratégia. Os pares e seu respectivo Sharpe são mostrados na tabela 3.16 e as rentabilidades líquidas dos pares com operação aberta são mostrados no gráfico 3.13. Podemos observar que há uma diversidade de setores presentes na tabela, incluindo bancos (BBAS3, BBDC4), bebidas (ABEV3), energia elétrica (ELET3, ENBR3), varejo (LREN3), construção (MRVE3), agronegócio (SLCE3), entre outros. Isso mostra que não há um setor específico que se destacou no período. O par ITSA4 e TIMS3 possui o maior índice de Sharpe, indicando que ele obteve o melhor desempenho em relação ao risco assumido, além disso, a rentabilidade média desse

par é de 8,15%, o que é relativamente alta comparada com outros pares. Já os pares BBAS3 e YDUQ3 e MULT3 e PCAR3 tiveram rentabilidades médias acima de 10%, mas índices de Sharpe relativamente baixos, o que indica que esses pares podem ter sido mais arriscados do que outros pares com índices de Sharpe mais elevados.

Tabela 3.16: Índices de Sharpe no período de treino

ITSA4_TIMS3	365,57%
ELET3_NTCO3	67,64%
ITSA4_PCAR3	38,06%
ABEV3_ENGI11	26,84%
BBAS3_LREN3	19,95%
MULT3_PCAR3	18,3%
GGBR4_NTCO3	14,93%
LREN3_TIMS3	12,41%
ENBR3_TIMS3	12,35%
BBDC4_TIMS3	11,09%
BBAS3_ENGI11	9,02%
LREN3_PCAR3	8,58%
BBAS3_YDUQ3	7,87%
BRFS3_TIMS3	7,46%
MRVE3_NTCO3	6,72%
SLCE3_VALE3	6,38%

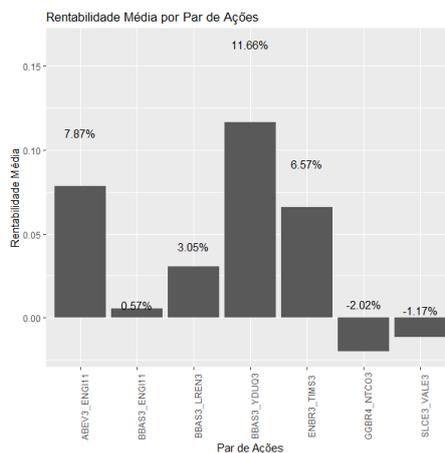


Figura 3.13: Rentabilidade dos Pares da Janela 2020-2021

3.3.10 Janela 2021-2023

Para a janela 2021-2023, tivemos operação para 13 pares, dos 20 que entraram na estratégia. Os pares e seu respectivo Sharpe são mostrados na tabela 3.17 e as rentabilidades médias dos pares que abriram operações são mostrados no gráfico 3.14. É possível notar que os pares com maior índice de Sharpe não tiveram as maiores rentabilidades médias. Por exemplo, o par B3SA3 e EZTC3 tiveram um índice de Sharpe mais alto do que o par CCRO3 e CIEL3, mas a rentabilidade média do par CCRO3 e CIEL3 é significativamente maior. No entanto, ainda podemos notar que a maioria dos pares com um alto índice de Sharpe também possui uma rentabilidade média razoavelmente alta, como EGIE3 e TOTS3, CCRO3 e YDUQ3 e CCRO3 e CSAN3. Esses pares pertencem a setores variados, incluindo energia, educação e infraestrutura. Além disso, alguns pares com um índice de Sharpe mais baixo apresentam rentabilidades médias negativas, como CCRO3 e CSAN3 e CPLE6 e MRFG3, pertencentes aos setores de infraestrutura e alimentos, respectivamente.

Tabela 3.17: Índices de Sharpe no período de treino

EGIE3_TOTS3	11,65%
B3SA3_EZTC3	10,4%
DXCO3_ECOR3	9,46%
DXCO3_RENT3	8,92%
CCRO3_CIEL3	8,71%
CCRO3_YDUQ3	7,33%
CCRO3_CSAN3	7,19%
TIMS3_VIVT3	7,06%
B3SA3_FLY3	5,99%
CCRO3_NTCO3	5,71%
CSAN3_VALE3	5,54%
CPLE6_MRFG3	5,53%
CCRO3_SANB11	5,28%
DXCO3_RADL3	5,25%
EZTC3_FLY3	4,98%
CCRO3_ITUB4	4,98%
BBDC4_LREN3	4,49%
ECOR3_LREN3	4,47%
ENEV3_EZTC3	4,4%
COGN3_FLY3	3,92%

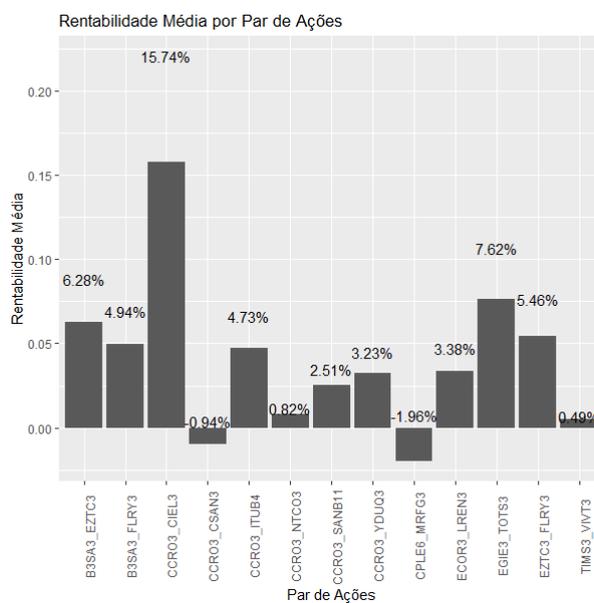


Figura 3.14: Rentabilidade dos Pares da Janela 2021-2023

4 Conclusão

Neste trabalho identificamos pares cointegrados, dentre um *pool* de papéis negociados na IBOVESPA. Utilizamos uma estratégia de *pairs trading* onde uma operação era aberta sempre que o resíduo da relação de cointegração ultrapassasse, em módulo, o threshold de dois desvios padrão. Os resultados obtidos foram bastante satisfatórios, resultando em rentabilidades médias positivas para todas as janelas em que foram testados os pares. Em alguns casos, houveram rentabilidades negativas que poderiam ter sido evitados caso fosse utilizado outras estratégias de stop-loss, onde vimos que em alguns casos, a estratégia rapidamente ultrapassou o ponto escolhido. Além disso, grande parte dos pares selecionados não puderam ser utilizados pois não houve oportunidade de operação na janela de trade, em alguns casos, como o da janela 2019-2020, apenas 7 pares dos 20 selecionados tiveram oportunidades de operação na janela escolhida. É interessante notar que a grande maioria dos pares não se repetiram entre as janelas escolhidas, mostrando a importância do rebalanceamento através das janelas móveis. Em diversos casos, os pares tinham relações fortes de cointegração até determinado ponto da amostra, perdendo tal relação nas janelas seguintes, provavelmente, caso não fosse utilizado o rebalanceamento dos pares, a rentabilidade dos pares seria menor e diversos pares não teriam sua relação de cointegração identificada.

Dos 105 pares selecionados, apenas 24 tiveram rentabilidade média líquida negativa, a janela com maior rentabilidade média foi a de 2015-2016, com rentabilidade líquida média de 10,65%, em 14 operações, muito influenciada por duas operações que obtiveram retorno líquido de aproximadamente 28%, sendo uma operação envolvendo EZTC3 e USIM5, com tempo total de 28 dias de operação e outra envolvendo CMIG4 e EZTC3, com 45 dias de operação.

Referências Bibliográficas

- Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*.
- Caldeira, J. (2011). Arbitragem estatística e estratégia long-short pairs trading, abordagem com cointegração aplicada a dados do mercado brasileiros. *Encontro de Economia da Região Sul-ANPEC SUL*.
- Caldeira, J. e Moura, G. (2013). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Available at SSRN 2196391*.
- Caldeira, J. e Portugal, M. (2010). Estratégia long-short, neutra ao mercado, e index tracking baseadas em portfólios cointegrados. *Revista Brasileira de Finanças*.
- Engle, R. e Granger, W. (1987). Co-integration and error correction: Representation, estimation, and testing. *The Econometric Society*.
- Herlemont, D. (2000). Pairs trading, convergence trading, cointegration. *YATS Working Papers*.
- Lucas, A. e Franses, P. (1997). Outlier robust cointegration analysis. <https://www2.math.ethz.ch/finance/>.
- Silva, F., Ziegelmann, F., e Caldeira, J. (2023). A pairs trading strategy based on mixed copulas. *The Quarterly Review of Economics and Finance*.
- Vidyamurthy, G. (2004). Pairs trading, quantitative methods and analysis. *Journal of Mathematical Finance*.
- Xie, W. e Wu, Y. (2013). Copula-based pairs trading strategy. *SSRN Electronic Journal*.