

Projeto Orientado em Computação II

Pesquisa mista de Clusterização dos Setores do Ibovespa

Estudante: Pedro Henrique Andrade¹
Prof. Orientador: Cristiano Arbex Valle¹

¹ Departamento de Ciências da Computação
Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG – Brasil

`pedro.handrade2000@gmail.com, arbex@dcc.ufmg.br`

Abstract. *Este relatório descreve o desenvolvimento de uma metodologia para a clusterização dos setores do Ibovespa, levando em consideração correlações entre os ativos e sensibilidades a variáveis macroeconômicas. A pesquisa é dividida em duas etapas principais: a primeira envolve a análise de correlação dos retornos dos ativos, enquanto a segunda foca na avaliação de métricas de risco e sensibilidade a variáveis macroeconômicas para uma clusterização refinada. O objetivo é facilitar a escolha de alocação de recursos de forma conjunta, proporcionando agilidade e clareza no processo de gestão de carteiras. Os resultados esperados incluem um modelo inovador que pode ser validado e utilizado para otimização de portfólios.*

1. Introdução

1.1. Caracterização do Problema

O mercado de ações é notoriamente dinâmico e imprevisível, onde as condições podem mudar drasticamente em um curto período. Essa volatilidade impõe desafios significativos na gestão de carteiras de ações, onde a capacidade de adaptar-se rapidamente às mudanças do mercado é crucial. Neste projeto, exploramos o desenvolvimento de uma metodologia de clusterização de setores dos ativos do Ibovespa, centrada na análise de correlações entre ativos e suas sensibilidades a variáveis macroeconômicas. Este enfoque reflete a necessidade de sistemas de gestão de portfólio que possam rapidamente se ajustar a novas informações de mercado e manter os riscos em níveis aceitáveis, garantindo assim a sustentabilidade e o sucesso a longo prazo.

1.2. Resumo POC I

Na POC I o desenvolvemos o "Otimizador de Carteira de Ações", uma ferramenta projetada para aprimorar a gestão de portfólios de ações, proporcionando escalabilidade e controle de risco em uma carteira de investimentos em ações. A metodologia do projeto foi dividida em duas etapas principais de otimização. A primeira etapa, a otimização setorial, foca na definição de scores de posicionamento para cada setor e na maximização de uma função objetiva, incorporando restrições como o Tracking Error variável. Essa abordagem mostrou resultados eficientes, oferecendo soluções rápidas e precisas para a alocação setorial. A segunda etapa, a otimização dos ativos, envolve a otimização individual de cada ativo, abordando os desafios relacionados ao elevado número de variáveis. Esta fase apresentou complexidades significativas, e as alternativas exploradas ao Tracking Error revelaram-se menos precisas. O backtesting com dados de um fundo de investimento confirmou a eficácia do modelo na otimização setorial, enquanto os resultados da otimização dos ativos indicaram a necessidade de investigação adicional em gestão de portfólios complexos. Este estudo contribui para o campo da gestão de portfólios, oferecendo insights sobre otimização em mercados dinâmicos e destacando áreas para pesquisa futura, especialmente na otimização de ativos sob condições complexas.

1.3. Motivação

Este estudo é motivado pela necessidade de proporcionar agilidade e clareza no processo de gestão de carteiras. Ao definir bem os setores de cada ativo, gestores de fundos e investidores individuais podem maximizar retornos, ao conseguir distinguir quais setores conseguem melhor performar em diferentes cenários de mercado. A clusterização de setores com características semelhantes pode facilitar a escolha de alocação de recursos de forma conjunta, oferecendo uma visão mais estratégica e informada.

1.4. Objetivos

Este projeto é guiado pelos seguintes objetivos detalhados:

1. Clusterizar os setores do Ibovespa levando em consideração correlação dos retornos e sensibilidade a variáveis macroeconômicas.
2. Ampliar o entendimento sobre o processo de gestão de portfólio e mercado financeiro.

3. Investigar a viabilidade de utilizar técnicas de clusterização avançadas para a identificação de setores com comportamentos similares.
4. Realizar backtesting dos clusters formados para validar a robustez e eficácia do modelo.

1.5. Breve Descrição da Estrutura do Trabalho

Este relatório está estruturado de forma a proporcionar uma compreensão clara e abrangente do projeto. Inicia-se com uma relação com trabalhos correlatos da literatura, oferecendo uma visão dos estudos anteriores e das teorias fundamentais relacionadas à clusterização e análise de dados financeiros. Após isso, a metodologia adotada é detalhada, explicando sobre a coleta e tratamento dos dados, as duas etapas distintas de clusterização e as técnicas específicas utilizadas. Segue-se uma análise crítica dos resultados obtidos, avaliando a eficácia do modelo desenvolvido e discutindo suas implicações práticas. O relatório conclui com as principais conclusões, discute as limitações encontradas no estudo e propõe possíveis direções para pesquisas futuras.

2. Trabalhos Correlatos

2.1. Clusterização em Finanças

A aplicação de técnicas de clusterização em finanças tem sido amplamente explorada para identificar grupos de ativos com comportamentos semelhantes. Estudos anteriores [1] mostraram que a análise de correlação entre os retornos dos ativos pode ser uma ferramenta eficaz para a formação de clusters, facilitando a diversificação e a gestão de riscos em portfólios.

2.2. Financial Networks

A teoria de redes financeiras [1] permite a representação gráfica das relações financeiras entre entidades ou ativos em um mercado. Essas redes facilitam a análise das interconexões e a dinâmica do mercado financeiro, permitindo a identificação de clusters de ativos com características similares.

3. Referencial Teórico

Para a compreensão aprofundada do problema em questão, é essencial familiarizar-se com determinados conceitos e abordagens existentes no mercado financeiro e ciências de dados. A seguir, apresentamos uma breve descrição de cada um desses conceitos e ferramentas:

3.1. Financial Network

Uma rede financeira é uma representação gráfica das relações financeiras entre entidades ou ativos em um mercado. Essas redes permitem a análise das interconexões e a dinâmica do mercado financeiro, facilitando o entendimento da estrutura de dependências e a propagação de riscos e informações.

3.2. Clusterização

Refere-se ao processo de agrupar objetos semelhantes em conjunto, de modo que objetos no mesmo grupo (ou cluster) sejam mais similares entre si do que com aqueles em outros grupos. Na gestão de carteiras, a clusterização pode ser usada para identificar setores ou grupos de ativos com comportamentos semelhantes em termos de retorno e risco.

3.3. Beta

É uma medida da sensibilidade dos retornos de um ativo em relação aos retornos do mercado como um todo. Um beta maior que 1 indica que o ativo é mais volátil do que o mercado, enquanto um beta menor que 1 sugere que o ativo é menos volátil.

3.4. Variáveis Macroeconômicas

São indicadores que refletem o estado da economia em uma escala ampla, incluindo taxas de juros, taxas de inflação, taxas de câmbio, PIB, entre outros. No contexto de gestão de carteiras, a sensibilidade dos ativos a essas variáveis é crucial para a tomada de decisões estratégicas.

3.5. K-means

É um método popular de clusterização que divide um conjunto de dados em K grupos distintos, com base na similaridade dos dados. O algoritmo aloca cada ponto de dados ao cluster cujo centroide (o ponto médio de todos os pontos no cluster) é o mais próximo.

4. Contribuições

Este capítulo descreve detalhadamente as atividades conduzidas ao longo do projeto, evidenciando a metodologia empregada e as etapas cruciais para a realização da clusterização dos setores do Ibovespa. A abordagem metodológica foi estruturada de maneira a abarcar tanto a teoria quanto a prática, englobando desde a coleta e tratamento de dados até a aplicação de técnicas de clusterização avançadas. Cada seção deste capítulo aborda uma fase específica do projeto, refletindo a progressão lógica e as contribuições significativas realizadas durante o estudo.

4.1. Pré-Clusterização

A fase de pré-clusterização do projeto focou na coleta e tratamento de dados essenciais para a análise subsequente. Inicialmente, os dados de preço foram adquiridos da Bloomberg, oferecendo informações atualizadas e precisas do mercado. Durante o tratamento dos dados, um cuidado especial foi tomado com papéis que tiveram seus tickers alterados, assegurando a consistência e a precisão da base de dados. Os ajustes para dividendos e splits, essenciais para uma análise financeira precisa, foram realizados utilizando as informações providas pela Bloomberg, garantindo que os dados refletissem os valores reais dos ativos.

Na fase de cálculo dos dados, as métricas fundamentais calculadas foram o retorno, log do retorno diário, correlação entre os ativos e uma matriz de distâncias pré-calculadas das correlações. Estes cálculos são cruciais para compreender as dinâmicas do mercado e para a posterior clusterização dos setores.

4.2. Clusterização Baseada em Correlações

A primeira etapa da clusterização envolveu a análise de correlação dos retornos dos ativos do Ibovespa. Para isso, foi aplicada a técnica de agrupamento hierárquico, utilizando o algoritmo de linkage de Ward e a distância euclidiana como medida de similaridade. Este método permitiu a formação de uma rede financeira preliminar, facilitando a visualização da estrutura de clusters e a identificação de setores com comportamentos de mercado semelhantes.

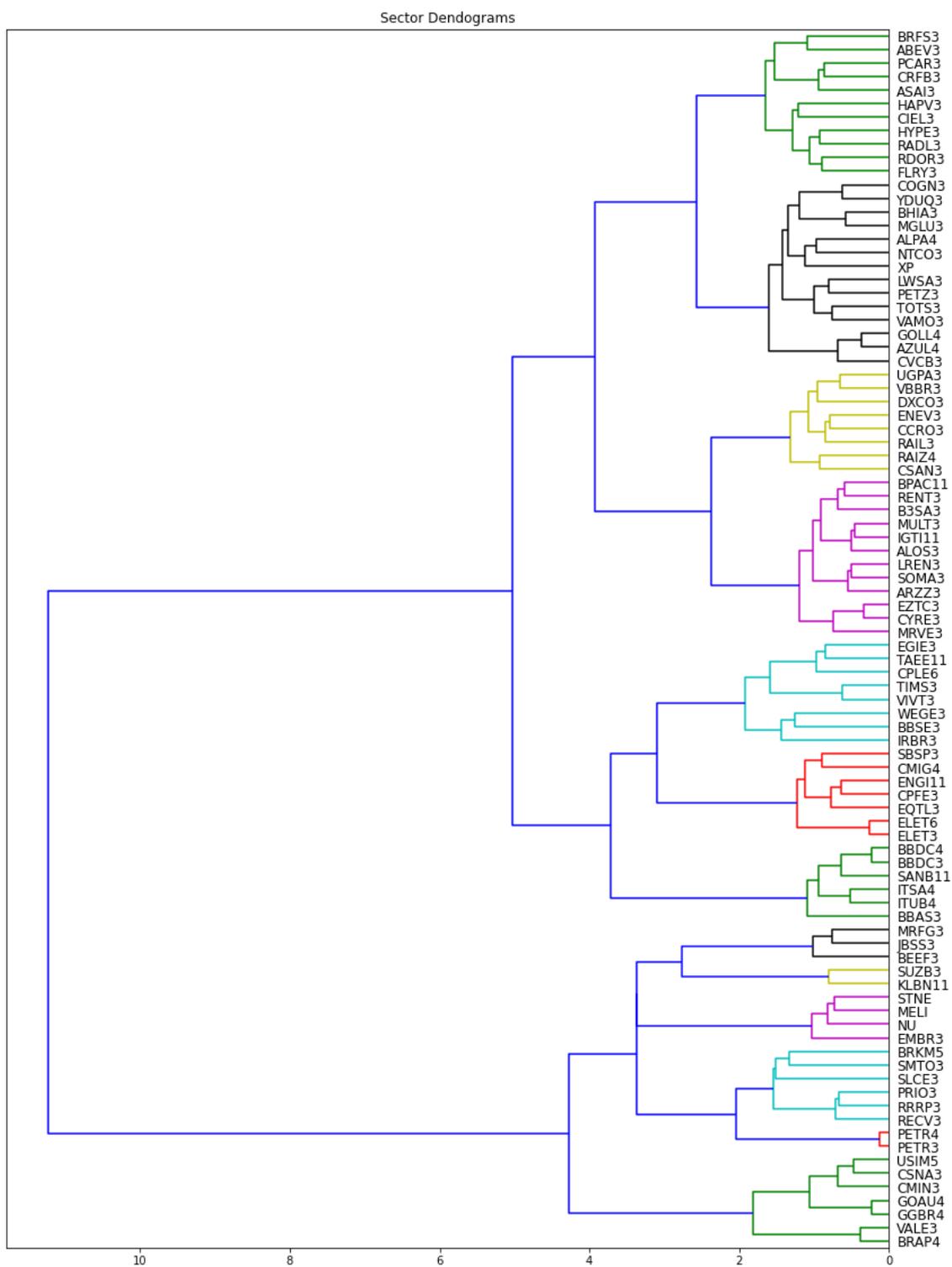


Figure 1. Dendograma da Clusterização de Correlações

Na primeira etapa, foram identificados clusters preliminares com base nas correlações de retorno, como pode ser visto no dendograma acima. Essa estrutura da árvore hierárquica facilita a identificação de setores com comportamentos de mercado semelhantes. Para chegar nesses resultados, foram utilizados dados de retornos de outubro de 2021 até out-

ubro de 2022, e um $K = 13$ clusters, usando o Hierarchical Cluster da biblioteca scikit-learn do Python[3]. Antes de passar para o modelo, foi precomputado uma matriz de distância das correlações, pois isso permite uma análise mais precisa das relações entre os ativos, ajudando a capturar as semelhanças e diferenças de maneira mais efetiva. Além disso, o número de clusters foi 13 pois é o que captura o "elbow" da curva plotada do score de Calinski Harabasz[4], como mostra a figura abaixo.

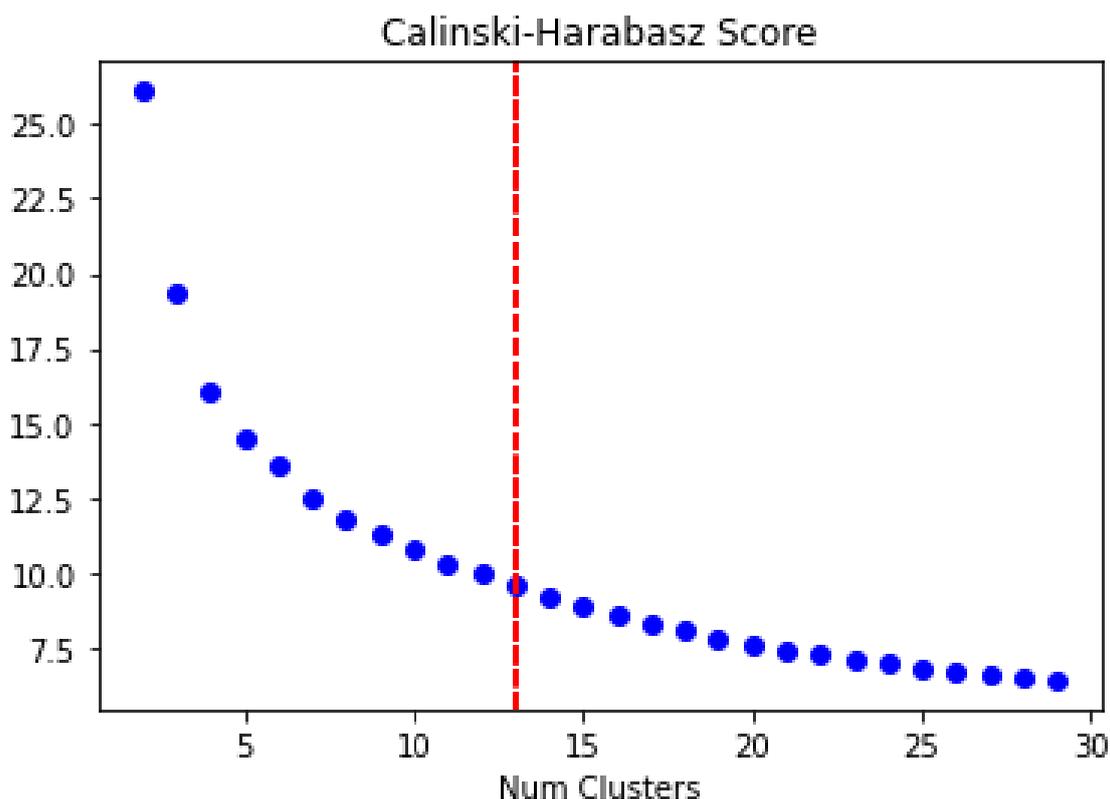


Figure 2. Calinski Harabasz Score

4.3. Clusterização Baseada em Sensibilidade e Risco

Na segunda etapa de clusterização, teve o objetivo de refinar os clusters identificados na primeira etapa, incorporando a análise de sensibilidade a variáveis macroeconômicas e métricas de risco. Esta abordagem visa proporcionar uma visão mais completa e detalhada dos setores do Ibovespa, considerando não apenas os retornos, mas também o perfil de risco e a reação dos ativos a mudanças macroeconômicas.

Primeiramente, foi realizado o cálculo da volatilidade (*Vol*) dos papéis, que é uma medida fundamental do risco de um ativo. Em seguida, o beta (*Beta*) dos papéis, que indica a sensibilidade dos retornos dos ativos em relação aos retornos do mercado como um todo.

Além disso, foi calculado a correlação de cada papel com várias variáveis macroeconômicas, incluindo:

- **Câmbio USDBRL:** Taxa de câmbio entre o dólar americano e o real brasileiro.
- **Taxa de Juros Nominal de 10 anos no Brasil:** Taxa de juros nominal para títulos de 10 anos.

- **Taxa de Juros Real de 10 anos no Brasil:** Taxa de juros real ajustada pela inflação para títulos de 10 anos.
- **Índice de Commodities:** Índice que reflete os preços das commodities.

O output da etapa anterior, que consistia nos clusters baseados na correlação entre os ativos, foi utilizado como input para esta nova etapa de clusterização. Esta integração permite que os clusters iniciais sejam refinados com base em informações adicionais de risco e sensibilidade macroeconômica.

Para padronizar os dados e garantir que todas as variáveis estivessem no mesmo range, foi aplicado a técnica de *MinMaxScaler*. Este método normalizou os dados, ajustando os valores para um intervalo entre 0 e 1, o que é crucial para garantir que nenhuma variável domine o processo de clusterização.

Com os dados preparados e normalizados, utilizamos o algoritmo de *K-means* da biblioteca *scikit-learn*, definindo o número de clusters $K = 13$. Esta escolha foi baseada na análise do "elbow" da curva de inércia[2].

Os resultados dessa etapa mostraram um refinamento significativo dos clusters, proporcionando insights mais detalhados sobre a estrutura do mercado e os comportamentos dos setores em diferentes condições macroeconômicas. A integração das métricas de risco e sensibilidade macroeconômica permitiu a formação de clusters mais robustos e informativos, que podem ser utilizados para uma alocação de recursos mais estratégica e eficiente.

Estes resultados confirmam a viabilidade da abordagem e oferecem uma base sólida para a gestão de portfólios em diferentes cenários de mercado. A próxima etapa do projeto envolveu a validação dos clusters identificados através de backtesting, garantindo que a abordagem seja não apenas teoricamente robusta, mas também aplicável e eficaz em condições de mercado realistas.

5. Resultados

Os resultados finais do projeto apresentaram avanços significativos na clusterização dos setores do Ibovespa.

5.1. Clusterização Baseada em Correlações

Na análise dos clusters preliminares, observou-se uma clara distinção entre setores com características semelhantes, como pode ser visualizado no dendograma apresentado anteriormente. Essa estrutura hierárquica confirma a eficácia da técnica de agrupamento hierárquico utilizada e fornece uma base sólida para a segunda etapa de clusterização.

5.2. Clusterização Baseada em Sensibilidade e Risco

A segunda etapa de clusterização refinou os resultados da primeira fase, incorporando a análise de sensibilidade a variáveis macroeconômicas e métricas de risco. A aplicação do K-means permitiu a identificação de clusters mais refinados, considerando não apenas os retornos, mas também o perfil de risco e a reação dos ativos a mudanças macroeconômicas. Estes resultados confirmam a viabilidade da abordagem.

Nesta etapa, foram verificados as sensibilidades de cada setor predefinido e o backtest desses setores em alguns acontecimentos históricos no mercado de ações. Além disso,

destaca-se que houve um ajuste fino dos setores, resultando em novos nomes como Bancos, CM HB, CM LB, Defenders, Strikers, Yield, Domésticos HB, Domésticos LB, Off-shore, Cap Goods e Outros. Esse ajuste foi necessário devido a idiosincrasias dos papéis, causadas por movimentos separados dos ativos. Também foram incluídos alguns papéis listados fora do Brasil, mas com operações locais, para analisar seu comportamento. Esses papéis foram NU, STNE, MELI e XP.

5.3. Sensibilidades dos Setores

Os gráficos de radar a seguir ilustram as sensibilidades dos diferentes setores com base nas variáveis analisadas, incluindo volatilidade (*Vol*), beta (*Beta*), câmbio USDBRL, taxas de juros nominais e reais de 10 anos no Brasil, e o índice de commodities. Esses gráficos oferecem uma visão clara das diferenças e similaridades entre os setores.

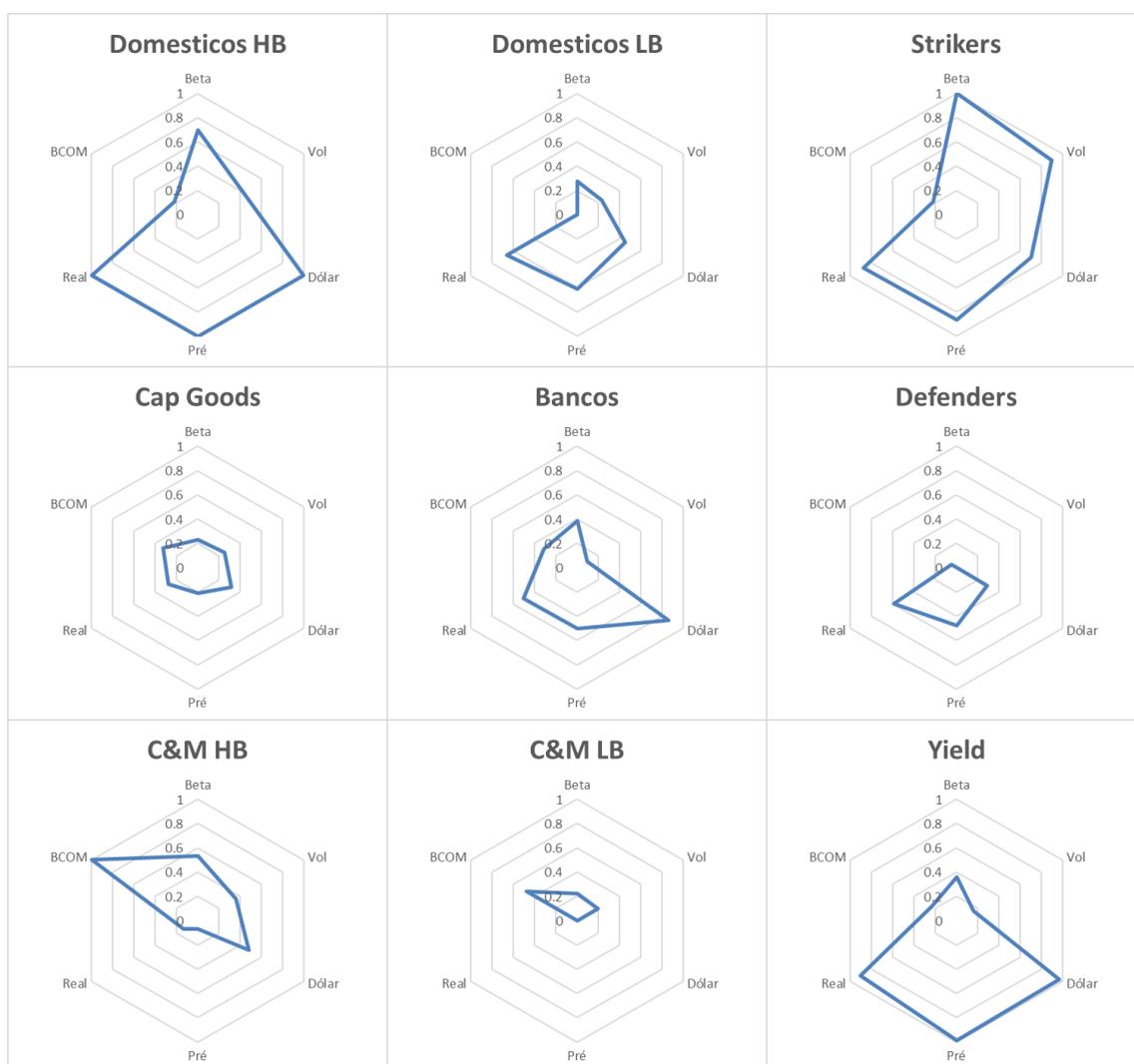


Figure 3. Gráficos de Radar: Sensibilidades dos Setores

5.4. Desempenho em Cenários Históricos

Os gráficos de backtest foram utilizados para validar a robustez dos setores identificados em diferentes cenários históricos. A análise incluiu períodos de crise, recuperação, e

eventos significativos no mercado financeiro. Os resultados dos backtests mostram como os setores performaram em cada cenário, fornecendo insights valiosos para a gestão de portfólios.

Rally pós 2008		Retomada covid		Fim do ciclo de alta de juros		Rally pós impeachment	
Mar 2009 - Jan 2010		Mai 2020 - Jul 2020		Jul 2022 - Ago 2022		Jan 2016 - Out 2016	
Dom LB	247.0%	Strikers	89.8%	Offshore	51.7%	Strikers	245.5%
Offshore	207.0%	Offshore	60.6%	Strikers	33.1%	C&M HB	242.5%
Dom HB	205.0%	Cap Goods	58.4%	Cap Goods	25.8%	Yield	108.9%
Strikers	169.0%	SMAL11	47.9%	Bancos	22.1%	Bancos	86.4%
SMAL11	165.0%	C&M HB	46.2%	Dom HB	20.9%	Offshore	75.3%
C&M LB	138.0%	Dom HB	44.1%	SMAL11	20.3%	IBOV	70.8%
C&M HB	107.0%	Yield	36.1%	C&M HB	18.6%	Dom LB	68.4%
Bancos	92.2%	Bancos	32.4%	IBOV	18.1%	SMAL11	68.2%
IBOV	87.5%	IBOV	31.9%	Yield	12.8%	Dom HB	65.3%
Cap Goods	83.8%	Dom LB	21.8%	Dom LB	11.3%	Defenders	61.1%
Yield	69.6%	Defenders	15.8%	C&M LB	10.3%	C&M LB	44.9%
Defenders	58.3%	C&M LB	6.3%	Defenders	3.4%	Cap Goods	-10.9%

Figure 4. Backtest: Cenários de alta (Equal Weighted)

Crise 2008	Covid	Política Monetária Restritiva		Caminhoneiros	Joesley	Trump					
Mai 2008 - Out 2009	Fev 2020 - Mar 2020	Jun 2021 - Dez 2021		Mai 2018	Mai 2017	Nov 2016 - Dez 2016					
Defenders	-6.4%	Dom LB	-19.3%	Cap Goods	1.8%	Offshore	8.2%	Cap Goods	0.0%	C&M LB	4.0%
Dom LB	-26.7%	Defenders	-27.7%	Defenders	0.1%	Cap Goods	-1.2%	Offshore	-2.7%	C&M HB	0.2%
Cap Goods	-31.1%	C&M LB	-36.7%	C&M LB	-8.3%	C&M LB	-1.4%	C&M LB	-3.7%	Dom LB	-6.6%
Yield	-31.6%	Yield	-44.4%	C&M HB	-14.5%	Defenders	-2.1%	Dom LB	-4.1%	IBOV	-7.9%
Bancos	-32.6%	Cap Goods	-44.7%	Yield	-18.5%	Dom LB	-2.8%	Defenders	-7.2%	Offshore	-8.7%
IBOV	-44.9%	Bancos	-44.8%	IBOV	-22.9%	SMAL11	-4.5%	IBOV	-8.7%	Dom HB	-9.9%
Dom HB	-48.4%	IBOV	-45.4%	Bancos	-26.2%	Bancos	-4.7%	SMAL11	-9.9%	Yield	-10.3%
C&M LB	-53.3%	Offshore	-46.9%	Dom LB	-29.9%	Yield	-5.9%	Dom HB	-10.0%	Bancos	-10.7%
Strikers	-54.3%	SMAL11	-49.1%	SMAL11	-29.9%	Dom HB	-6.2%	Yield	-10.9%	Defenders	-10.9%
C&M HB	-54.7%	Dom HB	56.5%	Dom HB	-34.7%	Strikers	-6.2%	Bancos	-11.9%	Cap Goods	-12.1%
Offshore	-58.3%	C&M HB	-63.2%	Strikers	-45.7%	IBOV	-7.8%	C&M HB	-11.9%	SMAL11	-13.6%
		Strikers	-67.0%	Offshore	-54.0%	C&M HB	-17.4%	Strikers	-12.9%	Strikers	-14.7%

Figure 5. Backtest: Cenários de baixa (Equal Weighted)

5.4.1. Análise Detalhada dos Resultados

A clusterização baseada em sensibilidade e risco, complementada com ajustes manuais e a inclusão de ativos offshore, permitiu uma análise detalhada das características dos setores do Ibovespa. Os resultados dos gráficos de radar e dos backtests confirmam a eficácia da metodologia adotada, proporcionando uma visão clara e detalhada das sensibilidades dos setores e seu desempenho em diferentes cenários de mercado.

Esta abordagem oferece uma ferramenta poderosa para gestores de fundos e investidores, permitindo uma alocação de recursos mais estratégica e baseada em dados robustos e análise detalhada.

6. Conclusões e Trabalhos Futuros

6.1. Conclusões

O projeto demonstrou sucesso significativo na clusterização dos setores do Ibovespa, com a abordagem baseada em correlações e sensibilidade a variáveis macroeconômicas mostrando-se eficaz. A utilização de técnicas avançadas de clusterização permitiu a identificação de setores com comportamentos semelhantes, proporcionando uma visão mais clara e estratégica para a gestão de carteiras de ações. Os resultados obtidos até o momento confirmam a validade da metodologia adotada e oferecem uma base sólida para a continuação do projeto.

6.2. Trabalhos Futuros

Diante dos desafios encontrados, sugere-se para futuras pesquisas:

- Refinar os métodos de clusterização, incorporando análises mais profundas e variadas, incluindo novos indicadores econômicos.
- Explorar novas métricas de risco que possam oferecer um equilíbrio entre precisão e facilidade de cálculo em portfólios de grande escala.
- Realizar estudos de caso em diferentes mercados e contextos para avaliar a aplicabilidade e eficácia da metodologia em cenários variados.

Este projeto abriu novas perspectivas na clusterização de setores do Ibovespa e estabeleceu uma base sólida para futuras investigações no campo da gestão de investimentos. A continuidade dessas pesquisas é essencial para desenvolver estratégias de otimização mais robustas e adaptáveis às dinâmicas complexas do mercado financeiro.

7. Dados Finais

Ticker	Vol	Beta	USDBRL	Juros Nominais	Juros Reais	Commodities	Cluster	Setor
TAE11	0.17	0.34	-0.26	-0.29	-0.29	0.01	4	Defenders
CPLE6	0.30	0.67	-0.32	-0.29	-0.29	0.05	4	Defenders
EGIE3	0.18	0.44	-0.41	-0.37	-0.33	0.02	1	Yield
CPFE3	0.27	0.66	-0.36	-0.34	-0.31	0.03	1	Yield
ENGI11	0.27	0.67	-0.39	-0.40	-0.37	-0.02	1	Yield
EQTL3	0.25	0.70	-0.37	-0.35	-0.32	0.03	1	Yield
ELET6	0.32	0.91	-0.43	-0.39	-0.33	0.01	1	Yield
CMIG4	0.29	0.91	-0.47	-0.43	-0.37	0.07	1	Yield
SBSP3	0.33	0.92	-0.41	-0.39	-0.33	0.05	1	Yield
ELET3	0.34	1.06	-0.46	-0.41	-0.35	0.04	1	Yield
ALPA4	0.55	1.38	-0.29	-0.27	-0.26	0.00	2	Domesticos HB
NTCO3	0.60	1.40	-0.29	-0.29	-0.27	-0.01	2	Domesticos HB
XP	0.58	1.66	-0.42	-0.33	-0.28	0.05	7	Domesticos HB
TOTS3	0.39	1.19	-0.46	-0.35	-0.34	0.07	1	Outros
COGN3	0.53	1.42	-0.35	-0.35	-0.40	0.05	7	Strikers
PETZ3	0.61	1.59	-0.37	-0.34	-0.36	0.03	7	Strikers
YDUQ3	0.63	1.59	-0.37	-0.37	-0.40	0.05	7	Strikers
BHIA3	0.72	1.77	-0.31	-0.35	-0.37	0.02	7	Strikers
LWSA3	0.73	1.87	-0.37	-0.34	-0.33	0.05	7	Strikers
AZUL4	0.75	2.07	-0.43	-0.41	-0.35	-0.01	7	Strikers
MGLU3	0.78	2.09	-0.34	-0.30	-0.33	0.07	7	Strikers
CVCB3	0.78	2.10	-0.40	-0.40	-0.36	0.08	7	Strikers
GOLL4	0.70	2.18	-0.50	-0.45	-0.39	0.03	7	Strikers
WEGE3	0.31	0.73	-0.37	-0.25	-0.22	0.10	9	Cap Goods
VIVT3	0.22	0.41	-0.22	-0.21	-0.24	-0.03	4	Defenders
BBSE3	0.24	0.43	-0.30	-0.24	-0.22	0.00	4	Defenders
TIMS3	0.24	0.46	-0.29	-0.26	-0.27	-0.02	4	Defenders
IRBR3	0.67	1.17	-0.19	-0.19	-0.17	0.02	2	Outros
BRKM5	0.49	1.05	-0.21	-0.12	-0.14	0.16	6	CM HB
PRI03	0.46	1.11	-0.33	-0.08	-0.09	0.42	5	CM HB
RECV3	0.49	1.17	-0.40	-0.22	-0.16	0.38	5	CM HB
RRRP3	0.54	1.32	-0.32	-0.16	-0.13	0.39	5	CM HB
SLCE3	0.34	0.57	-0.30	-0.14	-0.09	0.35	10	CM LB
SMT03	0.40	0.95	-0.33	-0.22	-0.17	0.20	9	CM LB
GOAU4	0.33	0.97	-0.35	-0.15	-0.10	0.32	10	CM HB
GGBR4	0.35	0.97	-0.33	-0.15	-0.10	0.30	10	CM HB
CMIN3	0.43	1.19	-0.33	-0.15	-0.16	0.25	5	CM HB
USIM5	0.42	1.27	-0.30	-0.17	-0.12	0.24	5	CM HB
CSNA3	0.49	1.53	-0.37	-0.13	-0.09	0.33	5	CM HB
BRAP4	0.33	0.88	-0.29	-0.06	0.01	0.39	10	CM LB
VALE3	0.35	0.91	-0.23	-0.05	0.01	0.35	10	CM LB
SANB11	0.28	0.87	-0.36	-0.24	-0.22	0.07	12	Bancos
BRFS3	0.57	1.27	-0.29	-0.25	-0.23	-0.07	2	CM LB
HAPV3	0.71	1.65	-0.26	-0.29	-0.27	-0.02	2	Domesticos HB
RDOR3	0.44	1.32	-0.44	-0.39	-0.41	0.06	9	Domesticos HB

Ticker	Vol	Beta	USDBRL	Juros Nominais	Juros Reais	Commodities	Cluster	Setor
ABEV3	0.23	0.53	-0.27	-0.25	-0.19	-0.07	4	Domesticos LB
CRFB3	0.41	1.04	-0.35	-0.32	-0.28	0.08	9	Domesticos LB
PCAR3	0.48	1.18	-0.32	-0.30	-0.30	0.10	9	Domesticos LB
RADL3	0.27	0.65	-0.34	-0.36	-0.33	-0.03	12	Domesticos LB
HYPE3	0.28	0.77	-0.32	-0.29	-0.34	0.04	12	Domesticos LB
ASAI3	0.39	0.81	-0.24	-0.26	-0.26	-0.02	12	Domesticos LB
FLRY3	0.36	0.90	-0.34	-0.29	-0.33	-0.04	12	Domesticos LB
JBSS3	0.33	0.50	-0.16	-0.07	-0.06	0.01	6	CM LB
BEEF3	0.42	0.65	-0.18	-0.13	-0.14	0.07	6	CM LB
MRFG3	0.44	0.74	-0.19	-0.16	-0.15	-0.02	6	CM LB
ITSA4	0.23	0.88	-0.48	-0.34	-0.30	0.10	12	Bancos
ITUB4	0.26	0.91	-0.40	-0.26	-0.24	0.08	12	Bancos
BBDC3	0.29	0.97	-0.41	-0.27	-0.27	0.10	12	Bancos
BBAS3	0.30	0.97	-0.44	-0.32	-0.27	0.05	12	Bancos
BBDC4	0.32	1.04	-0.35	-0.24	-0.24	0.07	12	Bancos
CIEL3	0.44	0.98	-0.36	-0.28	-0.24	0.10	9	Bancos
EMBR3	0.43	0.77	-0.22	-0.13	-0.12	0.07	6	Cap Goods
MELI	0.63	1.18	-0.27	-0.18	-0.16	0.09	8	Offshore
NU	0.77	1.24	-0.20	-0.18	-0.17	0.05	8	Offshore
STNE	0.84	1.81	-0.29	-0.23	-0.18	0.04	8	Offshore
SUZB3	0.30	0.31	0.01	0.06	0.05	0.10	11	CM LB
KLBN11	0.26	0.40	-0.10	0.01	-0.01	0.13	11	CM LB
PETR4	0.38	1.15	-0.32	-0.22	-0.15	0.21	0	CM HB
PETR3	0.39	1.18	-0.31	-0.21	-0.14	0.26	0	CM HB
ARZZ3	0.40	1.28	-0.42	-0.41	-0.41	0.04	3	Domesticos HB
RENT3	0.37	1.31	-0.50	-0.47	-0.48	0.05	3	Domesticos HB
B3SA3	0.40	1.39	-0.48	-0.41	-0.41	0.10	3	Domesticos HB
BPAC11	0.40	1.45	-0.45	-0.40	-0.40	0.06	3	Domesticos HB
CYRE3	0.46	1.48	-0.47	-0.51	-0.49	0.03	3	Domesticos HB
EZTC3	0.48	1.49	-0.41	-0.46	-0.44	-0.02	3	Domesticos HB
LREN3	0.46	1.51	-0.43	-0.39	-0.40	0.05	3	Domesticos HB
MRVE3	0.55	1.55	-0.38	-0.43	-0.45	0.00	3	Domesticos HB
SOMA3	0.51	1.64	-0.47	-0.44	-0.46	0.05	3	Domesticos HB
MULT3	0.34	1.10	-0.47	-0.45	-0.43	0.07	3	Yield
ALOS3	0.35	1.18	-0.48	-0.51	-0.51	0.06	3	Yield
IGTI11	0.37	1.20	-0.47	-0.43	-0.45	0.06	3	Yield
CSAN3	0.34	1.25	-0.51	-0.37	-0.37	0.27	3	CM LB
UGPA3	0.38	1.25	-0.47	-0.37	-0.39	0.15	3	Domesticos HB
VBBR3	0.37	1.28	-0.48	-0.40	-0.39	0.10	3	Domesticos HB
RAIZ4	0.45	1.30	-0.36	-0.27	-0.29	0.16	0	Domesticos HB
DXCO3	0.44	1.36	-0.43	-0.38	-0.36	0.09	3	Domesticos HB
VAMO3	0.49	1.50	-0.43	-0.42	-0.42	0.10	3	Strikers
CCRO3	0.31	0.94	-0.48	-0.40	-0.44	0.06	3	Yield
RAIL3	0.32	0.95	-0.42	-0.37	-0.39	0.10	3	Yield

Ticker	Vol	Beta	USDBRL	Juros Nominais	Juros Reais	Commodities	Cluster	Setor
ENEV3	0.34	0.96	-0.39	-0.38	-0.38	0.07	3	Yield

References

- [1] Gautier Martia Frank Nielsenc Mikołaj Bińkowski Philippe Donnatb. ““A review of two decades of correlations, hierarchies, networks and clustering in financial markets””. In: *Arxiv* vol. 7.nº 1 (November 3, 2020). DOI: <https://doi.org/10.48550/arXiv.1703.00485>.
- [2] *Inertia K-means*. URL: https://www.w3schools.com/python/python_ml_k-means.asp#:~:text=The%20elbow%20method%20lets%20us,K%20based%20on%20our%20data. (visited on 07/28/2024).
- [3] *Scikit-learn Documentation*. URL: <https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering> (visited on 07/28/2024).
- [4] *Scikit-learn Documentation*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.calinski_harabasz_score.html (visited on 07/28/2024).