

Uma Análise Preliminar da Dinâmica Volatilidade-Retorno de Ações das Grandes Empresas Utilizando Técnicas de Aprendizado Não-Supervisionado

Camila C. Figueiredo¹, Juliana V. Valerio², Giseli R. Lopes³

Instituto de Computação, Universidade Federal do Rio de Janeiro (UFRJ), RJ

Laura O. F. Moraes⁴

Centro de Ciências Exatas e Tecnologia, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), RJ

O estudo aplica a Análise de Componentes Principais (PCA) aos históricos de preços de abertura e fechamento das 500 maiores empresas listadas na bolsa[2], considerando dados de $D = 1032$ dias. A escolha do PCA, como discutido por Abdi e Williams [1], se justifica pela capacidade da técnica de reduzir a dimensionalidade, destacando os principais padrões de variância. O conjunto de dados aplicado no PCA contém o retorno diário das empresas ($E = 500$), e na matriz reduzida, de 1032 para 5 colunas, foi adicionado o vetor de volatilidade, calculado com o desvio padrão do retorno diário. O objetivo é identificar padrões que permitam agrupar as empresas em clusters de acordo com categorias de risco: conservador — baixo retorno e baixa volatilidade —, moderado — retorno e volatilidade médios — e arrojado — alto retorno e alta volatilidade.[3].

O retorno diário foi calculado pela diferença entre preços de fechamento e abertura, e a volatilidade pelo desvio padrão do retorno. Após o PCA[4], os dados foram agrupados com K-Means, identificando clusters de empresas com comportamentos similares.

A Figura 1 mostra a distribuição dos *clusters* em termos de retorno e volatilidade, enquanto a Figura 2 apresenta uma análise mais detalhada com um *boxplot*. Observa-se que os *clusters* azul (0) e verde (2) têm distribuições equilibradas próximas de 1, indicando balanço em relação à mediana. Já o *cluster* laranja (1) é assimétrico, com concentração em uma ponta, sugerindo maior instabilidade dessas empresas.

Ambos os três *clusters* apresentam uma distribuição assimétrica em relação a volatilidade, então não existe exatamente uma média desses valores. Porém, é importante ressaltar que o *cluster* verde (2) se destaca por apresentar uma assimetria menor que a média quando comparada aos outros dois.

O *cluster* verde (2) apresenta retorno positivo e baixa volatilidade, sendo mais adequado para investidores conservadores, que preferem evitar investimentos arriscados e com grandes variações. Já o azul (0) tem um retorno ligeiramente maior que o *cluster* verde, mas com alta volatilidade, atraindo investidores moderados — aqueles que buscam retornos mais elevados, mas com uma menor disposição para correr riscos. Por fim, o *cluster* laranja (1) oferece altos retornos e alta volatilidade, sendo recomendado para investidores mais arrojados, que estão dispostos a assumir maiores riscos em busca de retornos mais elevados.

Um comportamento intrigante é observado no *cluster* azul (0), que apresenta retorno tão baixo quanto o verde (2), mas demonstra uma volatilidade quase tão alta quanto o *cluster* laranja (1).

¹camilacaleones@ic.ufrj.br

²juvianna@ic.ufrj.br

³giseli@ic.ufrj.br

⁴laura@uniriotec.br

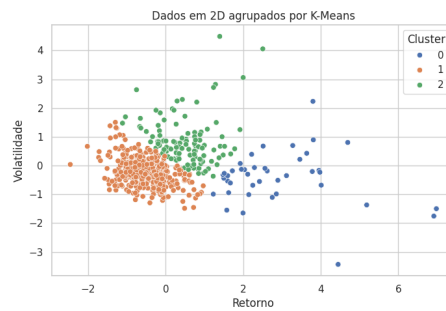


Figura 1: Distribuição dos clusters em relação ao retorno e volatilidade. Fonte: dos autores

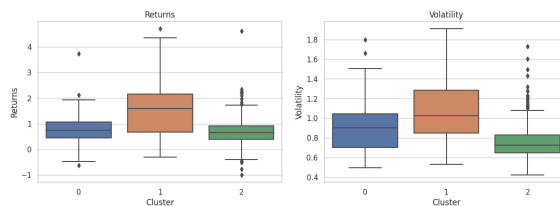


Figura 2: Classificação das empresas. Fonte: dos autores

Essa observação sugere que o *cluster* azul (0) pode representar um caso interessante e único, caracterizado por um equilíbrio indesejável entre um retorno baixo e uma volatilidade média.

Em relação aos retornos, o *cluster* laranja (1) se destaca por apresentar um retorno maior em comparação aos clusters azul (0) e verde (2). Enquanto os clusters azul (0) e verde (2) apresentam retornos medianos quase equivalentes, é importante destacar que o *cluster* verde (2) se destaca por ter um retorno menor em comparação aos demais.

Esses tipos de análises apresentadas facilitam o entendimento de investidores inexperientes sobre as características dos grupos de empresas. Dessa forma, podem ser utilizadas como subsídios para a recomendação de investimentos de acordo com o perfil de cada investidor.

Referências

- [1] H. Abdi e L. J. Williams. “Principal component analysis”. Em: **WIREs Computational Statistics** 13.2 (2021), e1519. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.1519>.
- [2] S. Charts. **S&P 500 - Standard and Poor’s 500**. <https://www.slickcharts.com/sp500>.
- [3] C. E. Mendes, J. P. Silva e M. A. Barbosa. “Análise de Cluster em Dados de Volatilidade de Mercados: Uma Aplicação em Índices Setoriais”. Em: **Anais do Congresso Nacional de Matemática Aplicada e Computacional (CNMAC)**. 2020, pp. 999–1010. URL: <http://www.sbmac.org.br/anais/2020/999.pdf>.
- [4] M. S. Torres, A. F. Brito e P. A. Nogueira. “Análise de Componentes Principais em Dados Financeiros: Uma Aplicação para Ações de Empresas Brasileiras”. Em: **Anais do Congresso Nacional de Matemática Aplicada e Computacional (CNMAC)**. 2019, pp. 1234–1245. URL: <http://www.sbmac.org.br/anais/2019/1234.pdf>.