UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO ACADÊMICO DO AGRESTE

ELIEL JOCSÃ MARINHO DE OLIVEIRA

OTIMIZAÇÃO DE CARTEIRAS COM INTELIGÊNCIA ARTIFICIAL: UMA APLICAÇÃO PRÁTICA DA TEORIA DE MARKOWITZ EM PYTHON

CARUARU

2024

UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO ACADÊMICO DO AGRESTE

CIÊNCIAS ECONÔMICAS

ELIEL JOCSÃ MARINHO DE OLIVEIRA

OTIMIZAÇÃO DE CARTEIRAS COM INTELIGÊNCIA ARTIFICIAL: Uma aplicação prática da teoria de Markowitz em Python

Trabalho de Conclusão de Curso apresentado ao Curso de Ciências Econômicas da Universidade Federal de Pernambuco – Centro Acadêmico do Agreste, como requisito parcial para a obtenção do título Bacharel em Ciências Econômicas.

Área de concentração: Teoria Econômica

Orientador: Prof. Msc. Anderson Issao Saito

CARUARU

2024

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Oliveira, Eliel Jocsã Marinho de.
Otimização de carteiras com inteligência artificial: Uma aplicação prática da teoria de Markowitz em Python / Eliel Jocsã Marinho de Oliveira. - Caruaru, 2024.
74 : il., tab.

Orientador(a): Anderson Issao Saito
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Ciências Econômicas, 2024.

1. Finanças. 2. Teoria das carteiras de Markowitz. 3. Inteligência Artificial. I. Saito, Anderson Issao. (Orientação). II. Título.

330 CDD (22.ed.)

ELIEL JOCSÃ MARINHO DE OLIVEIRA

OTIMIZAÇÃO DE CARTEIRAS COM INTELIGÊNCIA ARTIFICIAL:
Uma aplicação prática da teoria de Markowitz em Python

TCC apresentado ao Curso de Ciências Econômicas da Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, como requisito para a obtenção do título de Bacharel em Ciências Econômicas.

Aprovado em: 12/03/2024.

BANCA EXAMINADORA

____
Profº. Msc. Anderson Issao Saito (Orientador)
Universidade Federal de Pernambuco

____
Profº. Dr. Klebson Humberto de Lucena Moura (Examinador Interno)
Universidade Federal de Pernambuco

____
Profº. Msc. José Cícero de Castro (Examinador Interno)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Em primeiro lugar agradeço a Deus, que me ajudou a superar os desafios durante toda a graduação. Agradeço, também, aos meus pais que incentivaram e fortaleceram com seu amor, apoio e cuidado diários. À minha esposa pela sua dedicação, amor e paciência. E dedico este trabalho, bem como todo o meu esforço aos meus filhos, os maiores presentes que recebi do Senhor.

Gostaria de agradecer também ao meu orientador, professor Msc. Anderson Saito, por toda dedicação e solicitude na execução deste trabalho, bem como aos professores Dr. Klebson Moura e Msc. Cícero Castro, que fazem parte da banca examinadora, pelas contribuições ao trabalho.

Sou profundamente grato aos colegas e amigos que aproveitaram comigo o conhecimento adquirido neste curso. Nossas conversas e embates nos corredores da universidade acerca dos conteúdos das aulas foram de fundamental importância para a assimilação do conhecimento.

Aos professores que se dedicaram para transmitir esse conhecimento, agradeço de coração. Sempre que me utilizar da Ciência Econômica para desenvolvimento profissional e pessoal, lembrei de vocês.

RESUMO

A gestão de investimentos tem se tornado cada vez mais complexa, demandando estratégias eficazes de alocação de ativos e gestão de risco. Nesse cenário, a Teoria das Carteiras de Markowitz continua sendo relevante e amplamente aplicada, proporcionando a base teórica para a otimização de carteiras de investimentos. Este trabalho explora a aplicação prática da Moderna Teoria do Portfólio de Markowitz na maximização do Sharpe Ratio e na busca por uma fronteira eficiente entre risco e retorno. Por meio da utilização de ações das 10 empresas mais valiosas da Bolsa de Valores de São Paulo (B3), o estudo demonstra a importância da diversificação e otimização de carteiras para investidores. Além disso, o desenvolvimento de uma aplicação em Python permite que investidores comuns utilizem a Teoria das Carteiras na composição de suas carteiras de ações, contribuindo para a democratização do acesso a essa ferramenta fundamentada e visa facilitar o acesso e a utilização por interessados de diferentes áreas e níveis de conhecimento no mercado financeiro.

Palavras-chave: Teoria das Carteiras de Markowitz; Python; Inteligência Artificial.

ABSTRACT

Investment management has become increasingly complex, demanding effective asset allocation and risk management strategies. In this scenario, Markowitz Portfolio Theory continues to be relevant and widely applied, providing the theoretical basis for optimizing investment portfolios. This work explores the practical application of Markowitz's Modern Portfolio Theory in maximizing the Sharpe Ratio and searching for an efficient frontier between risk and return. Through the use of shares from the 10 most valuable companies on the São Paulo Stock Exchange (B3), the study demonstrates the importance of diversifying and optimizing portfolios for investors. Furthermore, the development of an application in Python allows ordinary investors to use Portfolio Theory in the composition of their stock portfolios, contributing to the democratization of access to this grounded tool and to facilitate access and use by interested parties from different areas and levels of knowledge in the financial market.

Keywords: Markowitz Portfolio Theory; Python; Artificial Intelligence.

LISTA DE FIGURAS

LISTA DE GRÁFICOS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

O mercado de ações, componente vital dos mercados financeiros, é o mercado onde investidores negociam ações de empresas publicamente listadas em uma bolsa de valores. Este mercado desempenha um papel de destaque na economia global e é relevante tanto para as empresas que buscam financiamento quanto para os investidores em busca de oportunidades de crescimento e rentabilidade. Ele é definido por Alexandre Assaf Neto, renomado autor e especialista em finanças, como um segmento do mercado de capitais onde ocorre a negociação de títulos emitidos por empresas e outros entes (Assaf Neto, 2019, p. 85). Esses títulos, conhecidos como ações, representam uma parte da propriedade da empresa emissora e conferem direitos aos acionistas.

Desse modo, o mercado de ações é um instrumento fundamental para empresas financiarem projetos de expansão e inovação, permitindo-lhes captar capital de uma ampla base de investidores (Bodie et al., 2014, p. 3). Como também destaca Assaf Neto (2019), o mercado de ações é um dos pilares do sistema financeiro, desempenhando um papel fundamental na mobilização de recursos para empresas e na alocação de capital. Essa alocação eficiente de capital é essencial para o crescimento econômico, uma vez que permite que empresas financiem projetos e inovações, o que impulsiona o desenvolvimento econômico. A B3 (Brasil, Bolsa Balcão), bolsa brasileira sediada na capital de São Paulo, segundo a World Exchange (2023), obteve uma capitalização de mercado[1], em julho de 2023, na casa dos 830.449,19 milhões de dólares. Foram 353 empresas listadas, dentre as quais 345 empresas brasileiras.

Os principais agentes do mercado de ações, conforme observado por Assaf Neto (2019), incluem investidores individuais, instituições financeiras, empresas listadas, reguladores e intermediários financeiros, como corretoras e bancos de investimento. Cada um desses agentes desempenha um papel crítico na dinâmica do mercado de ações. Como já observado no parágrafo anterior, as empresas são a parte financiada nas operações do mercado de ações, enquanto os investidores são a parte financiadora no processo. Por outro lado, para o investidor, denominado de

---

[1] Valor total das ações emitidas de empresas nacionais.

acionista, como enfatiza o autor brasileiro Gustavo Cerbasi, investir em ações é uma das formas mais eficazes de participar dos resultados das empresas e do crescimento econômico (Cerbasi, 2018, p. 105).

No entanto, podemos observar alguns problemas na capacidade dos agentes demandantes, isto é, os acionistas. Em um mercado dinâmico e com uma multiplicidade muito grande de opções de investimento, muitas vezes o investidor leigo não consegue reunir condições de realizar bons investimentos, sobretudo no Brasil onde é notória a ausência de educação financeira para a maioria da população, como aborda DE MEDEIROS & DE MEDEIROS (2011).

A primeira das barreiras encontradas para o investidor é o risco. Hull (2017) destaca que o risco é algo inerente ao mercado de ações. O mercado de ações está sujeito a riscos significativos, especialmente devido à volatilidade dos preços das ações (Assaf Neto, 2019, p. 88). Isso significa que os preços das ações podem flutuar substancialmente em curtos períodos, apresentando desafios consideráveis para investidores que buscam minimizar suas perdas.

Outra questão importante é que investir requer uma análise criteriosa das empresas e das condições macroeconômicas (Assaf Neto, 2019, p. 88). Isso implica que os investidores precisam ter habilidades analíticas sólidas e acesso a informações de qualidade para tomar decisões informadas. Desse modo, a participação de leigos, sem pré-capacitação na área, pode levar a tomadas de decisão precipitadas. Assaf Neto (2019) adverte ainda que a falta de preparação e conhecimento pode levar a escolhas financeiras inadequadas. Malkiel (2015) adverte que a acessibilidade crescente ao mercado de ações atraiu investidores inexperientes que frequentemente tomam decisões precipitadas, sujeitas a erros financeiros significativos.

Como Cerbasi (2018) destaca, investir em ações é uma maneira eficaz de participar dos resultados das empresas e do crescimento econômico. No entanto, os desafios inerentes a este mercado, incluindo o risco, a complexidade e a participação de investidores leigos, exigem abordagens cuidadosas e estratégias bem definidas para navegar com sucesso nesse ambiente.

O avanço da tecnologia tem se tornado um fator facilitador para o acesso dos gestores a ferramentas para uma melhor tomada de decisão. (SHARDA E*t al.,* 2019).

À medida que o mercado de ações evolui, testemunhamos uma revolução na forma como os investidores abordam a tomada de decisão. A utilização da programação para resolução de problemas de processamento e análise de dados financeiros aparece como um facilitador, dentre as linguagens mais utilizadas podemos citar R e Python, sendo esta última conhecida por uma menor curva de aprendizado. Assim, a análise de dados com Python tem desempenhado um papel crescente no processo de avaliação de ativos e na previsão de tendências de mercado e será o ferramental utilizado neste trabalho. Python tem se destacado como a linguagem preferida[2] por profissionais de *data science* devido à sua facilidade de uso e à vasta gama de bibliotecas disponíveis para análise de dados. Isso oferece aos investidores uma vantagem significativa ao processar grandes volumes de dados em tempo real e identificar padrões complexos (McKinney, 2017).

Além disso, a inteligência artificial (IA) desempenha um papel cada vez mais central na tomada de decisões de investimento. Algoritmos de IA, especialmente os baseados em aprendizado de máquina, têm a capacidade de analisar dados em tempo real e prever tendências de mercado com alta precisão. A IA capacita investidores a identificar insights ocultos em grandes conjuntos de dados e adaptar suas estratégias de investimento de acordo com as mudanças nas condições do mercado (Russell & Norvig, 2020).

Nesse contexto, procuraremos responder se o uso do ferramental tecnológico pode propiciar melhores resultados a investidores de maneira automatizada. O presente trabalho tem como objetivo geral desenvolver uma aplicação, utilizando sobretudo a linguagem de programação Python e algoritmos de inteligência artificial, que retorne ao usuário, dentre um conjunto de ações, qual a melhor combinação de pesos na composição de um portfólio[3]; para isso, nos debruçaremos sobre a teoria moderna da composição de carteiras de Markowitz, utilizando-a como referência para o funcionamento da aplicação e obtenção dos pesos ótimos da carteira. Em seguida, simularemos a composição de um portfólio de ações contendo as 10 ações mais

[2] O site <https://www.tiobe.com/tiobe-index> apresenta o índice das linguagens de programação mais utilizadas por desenvolvedores e engenheiros de software, bem como o número de cursos disponíveis. Atualmente (Fevereiro de 2024) a linguagem ocupa o 1º lugar, posição que já ocupava no ano anterior.

[3] Conjunto de ativos, neste caso ações de empresas. Também chamado de carteira de ações.

valiosas da bolsa de valores de São Paulo, obtendo carteiras otimizadas, utilizando a aplicação que será desenvolvida. Por fim, compararemos os seus desempenhos com o de outra carteira composta apenas com o índice da bolsa de valores do Brasil, a fim de verificar se há melhora no retorno em relação às demais.

Portanto, este trabalho servirá como contribuição para a democratização do acesso a uma ferramenta teoricamente bem fundamentada para montagem e análise de carteiras. Assim, pessoas de diferentes áreas, com diferentes níveis de conhecimento, poderão simular, de modo automatizado, a melhor composição de carteira para uma determinada coleção de ações.

2 REFERENCIAL TEÓRICO

O mercado de capitais realiza a importante tarefa de interligar poupadores e investidores, o que o caracteriza como uma grande fonte de financiamento do setor privado. Exerce, portanto, papel fundamental no processo de desenvolvimento econômico de um país. Neste capítulo exploraremos o funcionamento do mercado de capitais brasileiro, o seu lugar no sistema financeiro nacional, seu órgão regulador, a principal bolsa de valores brasileira e suas principais características. Trataremos também da natureza dos dados dos pregões da bolsa divulgados diariamente pela principal bolsa de valores do mercado mobiliário brasileiro, a B3[4].

Neste capítulo apresentaremos, também, a Moderna Teoria do Portfólio de Markowitz[5], sua relevância, principais conceitos e modelização matemática. Essa apresentação partirá de uma visão mais genérica de como se dá o processo de investimento, à luz da teoria de Markowitz, e seguirá abordando os conceitos-chaves da teoria, apresentando suas formulações matemáticas e exemplificando graficamente questões pertinentes como a da Fronteira Eficiente de Ativos com Risco e Retorno, importante ferramenta para determinação de um conjunto dentre o qual escolhemos a carteira ótima. Abordaremos, também, o índice de Sharpe[6] como a métrica utilizada para comparar os resultados dos portfólios de maneira objetiva, dentre os pontos ao longo da Fronteira Eficientes dos Ativos com Risco.

Por fim, apresentaremos a aplicação da Inteligência Artificial no problema, utilizando algoritmos de busca local, capazes de realizar a escolha da carteira ótima, a partir de diferentes caminhos. Abordaremos conceitos gerais e, em seguida, demonstraremos os algoritmos. Discorreremos desde o mais simples deles, o Hill Climb (subida de encosta) até o mais complexo dentre os abordados neste trabalho,

---

[4] B3 (Brasil, Bolsa, Balcão) é a sigla para a Bolsa de Valores oficial do Brasil.

[5] Harry Markowitz apresentou em sua tese de doutorado em economia em 1952, para a Universidade de Chicago o que seria a primeira teoria a propor a gestão de carteiras de ativos com base em sua relação risco-retorno. Recebeu em 1990, o prêmio Nobel de Economia conjuntamente alguns de seus discípulos por sua contribuição à teoria das carteiras (PINHEIRO, 2019).

[6] William F. Sharpe foi um notável economista, discípulo de Markowitz, e criador do Modelo de Precificação de Ativos (CAPM). Neste modelo o rendimento dos ativos estão relacionados com o rendimento do mercado em que estão inseridos. Foi ganhador, juntamente com Markowitz, do Nobel de Economia em 1990 (PINHEIRO, 2019).

o algoritmo genético, que se inspira no processo evolutivo biológico para chegar a um resultado final assertivo.

2.1 MERCADO DE CAPITAIS NO BRASIL

O mercado de capitais é caracterizado por Assaf Neto (2015) como uma espécie de ponte entre os poupadores/investidores e empresas que necessitam de financiamento a longo prazo. Este financiamento, segundo o autor, pode ser realizado através de diferentes instrumentos ou títulos, dentre os quais: *comercial papers*, destinados ao financiamento de capital de giro; debêntures, que são títulos privados de crédito emitidos por empresas; e ações, que são definidas como a menor fração do capital social de uma sociedade anônima.

Ações conferem direitos aos seus detentores a depender do seu tipo. Em geral, existem três tipos de ação: ordinária, preferencial e de gozo ou fruição. Assaf Neto (2019) destaca que as ações preferenciais conferem preferência em relação ao recebimento dos dividendos[7], bem como no reembolso em caso de liquidação (encerramento) da sociedade. Por conseguinte, as ações ordinárias conferem um direito que não se aplica às ações preferenciais: o direito a voto na assembleia de acionistas (ASSAF NETO, 2015). Assim sendo, eles deliberam sobre os rumos da empresa e têm seus direitos de recebimento de dividendos e reembolso de capital em caso de liquidação condicionada às preferências dos acionistas preferenciais. O autor também define a ação de gozo ou fruição como sendo de posse e propriedade dos fundadores da empresa, a qual lhes permite o direito de receber dividendos, preferência na aquisição de novas ações e conserva o direito ao voto.

No Brasil, o mercado de valores mobiliários, como ações e debêntures, têm seu funcionamento regulado pela CVM (Comissão de Valores Mobiliários). Na estrutura do SFN (Sistema Financeiro Nacional), a CVM faz parte do subsistema normativo (ASSAF NETO & GUASTI LIMA, 2008, p. 20). As suas atribuições são amplas, uma vez que é responsável por regulamentar, controlar e regular esse mercado. Ela fiscaliza as companhias emitentes dos títulos e até mesmo os investidores do mercado (ASSAF NETO & GUASTI LIMA, 2008, p. 21). Por fim, entre outras atribuições da Comissão de Valores Mobiliários, uma atribuição importante é a de supervisionar as bolsas de valores atuantes no país. (ASSAF NETO & GUASTI LIMA, 2008, p. 27).

---

[7] "parte dos resultados da empresa, determinada em cada exercício social e distribuída aos acionistas sob a forma de dinheiro" (Assaf Neto, 2019, p. 108)

As bolsas de valores, por sua vez, podem ser definidas como instituições auxiliares responsáveis pelo controle, fiscalização e divulgação dos negócios diários de compra e venda de títulos de empresas. (ASSAF NETO & GUASTI LIMA, 2008, p. 27). A B3 é a principal bolsa de valores nacional; sediada em São Paulo, possui mais 353 empresas listadas. (WORLD EXCHANGE, 2023).

2.1.1 Dados dos Pregões

Qualquer análise quantitativa pressupõe o conhecimento da estrutura dos dados com os quais estamos lidando. Portanto, um ponto que nos interessa para a análise dos dados das ações, para além da estrutura do sistema financeiro, do mercado e de seus órgãos e instituições, é o funcionamento dos pregões da bolsa e como os dados desses pregões são divulgados.

Nesse sentido, a estrutura dos dados divulgados pela própria B3 em seu Boletins Diário de Informações nos ajuda a compreender de que modo se organizam as informações dos preços dos ativos. A figura 1, mostra a estrutura do boletim diário de informações fornecido pela bolsa de valores. Os principais dados de preço para cada ação são Abertura, Mínimo, Máximo, Médio e Fechamento, além da Oscilação dos preços. Além disso, o documento também fornece informações acerca do volume de transações.

Figura 1 - Exemplo de Boletim Diário de Informações

BDI Boletim Diário de Informações

[B]³

REFERENTE A TERÇA-FEIRA - 19 DE SETEMBRO DE 2023 - N° 180

A Vista, termo e opções

Lote Padrão

Código	Empresa/Ação		Abertura	Mínimo	Máximo	Médio	Fechamento	Oscilação (%)	Ofertas		Negócios Realizados	
									Compra (R$)	Venda (R$)	Número	Quantidade
5GTK11	INVESTO 5GTK	CI	73,19	72,68	73,19	73,10	73,15	-0,05↓	70,56	73,16	9	350
A1AP34	ADVANCE AUTO	DRN	17,58	17,58	18,23	18,19	18,22	2,35↑	18,18	18,22	6	239
A1CR34	AMCOR PLC	DRN	44,69	44,65	44,69	44,67	44,65	-1,43↓	44,00	45,30	2	3
A1DM34	ARCHER DANIE	DRN	384,18	384,18	385,32	384,38	385,15	-0,04↓	385,15	-	3	38
A1EG34	AEGON NV	DRN	24,82	24,63	24,82	24,75	24,63	0,85↑	22,00	26,00	2	3

Fonte: B3 (2023).

Em geral, como verifica-se nos trabalhos de Pinheiro (2019) e Assaf Neto (2015), o preço utilizado para análise de dados de ações é o preço de fechamento de cada pregão. Outras análises, no entanto, podem ser feitas a partir dos demais dados. Para Assaf Neto (2015, p. 440), a variabilidade da ação em um dia é benéfica para os praticantes de d*ay trade*, por exemplo. Para eles uma maior variabilidade não configura necessariamente algo ruim, já que pode lhes propiciar mais momentos de oportunidade para obter ganho.

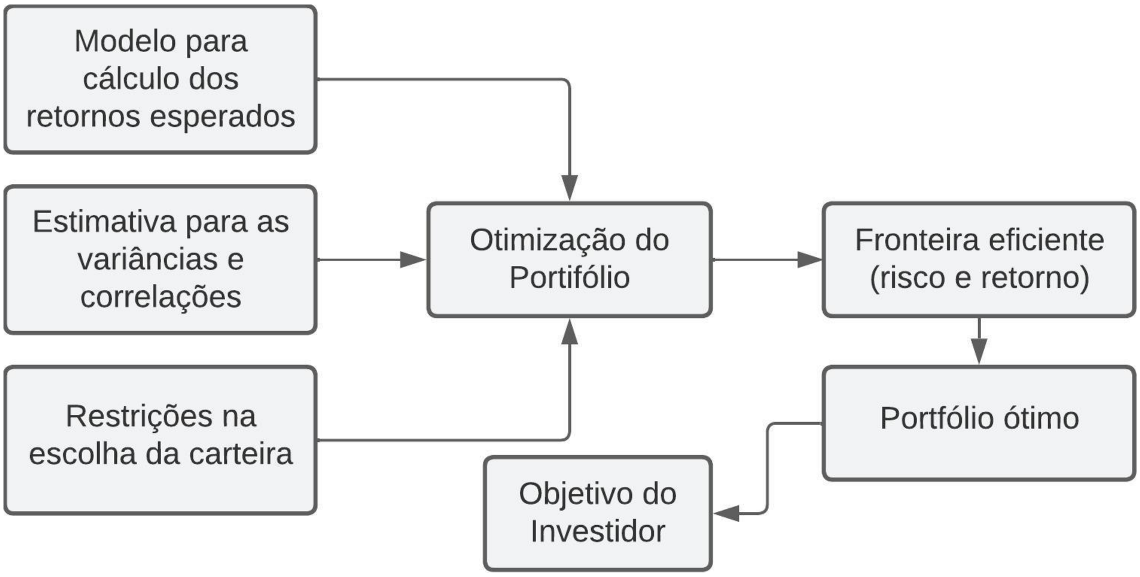
A B3 também disponibiliza uma planilha que lista o número de pregões realizados desde o ano 1964 até 2023. Essa planilha encontra-se no anexo A deste trabalho. Em média são realizados cerca de 250 pregões por ano, considerando que os pregões apenas são realizados em dias úteis. Essa informação é relevante, uma vez que a realização de cálculos de média anualizada tem de levar em consideração a quantidade de vezes que um determinado evento ocorreu em um ano. Em geral, o Banco Central do Brasil utiliza algo chamado base 252 para anualizar a taxa Selic e, portanto, o padrão de mercado para anualização é utilizar essa constante. Por isso, no nosso trabalho para a anualização de qualquer dado da nossa base tomaremos como parâmetro a base 252 (dias/ano).

2.2 TEORIA MODERNA DOS PORTFÓLIOS

A otimização de carteiras de investimentos realizada neste trabalho terá como base a Moderna Teoria do Portfólio (MTP) de Harry Markowitz, também conhecida como "*Portfolio Selection".* Como destacado por Araújo (2010), Markowitz publicou em 1952 um artigo seminal que é frequentemente considerado o marco inicial da moderna teoria financeira, pois formulou matematicamente o conceito de diversificação de investimentos.

Markowitz argumentou que a diversificação não se limita a reduzir o risco de ativos individuais, mas sim considerar como esses ativos interagem dentro de uma carteira de investimentos. Conforme apontado por Fabozzi (2002), no processo de investimento os investidores devem buscar maximizar o retorno esperado da carteira, ao mesmo tempo em que minimizam a variância dos retornos do portfólio. A otimização do portfólio resulta em uma fronteira eficiente entre risco e retorno e, dentre os pontos ótimos desta curva, o investidor escolhe od portfólio ótimo que preferir a depender de seu objetivo (maior retorno, menor risco ou índices de sua preferência), como expresso na Figura 2.

Figura 2 - Processo de investimento, segundo a teoria de Markowitz



Fonte: Fabozzi (2002). Alterada pelo autor.

No artigo "*Portfolio Selection*" (1952), publicado no *The Journal of Finance*, Markowitz introduziu a ideia de que os investidores devem diversificar seus recursos entre os ativos que oferecem o maior retorno esperado. Segundo Damodaran (2007), isso se baseia na regra da variância dos retornos esperados, que parte do princípio de que existe uma carteira ideal que proporciona o maior retorno com a menor variância. É importante notar que a busca pelo maior retorno nem sempre resultará na menor variância e vice-versa. A decisão final depende das preferências individuais do investidor ou de uma métrica objetiva, tal como o índice de Sharpe.

Essas premissas são importantes, pois configuram importantes alicerces para a modelagem feita por Markowitz (1952), bem como para os seus sucessores. As sete principais premissas da teoria moderna dos portfólios, segundo Zanini & Figueiredo (2005) são:

- “Os investidores avaliariam as carteiras apenas com base no retorno esperado e no desvio padrão dos retornos sobre o horizonte de tempo de um período”

(ZANINI & FIGUEIREDO, 2005). A escolha da carteira é baseada, portanto, no *trade-off* entre retorno esperado e risco da carteira. Assim, as curvas de utilidade do investidor são funções destas duas variáveis;

- “Os investidores seriam avessos ao risco. Se instados a escolher entre duas carteiras de mesmo retorno, sempre escolheriam o de menor risco” (ZANINI & FIGUEIREDO, 2005). Dado um nível de risco e retorno, a carteira escolhida será aquela que maximiza o retorno esperado com o menor risco. Dessa forma, o agente maximizará a utilidade esperada do investimento em certo intervalo de tempo;
- “Os investidores estariam sempre insatisfeitos em termos de retorno. Instados a escolher entre duas carteiras de mesmo risco, sempre escolheriam a de maior retorno” (ZANINI & FIGUEIREDO, 2005);
- “Seria possível dividir continuamente os ativos, ou seja, ao investidor seria permitido comprar mesmo frações de ações” (ZANINI & FIGUEIREDO, 2005);
- “Existiria uma taxa livre de risco, à qual o investidor tanto poderia emprestar quanto tomar emprestado” (ZANINI & FIGUEIREDO, 2005). Essa premissa, assim como no trabalho citado, não será relevante para este trabalho;
- “Todos os impostos e custos de transação seriam considerados irrelevantes” (ZANINI & FIGUEIREDO, 2005);
- “Todos os investidores estariam de acordo em relação à distribuição de probabilidades das taxas de retorno dos ativos. Isto significa que somente existiria um único conjunto de carteiras eficientes” (ZANINI & FIGUEIREDO, 2005).

Zanini & Figueiredo (2005) indicam ainda que, a partir dessas premissas, Markowitz esboça o que seriam as duas características fundamentais de uma carteira: O risco da carteira, que é mensurado pelo desvio padrão dos retornos dos ativos da carteira, e o retorno esperado da carteira, que é calculado pela média ponderada dos retornos individuais dos ativos da carteira.

A apresentação matemática da Teoria do Portfólio de Markowitz, portanto, é fundamentada em conceitos-chave da estatística e da economia financeira, como risco, retorno, covariância, correlação, fronteira eficiente e mercados eficientes. Segundo Assaf Neto (2013), esses conceitos formam a base para a construção e gestão de carteiras de investimentos e ajudam os investidores a tomarem decisões informadas e eficazes no ambiente financeiro contemporâneo. Versaremos nas subseções deste capítulo acerca desses diversos conceitos-chave e de sua modelagem matemática.

É importante destacar que a teoria do portfólio de Markowitz continua sendo relevante e amplamente aplicada na gestão de investimentos, servindo como uma pedra angular para estratégias de alocação de ativos e gestão de risco em um mundo financeiro cada vez mais complexo, como observado por Guerra (2009).

Ward, *et al* (2012) ratifica a relevância da MTP ao constatar que, conjuntamente às teorias dos sucessores de Markowitz, tais como Fama e Sharpe, a teoria dos portfólios continua sendo a base da maioria dos cursos de finanças e investimentos.

2.3 RENTABILIDADE DE UM PORTFÓLIO

O conceito de retorno esperado é de suma importância na avaliação de investimentos e na gestão de carteiras financeiras. Segundo a definição de Assaf Neto, o retorno pode ser compreendido como o ganho ou perda decorrente de um investimento durante um período específico (ASSAF NETO, 2018). O retorno esperado de uma carteira, por sua vez, é a média ponderada dos potenciais resultados desse investimento, calculados para diversos cenários previstos (ARAÚJO, 2010). Sendo essa média ponderada pela participação de cada ação no portfólio, também denominada de peso ação no portfólio. Essa abordagem nos conduz à compreensão de que o retorno está intrinsecamente ligado ao tempo e às variáveis que podem afetar o desempenho de um ativo.

Conforme conceitua Heij *et al.* (2004), a variável do retorno é tratada pela TMP como uma variável aleatória, incapaz de ser precisamente definida até o momento em que é efetivamente observada. Esta característica expressa a ideia de incerteza acerca dos retornos dos ativos. Métodos estatísticos, portanto, são fundamentais para a MTP, pois introduzem as noções de distribuição de probabilidade e valor esperado.

O cálculo do retorno diário dos ativos individuais é simplesmente dado pelo diferencial entre o preço final e inicial do ativo dividido pelo preço inicial do ativo, como vemos na equação (1), abaixo:

$$r_i = \frac{(P_1 - P_0)}{P_0} \tag{1}$$

Onde:

- r_i é o Retorno do ativo i;
- P_0 é o Preço Inicial do ativo;
- e P_1 é o Preço Final do ativo.

Considerando a natureza aleatória da variável, no caso de retornos equiprováveis, o retorno esperado para cada ativo (μ_i) é representado pela média aritmética dos retornos passados em uma população com T observações, sendo r_{ti} o retorno do ativo i no tempo t :

$$\mu_i = \frac{1}{T}\sum_{i=1}^{T} r_{ti} \tag{2}$$

Na gestão de uma carteira de investimentos, outrossim, a esperança do retorno total é calculada como a média ponderada dos retornos esperados de cada ativo individual (Markowitz, 1952). Portanto, o retorno esperado de uma carteira ($E\ (R_p)$) pode ser descrito nas formas algébrica e matricial, como a seguir:

$$E(R_p) = \mu_p = \sum_{i=1}^{N} X_i . \mu_i = \tag{3}$$

$$= [X_1\ X_2\ X_3 \dots X_N] \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ .. \\ \mu_n \end{bmatrix} = x'\mu$$

Onde:

- $E(R_p)$ representa o retorno esperado da carteira;
- N é o número de ativos na carteira;
- X_i denota o peso do ativo i na carteira;
- μ_i é o retorno esperado do ativo i.

Sendo, portanto, o retorno esperado de uma carteira de N ativos, calculado através da multiplicação da matriz transposta do vetor dos pesos de cada ativo pela matriz coluna dos retornos individuais esperados para cada um deles. Os pesos ótimos das ações de uma carteira são, portanto, o resultado mais importante da aplicação da moderna teoria dos portfólios, uma vez que a sua otimização pressupõe toda a combinação dos conceitos matemáticos e estatísticos por trás das variáveis do modelo. Desse modo, os pesos ótimos já estão considerando o risco e o retorno de cada ativo, além de uma condição estabelecida pelo investidor, como veremos adiante.

2.3.1 Modelo de Precificação de Ativos (CAPM)

De acordo com Pinheiro (2019), Harry Markowitz influenciou diversos outros modelos, dentre os quais um que viria a contribuir para a premiação de Willian Sharpe com o prêmio Nobel de Economia, o Modelo de Precificação de Ativos (CAPM). Embora não seja o enfoque principal deste trabalho, é importante mencionar que esse modelo é frequentemente considerado uma alternativa robusta para estimar retornos, especialmente quando se almeja reduzir a incerteza.

O CAPM é uma ferramenta fundamental que nos permite calcular o custo de capital de um investimento, ajudando assim na avaliação da viabilidade e rentabilidade dos ativos em questão. Sharpe (1964), proporciona uma estrutura teórica que relaciona o retorno esperado de um ativo com seu Beta e a uma taxa livre de risco.

A fórmula do CAPM para o cálculo do retorno esperado (E) de um ativo específico (i) é expressa da seguinte maneira:

$$E = R_f + \beta_i(E(R_m) - R_f) \tag{4}$$

Onde:

- R_f é o retorno de um ativo livre de risco na economia;
- $E(R_m)$ representa o retorno esperado de uma carteira diversificada de mercado, composta por todos os ativos de risco na economia;
- β_i corresponde ao coeficiente beta do ativo i, que quantifica o risco sistêmico[8] do ativo em relação às flutuações no retorno da carteira de mercado.

O Beta, como uma medida de risco sistemático, indica como um ativo se comporta em relação ao mercado como um todo. Um Beta superior a 1 implica que o ativo é mais volátil do que o mercado, enquanto um Beta inferior a 1 sugere uma volatilidade menor. A interpretação do resultado do Beta é crucial, pois ele influencia diretamente o cálculo do custo de capital e, consequentemente, a avaliação de um investimento (ASSAF NETO, 2018).

Quando um investidor compreende o Beta de um ativo, pode tomar decisões informadas sobre como esse ativo se encaixa em seu portfólio, levando em consideração a relação risco-retorno. Seu cálculo é realizado através de análises estatísticas que relacionam o desempenho histórico do ativo com o desempenho do mercado como um todo (ASSAF NETO, 2018). O Beta é, portanto, uma métrica que quantifica a sensibilidade do ativo às flutuações do mercado e é calculado da seguinte forma:

$$\beta = \frac{cov(\mu_r, \mu_m)}{\sigma_m^2} \tag{5}$$

Onde:

- $cov(\mu_r, \mu_m)$ é a covariância entre os retornos dos do ativo (μ_r) e os retornos do mercado (μ_m);
- e σ_m^2 é a variância dos retornos do mercado.

[8] O risco sistêmico será melhor explicado na seção 2.4.

O CAPM, portanto, oferece uma estrutura robusta para calcular o custo de capital de um ativo com base em seu risco sistemático, representado pelo Beta, e será uma das formas de avaliação de portfólios utilizadas na metodologia deste trabalho.

Essa métrica desempenha um papel essencial na avaliação de ativos e na gestão de portfólios, permitindo que investidores e analistas ponderem riscos e retornos. A partir da interpretação dos resultados do CAPM, os tomadores de decisão podem determinar se um ativo está em linha com as expectativas de mercado, equilibrando os potenciais ganhos com os riscos associados (ASSAF NETO, 2018). O resultado do CAPM pode ser lido, de maneira mais prática, como o retorno percentual mínimo necessário para que o investidor se exponha ao risco da carteira em questão.

Em resumo, o retorno esperado é uma métrica essencial para investidores, cujo cálculo pode ser abordado por diferentes métodos, sendo o CAPM frequentemente adotado quando se busca uma estimativa mais precisa e menos sujeita a incertezas (Fama, 1970).

2.4 RISCO DE UM PORTFÓLIO

Inicialmente, é importante compreender como o risco é medido para um ativo individual. Assaf Neto (2018) define o risco como a capacidade de medir as incertezas que cercam uma escolha com base no conhecimento das probabilidades de resultados. A medida de risco é frequentemente expressa como o desvio-padrão da taxa de retorno (σ), como descrito por Bodie, Kane e Marcus (2015). Essa medida representa a dispersão dos retornos históricos de um ativo ao longo do tempo.

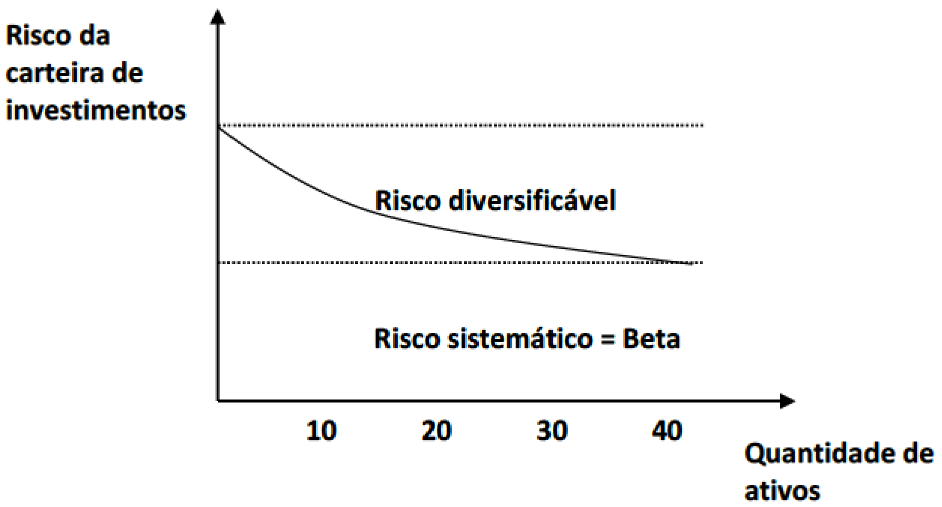
$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (R_i - \underline{R})^2} \tag{6}$$

Aqui, σ é o desvio-padrão, N é o número de observações de retornos, R_i é o retorno do ativo no período i, e $\underline{R}$ é a média dos retornos ao longo do período.

Ao construir uma carteira de ativos, no entanto, o risco não pode ser simplesmente calculado como uma média ponderada dos riscos individuais. Assaf Neto (2018) distingue o risco total de um ativo como a soma de sua parte sistêmica e não sistêmica. A parte sistêmica ou sistemática, de natureza conjuntural, não pode ser evitada, enquanto a parte não sistêmica ou não sistemática é inerente a cada ativo.

Risco sistemático é definido como sendo o risco “inerente a todos os ativos negociados no mercado, sendo determinado por eventos de natureza política, econômica e social.” (ASSAF NETO, 2018). Risco diversificável (não sistemático), por sua vez, é definido como sendo “identificado nas características do próprio ativo, não se alastrando aos demais ativos da carteira" (ASSAF NETO, 2018). A Figura 3 demonstra essa relação:

Figura 3 - ***Tipos de risco da carteira de investimentos***



Fonte: Artigo no site Yubb[9].

Uma consideração importante é que o risco de um ativo individual quando inserido em uma carteira não é igual ao risco quando interage com outros ativos da carteira. Cada ativo contribui de maneira diferente para o risco total da carteira. Além disso, quanto mais ativos compuserem a carteira, o risco tenderá a diminuir até certo ponto, devido à permanência do risco sistêmico (Assaf Neto, 2018).

[9] Disponível em: https://yubb.com.br/artigos/acoes/indice-beta-de-acoes-o-que-e-e-como-calcular

Assaf Neto, 2018, também afirma em relação ao risco não sistemático de um ativo que a “sua eliminação de uma carteira é possível pela inclusão de ativos que não tenham correlação positiva entre si”. Assaf Neto, 2018, no entanto, salienta que a taxa de diminuição do risco de uma carteira é decrescente à medida que se aumenta o número de ativos; isto se dá, pois, no limite, permanecerá um nível de risco sistemático impossível de ser eliminado da carteira.

Para entendermos melhor o conceito da diversificação diminuindo o risco de uma combinação de ativos, comecemos analisando a matematização do problema com dois ativos. O cálculo que relaciona a covariância dos retornos e a correlação dos mesmos é desenvolvido por Assaf Neto, 2018. O risco para uma carteira com dois ativos, X e Y, pode ser dado pelas seguintes expressões, que utilizam a covariância ou a correlação entre os dois ativos no cálculo do risco total:

- Utilizando a covariância dos retornos dos ativos temos:

$$\sigma_p = \sqrt{(w_X^2 \, . \, \sigma_X^2) + (w_Y^2 \, . \, \sigma_Y^2) + \; 2. (w_X \, . \, w_Y \, . \, cov_{XY})} \qquad 7)$$

- Já utilizando a correlação entre eles:

$$\sigma_p = \sqrt{(w_X^2 \, . \, \sigma_X^2) + (w_Y^2 \, . \, \sigma_Y^2) + \; 2. (w_X \, . \, w_Y \, . \, \rho_{XY} \, . \, \sigma_X \, . \, \sigma_Y)} \qquad 8)$$

Essas equações, portanto, se equivalem. Apenas chegam ao mesmo resultado por caminhos distintos, isto se dá pois existe a seguinte relação entre covariância e correlação:

$$\rho_{XY} = \frac{cov_{XY}}{\sigma_X \, . \, \sigma_Y} \qquad (9)$$

Nas equações (7), (8) e (9) as variáveis são estão descritas da seguinte maneira:

- σ é o desvio padrão, para cada ativo;
- w é o peso, para cada ativo;

- cov_{XY} é a covariância dos retornos dos ativos X e Y;
- ρ_{XY} é a correlação entre os retornos dos ativos X e Y;

Para calcular o risco de uma carteira com múltiplos ativos, a Teoria do Portfólio utiliza uma matriz de covariância, que representa as relações entre os ativos. A variância da carteira (σ_p^2) pode ser calculada com base nessa matriz, como indicado por Bodie, Kane e Marcus (2015):

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=i}^{N} w_i w_j Cov(R_i R_j) \tag{10}$$

Onde σ_p^2 é a variância da carteira, w_i e w_j são os pesos dos ativos i e j, respectivamente, e $Cov(R_i R_j)$ é a covariância entre os retornos de i e j.

A mesma transformação utilizada na análise com dois ativos, objetivando a apresentação do cálculo do risco através da correlação entre os retornos dos ativos, também pode ser feita para múltiplos ativos. Assim:

$$\sigma_p^2 = \sum_{i=1}^{N} \sum_{j=i}^{N} w_i w_j \, \rho_{ij} \sigma_i \, \sigma_j \tag{11}$$

Nas equações (10) e (11) as variáveis são estão descritas da seguinte maneira:

- σ é o Desvio Padrão do Portfólio (a volatilidade do portfólio);
- w_ie w_j são os pesos dos ativos individuais no portfólio;
- cov_{ij} é a covariância entre os ativos i e j;
- ρ_{ij} é a correlação entre os ativos i e j;

A construção de uma carteira, segundo a teoria de Markowitz, demanda o cálculo da matriz de covariância ou da matriz de correlação dos ativos. No entanto, quanto mais ativos na carteira, mais complexo se torna esse cálculo. Diversos métodos, como a matriz de covariância amostral, *Risk Metrics* e o método do encolhimento, são utilizados para calcular as matrizes de covariância, levando em consideração o número limitado de observações e minimizando os erros de estimação (Santos e Tessari, 2012). Algoritmos computacionais, como o que vamos desenvolver neste trabalho, são úteis no sentido de facilitar a realização de cálculos dessa natureza, possibilitando ainda a automação deles.

Por fim, a correlação entre os ativos desempenha um papel crucial na diversificação da carteira. A diversificação ideal, conforme enfatizada por Markowitz (1959), ocorre quando os ativos têm alta correlação, mas não perfeita. Isso permite reduzir o risco da carteira, mas sem eliminá-la completamente. Em outras palavras, a diversificação eficaz não requer que os ativos sejam completamente independentes, mas sim que suas variações não se movam perfeitamente em conjunto. Por outro lado, Moura (2009) argumenta que quanto mais correlacionados os pares de ativos, mais exposta às variações do mercado estará a carteira. Nesse caso, o risco da carteira será maior. Portanto, é importante encontrar um equilíbrio entre a diversificação e a correlação dos ativos para otimizar a gestão de riscos.

A correlação é uma medida estatística que indica o grau de linearidade entre dois ativos. Ela varia de -1 (correlação negativa perfeita) a +1 (correlação positiva perfeita), com 0 indicando ausência de relação. Além disso, o cálculo matricial do Risco (volatilidade) de um portfólio pode ser realizado também utilizando a matriz de correlações dos ativos, como demonstra a Figura 4.

Figura 4 - ***Matriz de correlação dos ativos***

$$\sigma_p^2 = \begin{bmatrix} W_1 & \cdots & W_i & \cdots & W_n \end{bmatrix} \overset{\text{MATRIZ DE CORRELAÇÃO}}{\begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{i,1} & \cdots & \sigma_{n,1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma_{i,1} & \ddots & \sigma_i^2 & \ddots & \sigma_{n,j} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \sigma_{n,1} & \cdots & \sigma_{n,j} & \cdots & \sigma_n^2 \end{bmatrix}} \begin{bmatrix} W_1 \\ \vdots \\ W_i \\ \vdots \\ W_n \end{bmatrix}$$

Fonte: Artigo no site Clube de Finanças[10]. Alterada pelo autor.

[10] Disponível em: http://clubedefinancas.com.br/materias/fronteira-eficiente/

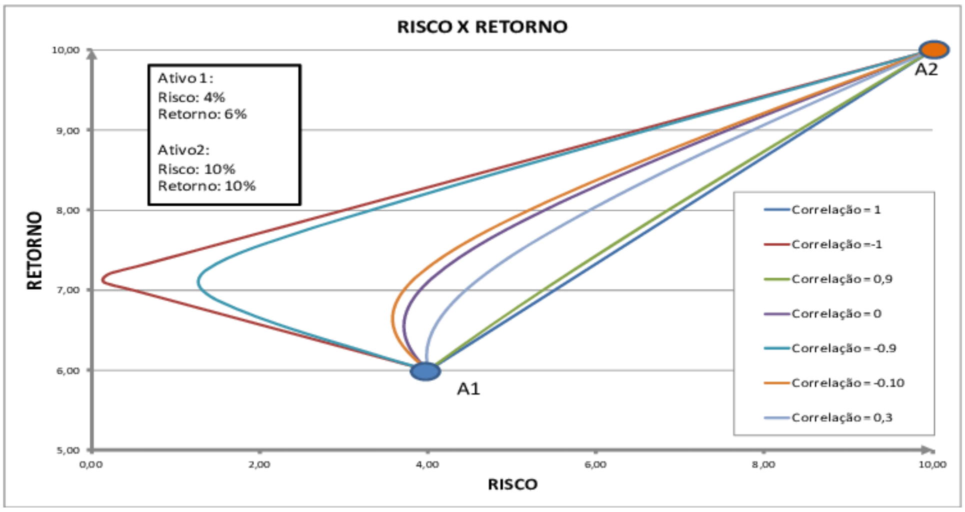
2.5. FRONTEIRA EFICIENTE DE ATIVOS E ÍNDICE DE SHARPE

A teoria de Markowitz não se limita apenas a calcular o risco e o retorno de uma carteira de ativos. Ela também oferece uma maneira de determinar a composição ideal da carteira, conhecida como a "fronteira eficiente de ativos" (Markowitz, 1952). Esta fronteira é um conceito fundamental na tomada de decisões de investimento e ajuda os investidores a selecionarem a melhor combinação de ativos para atingir um determinado nível de risco ou retorno.

A fronteira eficiente de ativos é uma curva que representa todas as possíveis combinações de ativos em uma carteira que oferecem o maior retorno esperado para um determinado nível de risco ou o menor risco para um determinado nível de retorno (Bodie, Kane & Marcus, 2015). Em outras palavras, ela mostra como um investidor pode otimizar sua carteira, equilibrando o retorno desejado com o risco aceitável.

Para construir uma fronteira eficiente, os investidores precisam considerar as taxas de retorno esperadas de cada ativo, a matriz de covariância ou a matriz de correlação que descrevem as relações entre os ativos e, é claro, seus objetivos pessoais de risco e retorno. A fronteira eficiente destaca as oportunidades de diversificação, mostrando como diferentes ativos interagem entre si em termos de risco e retorno. A Figura 5 exemplifica uma fronteira eficiente e demonstra a sua relação com o nível de correlação entre dois ativos, demonstrada por Miguel & Ramos (2017). Quando a correlação entre os ativos for igual a 1 tem-se a reta (correlação perfeita), quando a correlação for igual a -1 temos o outro caso extremo, ou seja, risco zero. Portanto, entre os valores -1 e 1 há os casos intermediários. O risco diminui à medida que a correlação entre os ativos é menor.

Figura 5 - Fronteira eficiente



Fonte: Miguel & Ramos, 2017.

A análise das carteiras na fronteira eficiente de ativos, conforme destacado por Markowitz (1959), é essencial para compreender como os investidores podem otimizar suas escolhas de portfólio. A fronteira eficiente representa as diferentes combinações possíveis de ativos que oferecem o melhor equilíbrio entre risco e retorno. Dentro dessa fronteira, existem portfólios que se destacam como ótimas escolhas.

Um portfólio é considerado eficiente se atender a três critérios essenciais estabelecidos por Markowitz (1959):

- **I - Carteira Legítima**: Para ser considerado eficiente, um portfólio deve ser uma "carteira legítima", o que significa que ele deve ter um retorno médio ponderado pelos ativos que o compõem. Além disso, o desvio padrão (risco) desse portfólio deve ser superior ao de uma carteira menos arriscada.
- **II - Melhor Retorno com Maior Risco**: Se outra carteira legítima tiver um retorno esperado superior ao do portfólio original, essa carteira alternativa também deve apresentar um risco (variação de retorno) maior. Isso garante que o portfólio seja uma escolha competitiva em termos de retorno versus risco.
- **III - Menor Risco com Menor Retorno**: Da mesma forma, se uma carteira legítima tiver um risco menor em comparação com o portfólio original, seu

retorno esperado também deve ser menor. Isso significa que o portfólio não está assumindo riscos desnecessários em troca de um retorno insuficiente.

Portfólios ineficientes geralmente não conseguem satisfazer esses critérios (II) e (III). Portanto, ficam abaixo da fronteira eficiente. Eles podem ser superados por outras carteiras que oferecem um equilíbrio mais favorável entre risco e retorno, estas são as carteiras eficientes e estão representadas exatamente sobre a linha da fronteira eficiente mencionada. A busca por carteiras eficientes na fronteira eficiente de ativos é um elemento crucial na construção de uma estratégia de investimento sólida.

Além disso, ao selecionar ativos específicos para compor uma carteira, os investidores também precisam considerar a relação entre retorno e risco. É aí que entra o Índice de Sharpe (Sharpe, 1966), uma medida que ajuda os investidores a avaliarem se um ativo ou carteira oferece um retorno adequado dado o nível de risco assumido. O Índice de Sharpe compara o retorno de um ativo ou carteira com o retorno de um ativo livre de risco (geralmente representado por títulos do governo). No Brasil, é prática comum utilizar a taxa de juros SELIC como uma proxy para o ativo livre de risco; sobre isso, discutiremos na próxima subseção que tratará das características do mercado de capitais brasileiro. A seguinte equação quantifica o índice de Sharpe:

$$S = \frac{R_p - R_f}{\sigma_p} \tag{6}$$

Onde:

- R_p é o retorno da carteira ou ativo;
- R_f é o retorno do ativo livre de risco;
- e σ_p é o desvio-padrão (risco) da carteira ou ativo.

Um Índice de Sharpe mais alto indica que um ativo ou carteira está oferecendo um retorno melhor em relação ao risco, enquanto um Índice de Sharpe mais baixo sugere que o retorno não está compensando adequadamente o risco (Sharpe, 1966). Assaf Neto (2018) estabelece que o índice Sharpe exprime a relação entre o prêmio pago pelo risco, isto é, o retorno e o próprio risco que o investidor assume ao realizar uma operação no mercado de ações.

Na prática, os investidores podem usar a fronteira eficiente para identificar as carteiras que oferecem o máximo de retorno para um determinado nível de risco (Markowitz, 1952). Em seguida, eles podem usar o Índice de Sharpe para escolher a carteira ideal entre as opções disponíveis na fronteira eficiente.

Como vimos, a fronteira eficiente de ativos e o Índice de Sharpe, permitem que os investidores construam carteiras que atendam às suas metas de retorno e tolerância ao risco, ajudando-os a tomar decisões informadas em um ambiente complexo de investimento.

3 METODOLOGIA

O desenvolvimento de uma pesquisa requer razões de ordem intelectual ou prática. Segundo Gil (2002, p. 17), as razões intelectuais podem ser descritas como decorrentes do próprio desejo pelo conhecimento, enquanto as razões práticas decorrem do desejo por um conhecimento que agregue melhorias à forma de se fazer algo em relação a como se faz hoje.

Neste sentido, e ciente de que as razões supramencionadas não se anulam, este trabalho realiza, com natureza quantitativa, uma pesquisa que se utilizará de técnicas exploratórias e avaliativas acerca da utilização das ferramentas de Data Science e Inteligência Artificial no mercado de capitais brasileiro.

É importante salientar, ao se tratar de pesquisa quantitativa, o que diz Apolinário (2004). Para ele, a pesquisa quantitativa lida com fatos. Desta forma, as variáveis devem estar bem definidas (seu escopo, tempo, unidade de mensuração etc) e as análises bem fundamentadas em critérios matemáticos.

Uma das formas de emprego da pesquisa exploratória, segundo Marconi e Lakatos (2002), é o estudo exploratório-descritivo combinado. As autoras definem este como sendo “estudos exploratórios que têm por objetivo descrever completamente determinado fenômeno” (MARCONI & LAKATOS, 1990, p.85). Neste trabalho usaremos a pesquisa exploratória no sentido de obter conhecimento acerca do processo de desenvolvimento da aplicação de otimização de carteiras com Python[11] e IA, isto é, como se dará a aplicação prática das ferramentas tecnológicas para otimização das carteiras de ações.

Já a técnica avaliativa, conforme descrita por Walliman (2014, p.6), proporciona a análise de “achados” e a conjectura de possíveis resultados a partir destes. Portanto, essa técnica nos possibilitará, primariamente, avaliar se o uso dessas tecnologias resultará em bons resultados e possibilitará uma utilização rápida e eficiente da teoria moderna dos portfólios.

[11] Python é uma linguagem de programação de softwares desenvolvida no fim da década de 1980 e amplamente conhecida pela sua sintaxe simples e eficaz. Desde seu lançamento tem ganhado crescente notoriedade e aceitação, sobretudo dos cientistas de dados e estatísticos, pela ampla gama de ferramentas que oferece através de suas bibliotecas.

O procedimento realizado será um estudo de caso, uma vez que obteremos dados históricos da bolsa de valores de São Paulo – B3. As ações utilizadas nesse estudo serão as que ocuparam as 10 primeiras posições no índice de ações mais valorizadas na B3, na data do estudo. O intervalo de tempo para os dados das cotações inicia-se no dia 01 de janeiro de 2017 até o dia 31 de dezembro de 2023. Esse estudo verificará a possibilidade da aplicação de *Data Scien*ce e IA com Python de forma simples e automatizada por investidores sem grande conhecimento técnico e os seus resultados práticos.

A realização do estudo de caso requer a instalação na máquina do usuário de um único software, para execução local dos códigos. (Ressalta-se que a obtenção dos dados através do YFinance depende de acesso à internet). Essa etapa deve ser pulada por aqueles que utilizarão algum serviço de execução do python na nuvem, tal como o Google Colab Notebooks. O software requerido é o seguinte:

Anaconda 3 – Este software inclui o interpretador Python e diversas bibliotecas pré-instaladas (numpy, pandas, mathplotlib etc). Além disso, sua instalação acompanha a instalação de uma IDE (Ambiente de Desenvolvimento Integrado) conhecida por sua versatilidade e facilidade de compreensão, o Jupyter Notebooks. Link para download do Anaconda 3 em sua última versão: <https://www.anaconda.com/download>[12].

Àqueles com mais disponibilidade de tempo ou curiosidade, um outro modo de executar o Jupyter Notebooks, com uma aparência distinta, é de maneira conjunta ao VS Code, uma IDE bastante popular para desenvolvedores. Nesse caso, a instalação do Anaconda 3 continua sendo um requisito. Deve-se instalar, também, uma extensão para o VS Code nomeada como Jupyter. Algumas configurações adicionais podem ser necessárias. No entanto, é importante frisar que para usuários leigos a simples instalação do Anaconda 3 e a utilização do Jupyter Notebooks se faz suficiente.

A subseção a seguir se dedica a abordar a importância da aplicação da linguagem Python, da ciência de dados e da inteligência artificial no mercado financeiro, sobretudo de capitais. Versaremos sobre como a aplicação de ferramentas

[12] Este link para download encontrou-se válido quando da redação deste trabalho, não sendo garantida a sua disponibilidade posteriormente.

como ciências de dados com Python e algoritmos de inteligência artificial[13] podem contribuir para aplicação dessa teoria. As contribuições da linguagem incluem a facilidade na coleta e tratamento de grandes quantidades de dados e, além disso, a automação de cálculos extensos de maneira replicável e simples. Apresentaremos, por fim, a função de uma biblioteca[14] para a linguagem, as principais bibliotecas utilizadas para análise de dados financeiros e as suas respectivas funções. Essas bibliotecas vão desde ferramentas de tratamento de dados e estatística como o Pandas e Scipy até bibliotecas que implementam algoritmos de inteligência artificial, como a *mlrose*. Desse modo, alguns algoritmos de busca local utilizados por essa biblioteca serão apresentados. Estes são fundamentais para lidar com buscas de um resultado ótimo dentre um conjunto de estados. Neste sentido, a sua aplicação na teoria dos portfólios constitui um avanço computacional importante na obtenção de pesos ótimos de uma carteira.

3.1 APLICAÇÃO DE DATA SCIENCE E IA COM PYTHON

O uso dos recursos tecnológicos em finanças é abordado por diversos autores que enfatizam os diversos benefícios que a utilização dessa tecnologia pode proporcionar a investidores, gestores, economistas e outros profissionais. Hilpisch (2014) escreveu um livro que se intitula, em tradução nossa, como Python para finanças, onde afirma que a relevância que a tecnologia tem para os agentes financeiros se deve aos ganhos de vantagem competitiva que esta proporciona. Cada vez mais, as instituições financeiras passam a se tornar tecnológicas.

Hilpisch (2014) aponta ainda que a importância do setor financeiro, com sua larga capacidade de investimento, na área da tecnologia constitui uma recíproca

---

[13] Inteligência Artificial (IA) refere-se à capacidade de máquinas e sistemas de computador realizarem tarefas que normalmente requerem inteligência humana, como aprendizado, raciocínio e resolução de problemas, usando algoritmos e dados. É um campo interdisciplinar que busca criar sistemas autônomos e adaptáveis para diversas aplicações. As origens desse campo de estudo e tecnologia remontam às décadas de 1950 e 1960. Para aprofundamento Acerca dos conceitos e contribuições da IA em diversas áreas do conhecimento recomenda-se o acesso ao seguinte link : <https://www.oracle.com/br/artificial-intelligence/what-is-ai>.

[14] Uma biblioteca Python funciona como uma caixa de ferramentas. Os desenvolvedores oferecem, geralmente através da chamada de funções, diversas ferramentas "prontas para uso". É comum a utilização da frase "você não precisa ficar reinventando a roda", como forma de explicar o conceito de biblioteca dentro do Python.

verdadeira ao parágrafo anterior. Sendo assim, o autor afirma que as instituições financeiras são responsáveis por empregar milhões de desenvolvedores no mundo. Pode-se observar, portanto, uma forte interligação entre esses dois setores de serviços.

Isso se deve, conjuntamente, por alguns fatores importantes. O primeiro deles é que a tecnologia funciona como facilitador para a execução de operações financeiras cada vez mais complexas. Grande parte dessa complexidade é constituída pelo segundo ponto: os vultosos volumes de dados que são gerados e analisados diariamente no mercado, fator que demanda cada vez mais dos sistemas computacionais. Em terceiro lugar, a análise em tempo real desses constitui uma importante vantagem competitiva para as empresas. Por fim, a tecnologia também vem sendo usada como barreira à entrada para esse mercado, uma vez que demanda qualificação e expertise, além de investimentos significativos (HILPISCH, 2014).

Nesse contexto, o uso do Python, para Hilpisch (2014), como uma linguagem e ecossistema, apresenta-se como uma escolha altamente adequada para o setor financeiro. Sua atratividade reside em uma série de vantagens notáveis, que vão desde uma sintaxe elegante até abordagens eficazes de desenvolvimento. Um dos pontos fortes do Python é sua vasta gama de bibliotecas e ferramentas prontamente disponíveis. Esse arsenal de recursos parece capaz de abordar a maioria das demandas levantadas pelos desenvolvimentos recentes na indústria financeira. Estes desenvolvimentos incluem a análise de dados em larga escala, com volumes substanciais e alta frequência de dados.

3.1.1 Características da Linguagem

A linguagem de programação Python foi criada por Guido van Rossum em 1991. Sua denominação curiosa é uma homenagem ao programa de televisão "*Monty Python's Flying Circus*", popular entre os programadores da época. Desde então, obteve um crescimento vertiginoso, consolidando-se como uma escolha primordial em diversas áreas, incluindo desenvolvimento web, ciência de dados, inteligência artificial e aprendizado de máquina. Sua comunidade de desenvolvedores ativos e comprometidos, contribuiu para a criação de uma ampla biblioteca de módulos e pacotes. Esses recursos têm uma aplicação versátil, desde a simplificação da visualização de dados até a execução de análises avançadas. Por isso, grandes empresas de tecnologia, como Google[15] e Microsoft[16], incorporam Python em suas operações, destacando sua relevância[17] no mercado atual (COSTA, 2023).

Como descrito na introdução deste capítulo, as qualidades que tornam a linguagem Python um facilitador são a sua elegância e características sintáticas. O Quadro 1 traz as principais características da linguagem e suas definições, de acordo com Hilpisch (2014). Dessas, destacamos, juntamente com o código aberto, isto é, não necessidade de licenciamento para uso e alterações na linguagem, seu uso multiplataforma como um fator crucial para sua popularidade e, principalmente, para a democratização do seu uso. Ao estar disponível no Linux, por exemplo, é facilmente embarcada em servidores e serviços em nuvem, tais como o Google Colab Notebooks, que permite a utilização da linguagem a partir de dispositivos com pouco poder de processamento.

Tabela 1 - *Principais características da linguagem*

Código Aberto (*Open Source*)	Python e a maioria das bibliotecas e ferramentas de suporte disponíveis são de código aberto e geralmente vêm com licenças bastante flexíveis e abertas.
Interpretada	Através do CPython, o interpretador da linguagem, o código Python é traduzido em tempo de execução em código de bytes executável (linguagem de máquina).
Multiparadígma	Python suporta diferentes paradigmas de programação e implementação, como orientação a objetos e programação imperativa, funcional ou procedural.
Multiuso	Python pode ser usado para desenvolvimento de código rápido e interativo, bem como para construção grandes aplicações; ele pode ser usado para operações de sistemas de alto ou baixo nível.

[15] A Google é bastante empenhada na propagação da linguagem Python, tanto que possui uma plataforma de serviço em nuvem para a execução de código Python de maneira remota e que dispensa a necessidade de grande poder computacional local para o usuário. Essa plataforma é o Google Colab Notebooks.

[16] A Microsoft também está empenhada em propagar a utilização da linguagem, tamanha sua versatilidade e adequação para tratamento e análise de dados. Recentemente, a última versão do Office 365 traz, no Excel, uma integração com código Python de maneira nativa.

[17] A linguagem Python está classificada como a primeira linguagem mais utilizada no mundo atualmente, segundo o Índice Tiobe. O índice pode ser consultado no seguinte link: <https://www.tiobe.com/tiobe-index>

Multi-Plataforma (*Cross-platform*)	Funciona em diversos sistemas operacionais: Windows, MacOS, Linux e outros.
Dinamicamente "tipada"	Os tipos em Python são geralmente inferidos durante o tempo de execução e não declarados estaticamente como na maioria das linguagens compiladas (Python é interpretada).
Identation aware	Python usa recuos para marcar blocos de código em vez de parênteses, colchetes ou ponto e vírgula.
Garbage collecting	Coleta o "lixo" (cache) da memória RAM de maneira automática. O desenvolvedor não precisa se preocupar com isso.

Fonte: Hilpisch (2014), adaptada pelo autor.

Hilpisch (2014), assim como Costa (2023), também destaca o ecossistema de bibliotecas no qual o Python está inserido. Elencamos, portanto, as principais bibliotecas destacadas pelo autor, bem como outras bibliotecas importantes na análise de dados financeiros e aplicação de algoritmos de otimização.

3.1.2 Bibliotecas

As bibliotecas, no Python, desempenham uma função muito importante. Os desenvolvedores não dispõem, como no Excel por exemplo, de um conjunto de recursos ou ferramentas para plotagem de gráficos, cálculos estatísticos, análise de dados ou otimizações em geral. Nesse contexto, as bibliotecas aparecem como ferramentas direcionadas para um uso específico. Essa característica enriquece as possibilidades do desenvolvedor, uma vez que cada ferramenta pode ser desenvolvida de maneira bastante específica. (HILPISCH, 2014). Elencaremos a seguir, de maneira breve, as bibliotecas utilizadas neste trabalho para o desenvolvimento da aplicação.

- YFinance - A biblioteca YFinance é especializada na obtenção de dados financeiros e informações sobre o mercado de capitais. Ela permite o acesso a uma ampla gama de dados financeiros, como preços de ações, informações sobre empresas, índices de mercado e séries temporais financeiras. Isso é vital para a análise de portfólio, modelagem de séries temporais financeiras e tomada de decisões de investimento. A yfinance torna mais acessível a tarefa de coletar e analisar dados financeiros para profissionais da área. É importante salientar que a biblioteca utiliza a API do Yahoo Finance para coletar os dados das variadas bolsas pelo mundo (PYTHON PACKAGE INDEX, 2023).

- Pandas - A biblioteca Pandas é uma ferramenta fundamental para a manipulação e análise de dados. Ela oferece estruturas de dados flexíveis, como DataFrames e Series, que permitem a importação, limpeza e transformação de dados com facilidade. Os dados de séries temporais fornecidos pelas bolsas de valores podem ser filtrados, agregados, pivotados e ser utilizados para cálculos estatísticos. Pandas é uma escolha essencial para cientistas de dados e investidores que precisam explorar e preparar dados para análise (PYTHON PACKAGE INDEX, 2023).

- Scipy - A biblioteca Scipy é uma coleção de módulos e funções que expande as capacidades do Python para tarefas relacionadas à ciência e engenharia. Ela inclui módulos para otimização, álgebra linear, integração numérica, estatísticas e processamento de sinais, entre outros. A scipy é uma ferramenta essencial para resolver problemas complexos que envolvem matemática avançada e cálculos científicos. Ela é frequentemente usada em aplicações que vão desde modelagem estatística até simulações numéricas (PYTHON PACKAGE INDEX, 2023).

- Numpy - Essencial para realizar cálculos numéricos eficientes. Ele permite a criação e manipulação de arrays multidimensionais, tornando-o ideal para tarefas de análise de dados, processamento numérico e computação científica. O NumPy é amplamente usado em áreas como ciência de dados, aprendizado de máquina e engenharia devido à sua capacidade de executar operações rápidas em grandes conjuntos de dados.

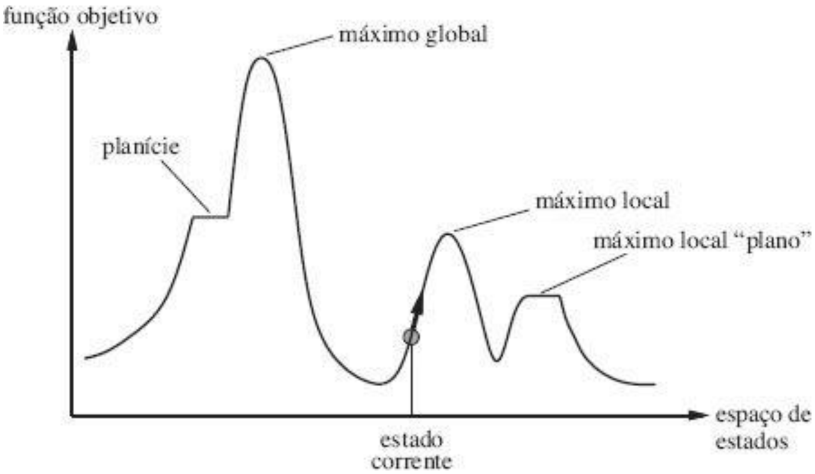
- Matplotlib e Plotly – Tratam-se de ferramentas de visualização de dados que oferece controle completo sobre a criação de gráficos e plots personalizados. Ela permite a criação de uma variedade de visualizações, incluindo gráficos de dispersão, gráficos de barras, gráficos de linha e histogramas. Com o Matplotlib, é possível personalizar cada aspecto de um gráfico, como cores, marcadores e títulos. Enquanto o principal diferencial do Plotly é fornecer gráficos dinâmicos que apresentam os dados de acordo com o “passar do mouse” do usuário (PYTHON PACKAGE INDEX, 2023).

- Mlrose - Responsável pela otimização através de algoritmos de inteligência artificial. Em sua documentação (HAYES, 2019), o desenvolvedor demonstra a aplicação de cinco algoritmos de IA para a otimização de problemas de maximização ou minimização. Três desses algoritmos são referenciados no livro “*Artificial Intelligence: A Modern Approach*” dos autores Russell & Norvig (2020). Os algoritmos apresentados no livro e implementados pela biblioteca são algoritmos de busca local . São eles:
    - Subida de encosta (*Hill Climb)*;
    - Têmpera simulada (*Simulated Anneling*);
    - e Algoritmo Genético (Genetic Alg).

3.1.3 Algoritmos de Otimização

Russell & Norvig (2020) apresentam os algoritmos de busca local como uma opção satisfatória para problemas de otimização, nos quais o objetivo é encontrar o melhor estado com base em uma função objetivo. A Figura 6 exemplifica com o que se depara o algoritmo de busca local. O conceito de estado atual é indispensável para compreender o funcionamento dos algoritmos de busca local. O "estado" se refere a uma configuração ou uma solução possível dentro do espaço de busca do problema, também chamada de espaço de estados. Essa configuração representa uma posição específica nesse espaço e pode ser representada por um conjunto de parâmetros, variáveis ou valores que caracterizam essa posição (RUSSELL & NORVIG, 2020).

Figura 6 - ***Espaço de estados (Algoritmos de busca local)***



Fonte: Russell & Norvig (2020).

O primeiro algoritmo referenciado por Russel & Norvig (2020) e implementado na biblioteca Mlrose, o da subida de encosta, funciona de maneira bastante simples: nas palavras de Russell & Norvig (2020): "Ele é simplesmente um laço repetitivo que se move de forma contínua no sentido do valor crescente, isto é, encosta acima. O algoritmo termina quando alcança um "pico" (máximo local) e é descrito pelos autores em linguagem de algoritmo, como na Figura 7, abaixo:

Figura 7 - ***Algoritmo Hill Climb***

```
função SUBIDA-DE-ENCOSTA(problema) retorna um estado que é um máximo local
  corrente ← CRIAR-NÓ(ESTADO-INICIAL[problema])
  repita
    vizinho ← um sucessor de corrente com valor mais alto
    se VALOR[vizinho] VALOR[corrente] então retornar ESTADO[corrente]
    corrente ← vizinho
```

Fonte: Russell & Norvig (2020).

O autor explica ainda que algoritmos como o da subida de encosta são frequentemente chamados de algoritmos ambiciosos. Isso se deve ao fato de que eles possuem uma certa ânsia[18] por achar um máximo, mas não sabem qual o próximo passo para determinar se este máximo é, de fato, o máximo global.

O segundo algoritmo busca solucionar uma limitação do primeiro, que por sua simplicidade pode ficar preso facilmente a um máximo local. Neste sentido a têmpera simulada funciona de maneira distinta: o algoritmo utiliza uma analogia bastante interessante, inspirada no processo físico de resfriamento e têmpera de materiais, onde um sistema é gradualmente resfriado até alcançar um estado de equilíbrio de menor energia. Ele começa com uma solução inicial e gera soluções vizinhas. Se uma solução vizinha for melhor, ela é aceita; mas até soluções piores podem ser aceitas com uma probabilidade decrescente, controlada pela "temperatura". Com o tempo, a temperatura diminui, tornando menos provável a aceitação de soluções piores. O algoritmo continua até atender a um critério de parada; sendo este o momento em que a temperatura (T) é mínima, proporcionando uma solução aproximada para o problema de otimização.

É usado em problemas complexos onde encontrar a solução ótima diretamente é difícil (RUSSELL & NORVIG, 2020). O algoritmo está demonstrado na figura 8, abaixo.

Figura 8 - *Algoritmo Simulated Annealing*

```
função TÊMPERA-SIMULADA(problema, escalonamento) retorna um estado solução
  entradas: problema, um problema
escalonamento, um mapeamento de tempo para "temperatura"
  atual ← CRIAR-NÓ(problema.ESTADO-INICIAL)
  para t = 1 até ∞ faça
    T ← escalonamento[t]
    se T = 0 então retornar corrente
    próximo ← um sucessor de atual selecionado aleatoriamente
    ΔE ← próximo.VALOR – atual.VALOR
    se ΔE > 0 então atual ← próximo
    senão atual ← próximo somente com probabilidade e^(ΔE/T)
```

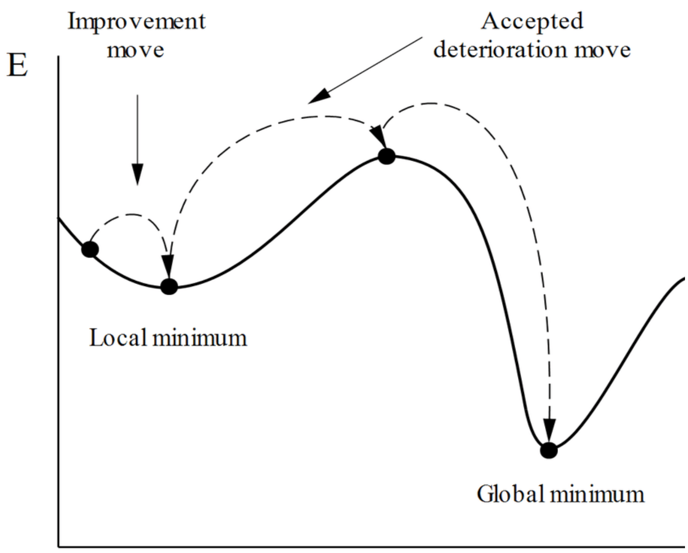
Fonte: Russell & Norvig (2020).

[18] Essa questão pode ser "contornada", se executarmos novamente o algoritmo e verificarmos se o resultado é maior que o anterior, e guardá-lo se este for o caso. Ao repetirmos o processo indefinidamente a tendência é encontrarmos máximos locais cada vez "melhores" até que se chegue a um máximo global. O problema desse procedimento, no entanto, é que nunca teremos certeza de que este máximo é, de fato, o máximo global.

Esse algoritmo, portanto, funciona de maneira mais dinâmica que o Hill Climb. Enquanto, o Hill Climb para sua execução no primeiro máximo local encontrado, e sua repetição ocorre de maneira "manual" pelo o usuário, o Simulated Anealling tem sua dinâmica própria de contornar o problema da ambição. A Figura 8 demonstra que quanto mais próximo de um resultado global menor a "temperatura" e, portanto, mais perto do fim da execução do algoritmo. Essa dinâmica é fundamentada na probabilidade de existir uma solução melhor, que vai diminuindo conforme o algoritmo vai verificando a escassez destas.

A figura 9 demonstra que a todo momento o algoritmo de Simulated Annealing avalia se é compensatório ou não realizar um esforço (gastar energia) para tentar encontrar um novo ponto ótimo; no caso da figura este seria um novo mínimo, até chegar ao mínimo global. Isso ocorre analogamente nos casos de máximo global, onde o algoritmo avalia se é compensatório, dada a probabilidade de encontrar máximo maior, fazer um esforço indo para um ponto vizinho mais baixo.

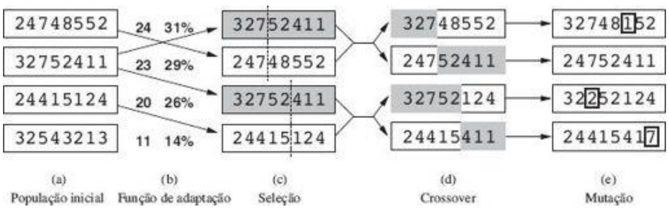
Figura 9 - *Dinâmica Simulated Annealing*



Fonte: Rosocha *et. al* (2015).

O último algoritmo de busca local implementado na biblioteca mlRose que iremos abordar, o mais sofisticado, envolve conceitos diferentes. O algoritmo genético (AG) é uma técnica de otimização que se inspira na seleção natural, evoluindo uma população de soluções candidatas para encontrar soluções próximas do ótimo. Como demonstram Russell & Norvig (2020), os AGs trabalham com uma população inicial de soluções representadas como cadeias de valores. Essas soluções são avaliadas por meio de uma função de adaptação que mede sua qualidade. A seleção de indivíduos para reprodução, o cruzamento que combina características de diferentes pais e a mutação ocasional introduzem diversidade na população. O processo evolutivo continua até que um critério de parada seja alcançado, e a melhor solução encontrada é considerada a solução final. A Figura 10 exemplifica as etapas desse processo.

Figura 10 - *Processo evolutivo do algoritmo genético*



Fonte: Russell & Norvig (2020).

O algoritmo genético segue uma abordagem inspirada na evolução biológica, onde a seleção natural é simulada para otimização. Como observado no texto, a função de adaptação atribui uma pontuação a cada indivíduo com base em sua qualidade. A seleção favorece os indivíduos mais aptos, mas também incorpora aleatoriedade para manter a diversidade. O cruzamento permite a combinação de características genéticas de pais diferentes, enquanto a mutação introduz variação. Esse processo iterativo gera soluções melhores ao longo das gerações e é repetido até que um critério de parada seja satisfeito, produzindo uma solução próxima do ótimo (RUSSELL & NORVIG, 2020). O algoritmo genético também está demonstrado, na íntegra, na Figura 11.

Figura 11 - *Algoritmo genético*

```
função ALGORITMO-GENÉTICO(população, FN-ADAPTA) retorna um indivíduo
  entradas: população, um conjunto de indivíduos
            FN-ADAPTA, uma função que mede a adaptação de um indivíduo
  repita
    nova_população ← conjunto vazio
    para i = 1 até TAMANHO(população) faça
      x ← SELEÇÃO-ALEATÓRIA(população, FN-ADAPTA)
      y ← SELEÇÃO-ALEATÓRIA(população, FN-ADAPTA)
      filho ← REPRODUZ(x, y)
      se (pequena probabilidade aleatória) então filho ← MUTAÇÃO(filho)
      adicionar filho a nova_população
      população ← nova_população
  até algum indivíduo estar adaptado o suficiente ou até ter decorrido tempo suficiente
  retornar o melhor indivíduo em população, de acordo com FN-ADAPTA
________________________________________________________________________________
função REPRODUZ(x, y) retorna um indivíduo
  entradas: x, y, indivíduos pais
  n ← COMPRIMENTO(x)c ← número aleatório de 1 a n
  retornar CONCATENA(SUBCADEIA(x, 1 c), SUBCADEIA(y, c + 1, n))
```

Fonte: Russell & Norvig (2020).

3.1.4 Método de Monte Carlo

É importante apresentarmos, também, um outro método de otimização que, ao contrário da IA, utiliza o simples poder computacional bruto para chegar aos resultados. O método de Monte Carlo consiste em qualquer algoritmo computacional estatístico baseado na amostragem aleatória e massiva para obter resultados numéricos. Desse modo, é classificado como um método de *brute force*, usa o poder computacional de processamento para realizar tarefas de cálculos repetidamente e, consequentemente, escolher o melhor dentre os resultados gerados aleatoriamente.

A utilização desse método é, sobretudo, custosa. Há um grande desperdício de poder computacional, uma vez que os resultados são gerados a partir de parâmetros aleatórios, o que pode resultar também em resultados piores tanto menor seja a amostra. O aumento da amostra, no entanto, demanda mais uso do poder computacional e, portanto, um aumento dos custos de processamento.

4 RESULTADOS

Tendo claro os conceitos da teoria dos portfólios de Markowitz e as funcionalidades da linguagem Python, bem como as aplicações de IA que nos serão úteis, desenvolvemos uma aplicação que permite a um investidor comum aplicar a Moderna Teoria das Carteiras na composição de sua carteira de ações.

É importante salientar que uma análise fundamentalista é indispensável em um cenário real de investimento. No entanto, partiremos do ponto onde o investidor já escolheu seus ativos e precisa alocá-los de maneira eficiente em sua carteira. Para isso, utilizaremos ações das 10 empresas mais valiosas da Bolsa de Valores de São Paulo (B3) em 2023. A Tabela 1 apresenta cada um desses ativos e o seu respectivo valor de mercado.

Partiremos de um cenário onde o investidor no início do ano de 2023, munido dos dados dos preços de fechamento das ações de 2018 a 2022[19] (1 de janeiro de 2018 a 31 de dezembro de 2022), estima os pesos ótimos que maximizam o Sharpe Ratio da carteira de investimentos para o período. Assim, ele realiza um investimento no início de 2023 nas ações mencionadas da Tabela 1, compondo sua carteira com os pesos ótimos calculados.

Ao fim do ano de 2023, faremos uma avaliação acerca do retorno e da evolução de patrimônio resultante da utilização dos pesos ótimos estimados na análise. Além disso, poderá comparar o desempenho da sua carteira com o desempenho da BOVA (Índice Bovespa), que sintetiza o desempenho da Bolsa brasileira de um modo geral. Além disso, o investidor poderá comparar também a sua composição de carteira, com uma composição igualitária de carteira (mesmo peso para todas as ações), refletindo se a aplicação da Moderna Teoria das Carteiras lhe rendeu bons resultados.

Durante essa análise, descreveremos os passos que a aplicação executa, descrevendo as suas principais funcionalidades e as informações geradas para o

[19] Utilizamos a maior série de tempo (com todos os anos completos) possível para as ações escolhidas. Uma vez que a ação com dados mais recentes aparece nos pregões a partir de 1 de março de 2017, qual seja a BPAC11.SA. Dessa forma, iniciaremos a base de dados da fase de otimização das carteira no primeiro ano em que há dados disponíveis para todas ações a partir do primeiro pregão do ano (2018).

investidor. Deste modo, além de analisar os resultados da aplicação da teoria, também verificaremos a construção e o uso desta aplicação.

Tabela 2 - *Índice das 10 maiores empresas da B3 em valor de mercado (2023)*

CÓDIGO[20] DA AÇÃO UTILIZADA	EMPRESA	VALOR DE MERCADO R$ (MIL)
PETR4.SA	PETROBRAS	475.335.629,6
VALE3.SA	VALE	333.390.105,9
ITUB4.SA	ITAÚ UNIBANCO	282.099.349,2
ABEV3.SA	AMBEV S/A	209.525.982,7
BBDC4.SA	BRADESCO	161.104.202,5
BBAS3.SA	BANCO DO BRASIL	151.093.439,5
WEGE3.SA	WEG	143.254.463,3
BPAC11.SA	BTGP BANCO	133.087.454,2
SANB11.SA	SANTANDER BRASIL	115.102.451,6
ITSA4.SA	ITAÚSA	98.669.320,09

Fonte: B3, adaptada pelo autor.

4.1 INSTALAÇÃO DAS BIBLIOTECAS NECESSÁRIAS

Previamente a todas as etapas de funcionamento da aplicação que desenvolvemos em Python, é preciso instalar todas as bibliotecas que utilizaremos do início ao fim do projeto. A figura 12 demonstra o código fonte necessário para fazer a importação de uma das bibliotecas, sendo extensível para as demais.

[20] O código das ações expressa, para além da empresa representada por aquele ativo, o seu tipo. Em termos gerais as ações identificadas com o numeral 3 são ordinárias, as identificadas com o numeral 4 são preferenciais e as identificadas com o numeral 11 representam BDRs, isto é *Brazilian Deposicional Recepts* (Ações de empresas estrangeiras representadas na Bolsa brasileira).

No nosso código utilizamos testes condicionais para verificar se cada biblioteca está disponível para importação. Caso contrário, a aplicação primeiramente realiza sua instalação para, em seguida, tentar realizar novamente a importação da biblioteca. Portanto, o processo é inteiramente automatizado.

Figura 12 - *Importação das bibliotecas*

```
# Importação das bibliotecas externas:

try:
    import yfinance as yf
except:
    %pip install yfinance
```

Fonte: Autoria Própria.

4.2 OBTENÇÃO E TRATAMENTO DOS DADOS HISTÓRICOS DE PREÇOS DOS ATIVOS

Após a instalação das bibliotecas que serão necessárias, a primeira etapa do funcionamento da aplicação é a obtenção dos dados históricos dos preços das ações escolhidas previamente. Esta coleta é realizada através da biblioteca Yahoo Finance[21]. Primeiramente, armazenamos em um vetor os códigos das ações, assim como descritos na Tabela 2. Em segundo lugar, devemos armazenar as datas de início e fim da série histórica que desejamos obter em suas respectivas variáveis. Por fim, a chamada do método *yf.download()* executa esta ação para cada ativo, recebendo como parâmetros a data inicial e final e código de um ativo. Como precisamos coletar os dados para todos os ativos, utilizamos a biblioteca Pandas para criar um *dataframe* que armazenará os dados de cada ativo em uma coluna única. As linhas correspondem, por sua vez, às datas dos pregões. Após o processo de download automatizado, os dados ficam organizados como na Figura 13.

[21] Esta biblioteca conecta a nossa aplicação com os dados da bolsa presentes no site do Yahoo Finance, importante braço da empresa de tecnologia mais conhecida pelo seu site de buscas. Portanto, facilita o processo de obtenção dos dados da bolsa. Uma vez que a utilização da API disponibilizada pela B3 é um tanto mais complexa e demanda maior conhecimento técnico.

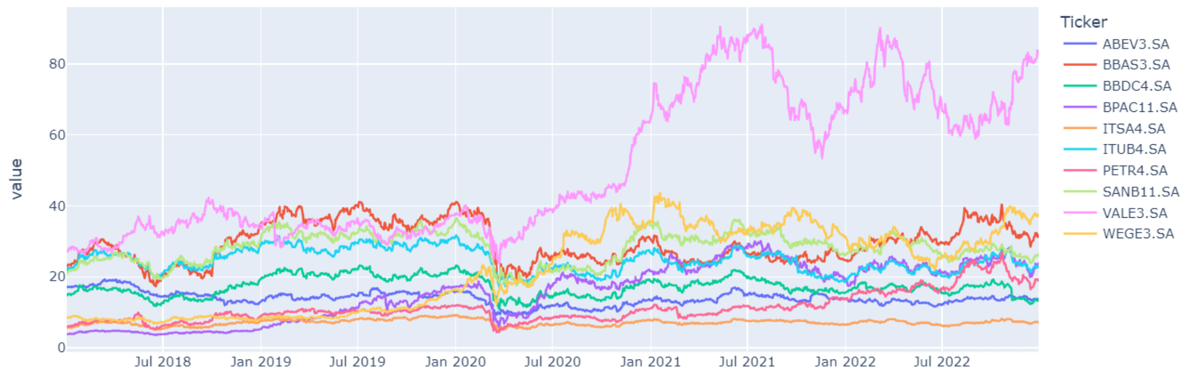
Figura 13 - *Dataframe com os preços de fechamento das ações*

Ticker Date	ABEV3.SA	BBAS3.SA	BBDC4.SA	BPAC11.SA	ITSA4.SA	ITUB4.SA	PETR4.SA	SANB11.SA	VALE3.SA	WEGE3.SA
2018-01-02	17.242147	22.969938	14.845123	3.879362	5.650487	21.889252	6.032850	21.234653	27.350491	8.752859
2018-01-03	17.281897	23.262907	14.913377	3.916309	5.717086	22.038906	6.087527	21.059479	27.186594	8.619922
2018-01-04	17.186504	23.486116	15.158057	3.930676	5.824664	22.497845	6.098464	21.312502	27.298042	8.535959
2018-01-05	17.250099	23.486116	15.243918	3.995665	5.834910	22.562685	6.134916	21.539583	27.724173	8.570942
2018-01-08	17.218302	23.541918	15.239624	4.044627	5.829786	22.487865	6.207820	21.956524	28.340406	8.640907
...	...	...	...	...	...	...	...	...	...	...
2022-12-23	13.893774	32.437153	13.721280	23.783394	7.419653	23.264914	19.409340	26.204805	81.185493	37.570316
2022-12-26	13.798742	32.391178	13.395013	23.191236	7.324529	22.803129	19.270262	25.907442	81.928505	37.267799
2022-12-27	13.798742	31.315460	13.331574	22.647615	7.246701	22.673832	19.401613	25.991074	83.884789	37.092148
2022-12-28	13.827251	31.876307	13.739404	23.773685	7.393710	23.135616	19.162085	26.390650	83.696678	37.765488
2022-12-29	13.798742	31.931471	13.730342	23.230066	7.359121	23.089439	18.930286	26.204805	83.593216	37.580074

Fonte: Autoria própria.

Utilizando a biblioteca matplotlib podemos, também, plotar o gráfico dos preços históricos das ações (Gráfico 1).

Gráfico 1 - *Histórico de preços das ações*



Fonte: Autoria própria.

Nenhum dos cálculos necessários nesta aplicação, como o de retorno e risco, por exemplo, pode ser feito sem que antes a base de dados esteja normalizada. Assim, os dados ficarão postos em termos de resultado e não estarão em grandezas distintas. O Gráfico 2 mostra os preços históricos do Gráfico 1, mas agora normalizados. É notável que há uma modificação na hierarquia dos ativos quando vistos em termos de resultado. Como podemos ver, a BOVA11.SA que antes ocupava a linha mais acima, fica longe disso em termos de desempenho. Isso se deve pelo fato de que em termos de grandeza ela é muito mais expressiva, no entanto, em termos de desempenho outras ações se sobressaem.

Gráfico 2 - *Histórico de preços normalizado*

Fonte: Autoria própria.

Nesta análise precisaremos, ainda, dividir os dados em dois períodos distintos, buscando avaliar o resultado da aplicação da teoria das carteiras por um investidor que utilizou a aplicação. Devemos, portanto, "quebrar" essa série histórica em duas, utilizando as funções de filtro da biblioteca Pandas ou baixando as duas séries separadamente.

A primeira série precisa incluir os preços a partir do dia 1 de janeiro de 2018 ao dia 31 de dezembro de 2022, para todas as ações e será usada pelo investidor para encontrar os pesos ótimos de sua carteira. A segunda incluirá os preços do ano de 2023, e ela será a série onde, de fato, os resultados serão analisados; em relação ao mercado de modo geral e a uma estratégia simplória.

Portanto, as análises de cunho estatístico e descritivo vão se ater à primeira série temporal, onde se aplicará a Teoria das Carteiras e a Otimização. A segunda série demonstrará o desempenho resultante dessa aplicação no ano seguinte ao da análise. Essa demonstração de desempenho está no subtópico 4.5 deste capítulo.

4.3 CÁLCULO DO RETORNO E RISCO

O cálculo dos retornos diários para todas as ações é feito através de um comando graças a biblioteca Pandas. Assim, o retorno diário é dado pela variação percentual do preço da ação em relação ao preço do dia anterior. A função da biblioteca pandas *pct_change()* retornou, a partir dos dados de preço, as variações diárias como mencionado na Figura 14.

Figura 14 - *Dataframe com os retornos diários das ações*

Ticker	ABEV3.SA	BBAS3.SA	BBDC4.SA	BPAC11.SA	ITSA4.SA	ITUB4.SA	PETR4.SA	SANB11.SA	VALE3.SA	WEGE3.SA
Date										
2018-01-03	0.002305	0.012754	0.004598	0.009524	0.011786	0.006837	0.009063	-0.008249	-0.005992	-0.015188
2018-01-04	-0.005520	0.009595	0.016407	0.003669	0.018817	0.020824	0.001797	0.012015	0.004099	-0.009741
2018-01-05	0.003700	0.000000	0.005664	0.016534	0.001759	0.002882	0.005977	0.010655	0.015610	0.004098
2018-01-08	-0.001843	0.002376	-0.000282	0.012254	-0.000878	-0.003316	0.011883	0.019357	0.022227	0.008163
2018-01-09	-0.002770	-0.009481	-0.009296	-0.006842	-0.015817	-0.010870	0.000000	-0.007081	-0.003701	0.016194
...	...	...	...	...	...	...	...	...	...	...
2022-12-23	0.013870	0.024688	0.017473	0.062446	0.016588	0.015726	0.047103	0.020261	0.006647	-0.011299
2022-12-26	-0.006840	-0.001417	-0.023778	-0.024898	-0.012821	-0.019849	-0.007166	-0.011348	0.009152	-0.008052
2022-12-27	0.000000	-0.033210	-0.004736	-0.023441	-0.010626	-0.005670	0.006816	0.003228	0.023878	-0.004713
2022-12-28	0.002066	0.017910	0.030591	0.049721	0.020286	0.020366	-0.012346	0.015374	-0.002242	0.018153
2022-12-29	-0.002062	0.001731	-0.000660	-0.022866	-0.004678	-0.001996	-0.012097	-0.007042	-0.001236	-0.004910

Fonte: Autoria própria.

O retorno anual pode ser calculado utilizando a média dos retornos diários multiplicada pelo número médio de pregões realizados por ano; adotamos 252 dias de pregão para o cálculo, conforme indicado no referencial teórico. Para nossa base de dados, temos as taxas de retorno anual para as ações analisadas, como na Tabela 2.

A partir dessas taxas de retorno anuais, conseguimos tão somente verificar os retornos que cada ação obtiveram por ano em média. Em comparação com a BOVA, por exemplo, esta teve um retorno anual médio de cerca de 12% no mesmo período. Algumas ações da nossa carteira, portanto, obtiveram nesse mesmo período um retorno de quase 4 vezes o retorno da Bolsa brasileira.

No entanto, como sinaliza a teoria das carteiras de Markowitz, ainda mais importante que o retorno de um ativo é o risco gerado por ele. Este risco é calculado através da variância e do desvio padrão do retorno. Desse modo, expressa a constância ou a volatilidade dos retornos do ativo.

Os riscos individuais dos ativos são calculados através da função *.std* que calcula o desvio padrão das colunas de uma base de dados, aplicada ao *dataframe* das taxas de retorno dos ativos. Tratando novamente em termos médios anuais, precisamos multiplicar cada desvio por 252. Assim, a Tabela 2 também inclui essas taxas calculadas para cada ativo da carteira.

Tabela 3 - ***Retorno e risco médios anuais das ações - (2018 - 2022)***

CÓDIGO DA AÇÃO UTILIZADA	EMPRESA	TAXA DE RETORNO MÉDIO ANUAL (%)	DESVIO PADRÃO MÉDIO ANUAL (R$)
PETR4.SA	PETROBRAS	35,48	7,72
VALE3.SA	VALE	31,24	6,52
ITUB4.SA	ITAÚ UNIBANCO	6,88	5,38
ABEV3.SA	AMBEV S/A	0,32	4,94
BBDC4.SA	BRADESCO	5,53	5,98
BBAS3.SA	BANCO DO BRASIL	15,28	6,57
WEGE3.SA	WEG	37,1	6,10
BPAC11.SA	BTGP BANCO	49,16	7,99
SANB11.SA	SANTANDER BRASIL	11,1	5,87
ITSA4.SA	ITAÚSA	10,25	4,95

Fonte: Autoria própria.

Em uma carteira de ações a interação entre estes riscos individuais, com a diversificação dos investimentos, diminui o risco total para o investidor. A nossa aplicação, realiza os cálculos estatísticos básicos do risco individual dos ativos e também da carteira. Mas, além disso, através da biblioteca Seaborn, apresenta a correlação entre os ativos em um mapa de calor (Gráfico 3), o que nos possibilita enxergar como os ativos estão correlacionados e, portanto, o quão diversificada está a nossa carteira.

As ações de um mesmo setor de atividade econômica apresentam, geralmente, uma correlação alta. Os bancos presentes na nossa carteira (BBAS3.SA, BBDC4.SA, BPAC11.SA, ITUB4.SA e SANB11.SA) são um ótimo exemplo desse fenômeno, pois possuem sempre correlação positiva substancial, forte ou muito forte entre si.

Podemos observar a partir desse mapa, também, a correlação positiva muito forte que existe entre os ativos ITUB4.SA e ITSA4.SA, o que se explica pelo fato de que a Itaúsa (ITSA4.SA) é uma holding que, dentre outras atividades, controla o Itaú Unibanco (ITUB4.SA). Assim, a Itaúsa, mesmo sendo uma empresa de investimentos que detém participações em diversas empresas, possui um alta correlação com a sua principal empresa representada, o Itaú Unibanco.

Gráfico 3 - *Matriz de Correlação - Heatmap*

Fonte: Autoria própria.

Os cálculos de risco da carteira, a partir dos riscos individuais dos ativos na aplicação são realizados através da matriz de covariância multiplicada pelos pesos dos ativos, o Gráfico 3 exemplifica a matriz de covariância calculada para a base de dados. Além disso, é importante salientar que, como os pesos ótimos ainda não foram calculados, quaisquer cálculos de risco serão feitos a partir de pesos “chutados” ou randômicos apenas como exemplificação. Neste caso, utilizaremos pesos iguais (10%) para todas as ações para apresentar os comandos e resultados do cálculo da volatilidade (Figura 15) e do risco não sistemático (Figura 16) da carteira. Os cálculos de multiplicação de matrizes são realizados através da função *dot* da biblioteca Numpy.

Figura 15 - *Cálculo da Matriz de Covariância*

```
# Calcula matriz de covariância

yearly_covariance = returns.cov() * 252
yearly_covariance
```

✓ 0.6s

Ticker Ticker	ABEV3.SA	BBAS3.SA	BBDC4.SA	BPAC11.SA	ITSA4.SA	ITUB4.SA	PETR4.SA	SANB11.SA	VALE3.SA	WEGE3.SA
ABEV3.SA	0.043395	0.017454	0.024330	0.029748	0.021016	0.020191	0.012260	0.020114	0.008128	0.014945
BBAS3.SA	0.017454	0.058377	0.031888	0.040354	0.030836	0.033997	0.029317	0.026700	0.008815	0.008406
BBDC4.SA	0.024330	0.031888	0.087454	0.064029	0.046006	0.052975	0.018620	0.045137	0.017662	0.012327
BPAC11.SA	0.029748	0.040354	0.064029	0.119682	0.049708	0.052950	0.027251	0.050966	0.024816	0.021405
ITSA4.SA	0.021016	0.030836	0.046006	0.049708	0.045752	0.044812	0.019520	0.036510	0.016429	0.012498
ITUB4.SA	0.020191	0.033997	0.052975	0.052950	0.044812	0.054698	0.019093	0.039831	0.015832	0.009692
PETR4.SA	0.012260	0.029317	0.018620	0.027251	0.019520	0.019093	0.101334	0.013090	0.019833	0.003221
SANB11.SA	0.020114	0.026700	0.045137	0.050966	0.036510	0.039831	0.013090	0.057640	0.016624	0.015076
VALE3.SA	0.008128	0.008815	0.017662	0.024816	0.016429	0.015832	0.019833	0.016624	0.071316	0.010254
WEGE3.SA	0.014945	0.008406	0.012327	0.021405	0.012498	0.009692	0.003221	0.015076	0.010254	0.064063

Fonte: Autoria própria.

Figura 16 - *Comandos para variância e volatilidade da carteira*

```
# Calcula a variância da carteira

portfolio_variance = np.dot(pesos_iguais, np.dot(yearly_covariance, pesos_iguais))
portfolio_variance
```

✓ 0.0s

```
0.03013042482143855
```

```
# Calcula a volatilidade da carteira

portfolio_vol = np.sqrt(portfolio_variance)
portfolio_vol
```

✓ 0.0s

```
0.17358117646057866
```

Fonte: Autoria própria.

O risco não sistemático precisa ser calculado de maneira que resulte da subtração entre a variância do portfólio, calculada como na Figura 16 e as variâncias individuais ponderadas pelos pesos. A figura 17 demonstra os passos utilizados para chegar ao resultado de 0,1584 de risco não sistemático da carteira. Esse valor, como vimos, é menor que a volatilidade da carteira (0,1735), o que apenas confirma a existência do risco sistemático.

Figura 17 - *Cálculo do Risco Não Sistemático*

```
variance_pesos_iguais = (returns.var()*252) * pesos_iguais
variance_pesos_iguais
```
✓ 0.0s

```
Ticker
ABEV3.SA      0.004339
BBAS3.SA      0.005838
BBDC4.SA      0.008745
BPAC11.SA     0.011968
ITSA4.SA      0.004575
ITUB4.SA      0.005470
PETR4.SA      0.010133
SANB11.SA     0.005764
VALE3.SA      0.007132
WEGE3.SA      0.006406
dtype: float64
```

```
sub = -variance_pesos_iguais.sum()
sub
```
✓ 0.0s

```
-0.07037101479671659
```

```
portfolio_variance
```
✓ 0.0s

```
0.03013042482143855
```

```
#Calcula o risco não sistemático

non_systematic_risk = (portfolio_variance - sub)
non_systematic_risk
```

```
0.1584850306655196
```

Fonte: Autoria própria.

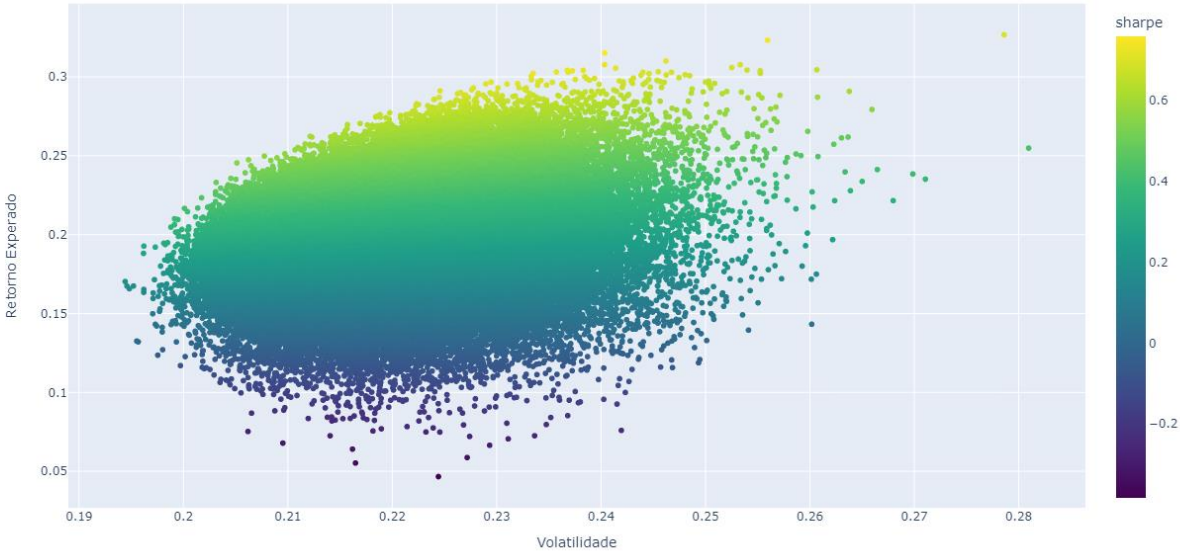
4.4 FRONTEIRA EFICIENTE DE ATIVOS

Os cálculos estatísticos, desde o retorno dos ativos individuais e da carteira até o seus riscos individuais e conjuntos, fornecem importantes direcionamentos para o investidor. Uma análise descritiva e estatística, por exemplo, já consegue afirmar se um investimento é mais seguro do que outro. No entanto, a aplicação tem como principal objetivo fornecer ao investidor um meio simples para aplicar a teoria.

Deste modo, o conceito de fronteira eficiente dos ativos é importante para demonstrar os pontos ótimos de combinação entre risco e retorno em qualquer carteira de investimentos. Estes pontos vão sendo determinados através das combinações de diferentes pesos para cada ativo da carteira. Empiricamente falando, a fronteira eficiente de ativos de Markowitz vem sendo apelidada de “Markowitz Bullet” ou, em tradução livre, Bala de Markowitz, pela semelhança que há entre o seu formato e o de um projétil.

A estimação dessa fronteira pode ser feita através de diversos métodos, dentre eles o método de Monte Carlo, onde o software atribui pesos randomicamente e recebe como resultado um índice de Sharpe para cada atribuição. A depender do número de simulações (tamanho da amostra) escolhido no momento da execução da aplicação, a fronteira estimada ganha a forma indicada na literatura como vemos na Figura 18 (utilizamos cem mil simulações de carteira através do método de Monte Carlo). Cada ponto da fronteira representa uma combinação de risco e retorno de uma composição de carteira. Os pontos dispostos na borda dessa “nuvem” de resultados formam uma fronteira ótima de investimentos.

Figura 18 - ***Fronteira Eficiente (Markowitz Bullet) - 100 Mil Simulações***



Fonte: Autoria própria.

4.5 OTIMIZAÇÃO DA CARTEIRA

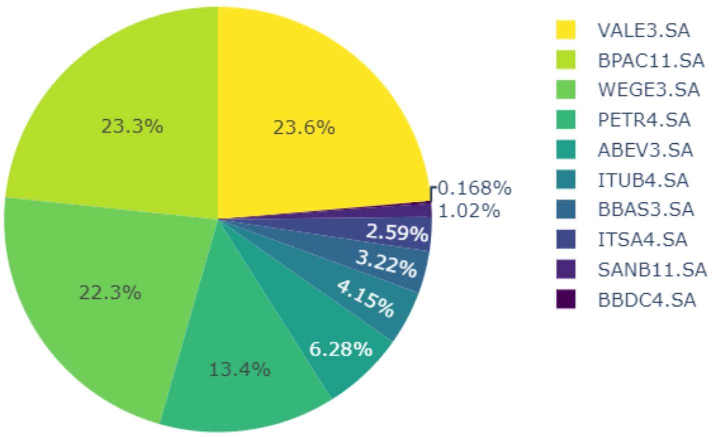
A partir dos cálculos de retorno e risco das ações, além de estimar e plotar a fronteira eficiente de ativos, precisamos determinar a melhor alocação para a carteira de ações. É importante relembrar que a escolha de uma alocação ótima necessita de um critério objetivo, uma vez que existem diversas combinações entre risco e retorno na fronteira eficiente de ativos. Portanto, como sinalizamos anteriormente, utilizaremos um bom e bastante utilizado indicador para comparar carteiras, o índice de Sharpe.

Para isso, utilizaremos as carteiras simuladas na estimação feita com o método de Monte Carlo e também os três algoritmos de busca local abordados anteriormente, quais sejam Hill Climb (Subida de Encosta), Simulated Annealing (Têmpera Simulada) e Algoritmo Genético.

Desse modo, todos os métodos e algoritmos que serão apresentados objetivam a maximização desse índice. Como explicado no referencial teórico, precisamos de um ativo livre de risco que representa um investimento seguro disponível no mercado. No Brasil, a taxa Selic determina a taxa básica de juros da economia, portanto, os ganhos dos investimentos mais seguros tendem a se aproximar da taxa Selic, dessa forma, utilizaremos a média da taxa Selic nos anos de 2018-2022 (período da base de dados) como ativo livre de risco. Para isso faremos o download da série histórica da taxa selic através da biblioteca externa SGS, que utiliza a API do Banco Central do Brasil para acessar essas informações.

A estimação da fronteira eficiente de ativos foi realizada através da simulação aleatória de cem mil carteiras distintas. Uma função presente na nossa aplicação (*best_portflolio*) permite calcular o índice de Sharpe para cada um delas e verifica qual a carteira com o maior índice. A carteira otimizada randomicamente (Gráfico 4). A carteira otimizada através do método de Monte Carlo foi bastante diversificada. Incluiu 4 ações com pesos preponderantes: VALE3.SA (23,6%), BPAC11.SA (23,3%), WEG3.SA (13,4%) e PETR4.SA (13,4%) e as outras 6 ações dividiram quase 20% do valor total. Essa foi a primeira "carteira otimizada", alcançando o índice de Sharpe 0,8615.

Gráfico 4 - ***Alocação da Carteira Otimizada Randomicamente***



Fonte: Autoria própria.

Os próximos três métodos de otimização se distinguem pelo algoritmo utilizado. No entanto, a biblioteca que realiza o procedimento é a mesma, a mlrose. E, por isso, alguns parâmetros demandados pela própria biblioteca para realizar a otimização de um conjunto de dados qualquer são comuns a todas elas. Uma função que retorne o Sharpe Ratio (resultado que se pretende maximizar) para cada conjunto que o algoritmo de busca local teste é o primeiro requisito. Em geral, essa função que processa os dados e retorna o resultado para cada iteração é chamada de *fitness_function*. Na nossa aplicação ela é definida como na Figura 19.

Figura 19 - *"Fitness_function" utilizada nos algoritmos de busca local*

```
def fitness_function(solucao):
    # retorno simples
    r = returns
    mean_returns = r.mean() * 252

    # matriz de covariância
    covariance = np.cov(r[1:].T)

    # gerando pesos aleatórios
    k = solucao
    w = k / sum (k)

    # retorno
    R = np.dot(mean_returns, w)

    # risco
    vol = np.sqrt(np.dot(w.T, np.dot(covariance, w))) * np.sqrt(252)

    # sharpe ratio

    risk_free = 0.1325
    sharpe = (R - risk_free)/vol

    return sharpe
✓ 0.0s
```

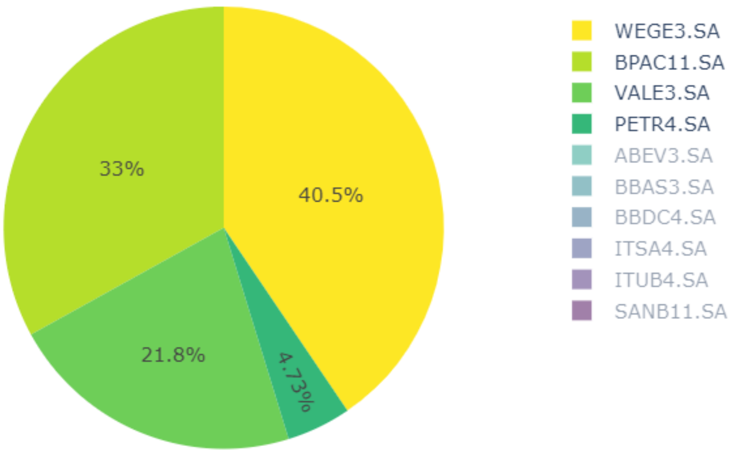
Fonte: Autoria própria.

Podemos observar que nela, assim como nos demais cálculos de retorno e risco, é utilizada a constante 252 para anualizar a base, como explicado no referencial teórico.

Além dessa função, um problema de maximização ou minimização é requerido para a realização da otimização, bastando apenas sinalizar se o problema se trata de uma maximização ou minimização e de uma otimização contínua ou discreta, além do número de variáveis, no nosso caso de ações. Neste caso, pretendemos maximizar o Sharpe Ratio (índice de Sharpe) em um intervalo contínuo de pesos. *mlRose.ContinuosOpt(maximize=True, length=10)*. Assim, descrevemos a função *fitness* e o problema de maximização comuns aos algoritmos.

O algoritmo da subida de encosta, “*hill_climb*”, apresentou pesos bastante concentrados em 4 ações, e irrisórios para as outras 6 (por isso aparecem com a legenda apagada no Gráfico 5), o que acaba por diminuir em termos práticos a quantidade de ações da nossa carteira. Um ponto que merece destaque é a concentração de quase metade do valor total alocado na ação WEG3.SA (40,5%), seguida por BPAC11.SA (33%), VALE3.SA (21,8%) e PETR4.SA (4,73%). Dessa forma, o algoritmo selecionou dentre as 10 ações uma combinação de pesos que alcançou um índice de Sharpe de 1,009. Como vimos no capítulo de metodologia, esse algoritmo funciona de maneira simples e objetiva para alcançar o máximo do espaço de dados.

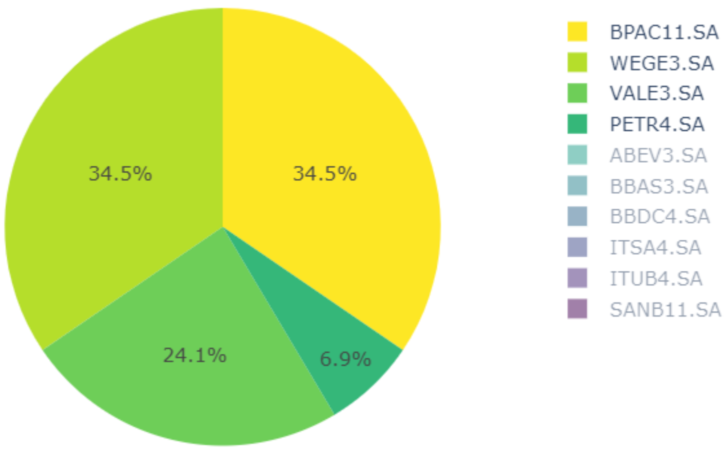
Gráfico 5 - *Alocação da Carteira Otimizada - Hill Climb*



Fonte: Autoria própria.

O algoritmo da têmpera simulada, “*simulated_annealing*”, por sua vez, retornou uma carteira com pesos significativos somente para as mesmas 4 ações retornadas pelo “*hill_climb*” (de mesmo modo as outras 6 ações se encontram com legenda apagada no gráfico 6). A alocação, por sua vez, não teve a mesma preponderância anterior: a ação BPAC11.SA obteve a maior alocação empatada com a ação WEG3.SA, ambas com 34,5%, seguidas de VALE3.SA (24,1%) e PETR4.SA (6,9%). Manteve, portanto, a concentração na alocação das ações. O índice de Sharpe obtido nesse caso foi 1,006, praticamente o mesmo valor alcançado pelo algoritmo anterior.

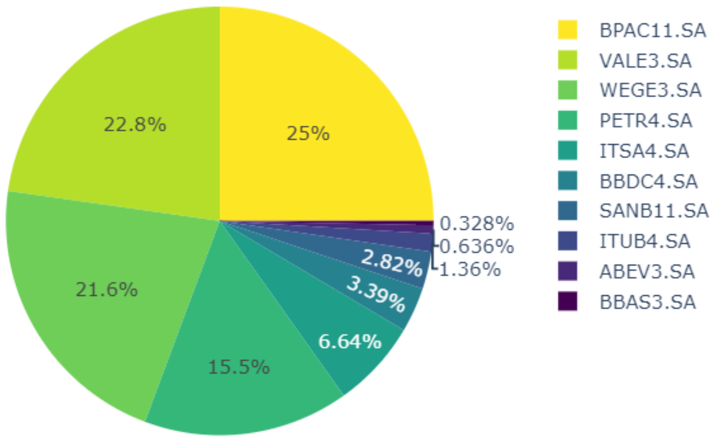
Gráfico 6 - ***Alocação da Carteira Otimizada - Simulated Annealing***



Fonte: Autoria própria.

Por fim, o algoritmo genético retornou a alocação mais variada dentre os três algoritmos (Gráfico 7), com pesos relevantes para todas as ações. Isso decorre da característica do algoritmo, uma vez que o “cruzamento” ou *crooss-over* resulta na alocação de todas ações na carteira. As três ações com maiores alocações de pesos continuaram as mesmas, no entanto nota-se uma dispersão meior na alocação do carteira em relação ao padrão apresentado nos dois algoritmos anteriores, tanto que as ações BPAC11.SA, VALE3.SA, WEG3.SA e PETR4.SA ficaram com alocações inferiores a 1/4 do valor total, enquanto todas as demais ações tiveram participação na carteira. O algoritmo genético atingiu um índice de Sharpe de 0,8763.

Gráfico 7 - *Alocação da Carteira Otimizada - Algoritmo Genético*



Fonte: Autoria própria.

Em todos os portfólios otimizados o valor do índice de Sharpe foi positivo, o que indica que o retorno da carteira está compensando o seu risco. Além disso, o comparativo (*benchmark*) das carteiras utilizando o índice de Sharpe é válido. Quanto maior for o seu índice de Sharpe, melhor o investimento em comparação aos demais. Entende-se que a magnitude (valor absoluto) do índice de Sharpe expressa como o retorno se sai em relação ao desvio-padrão do investimento em relação à taxa de risco. Um adendo importante a ser feito, ainda, é que o índice de Sharpe é um indicador de retorno ajustado ao risco e, portanto, um maior índice de Sharpe não resulta necessariamente em maior retorno e ganho, como verificaremos adiante. Além disso, mesmo com um índice de Sharpe menor, por apresentar maior diversificação em relação às outras carteiras, esta tende a apresentar um menor risco, como enfatiza a teoria das carteiras de Markowitz.

4.6 COMPARATIVO DE DESEMPENHO NO ANO DE 2023

A partir do histórico de preços das ações de 1 de janeiro de 2018 a 31 de dezembro de 2022, utilizando o método de Monte Carlo e três algoritmos de busca local, estimamos 4 carteiras com alocações ótimas. Nesta etapa, apresentaremos um comparativo dos resultados da aplicação desses pesos ótimos durante o ano de 2023.

Para isso, consideramos a compra fictícia de 10 mil reais em ações, nos percentuais apontados pelas carteiras ótimas, no primeiro dia de bolsa do ano de 2023. Acompanhamos a evolução do capital até o último dia de bolsa de 2023.

Faremos, portanto, um comparativo entre as carteiras otimizadas pelos métodos abordados e uma carteira composta somente por ações BOVA11.SA, que representam o índice IBOVESPA; ambas também iniciarão com 10 mil reais. O *dataframe* na Figura 20 mostra a evolução do capital de todas essas carteiras (5 primeiros e 5 últimos dias) e nos permite entender como o Gráfico 8 é composto: diariamente o retorno dos ativos de cada carteira, quando ponderados por seus respectivos pesos na carteira, resultarão no retorno diário da carteira, o que implica no valor total da carteira.

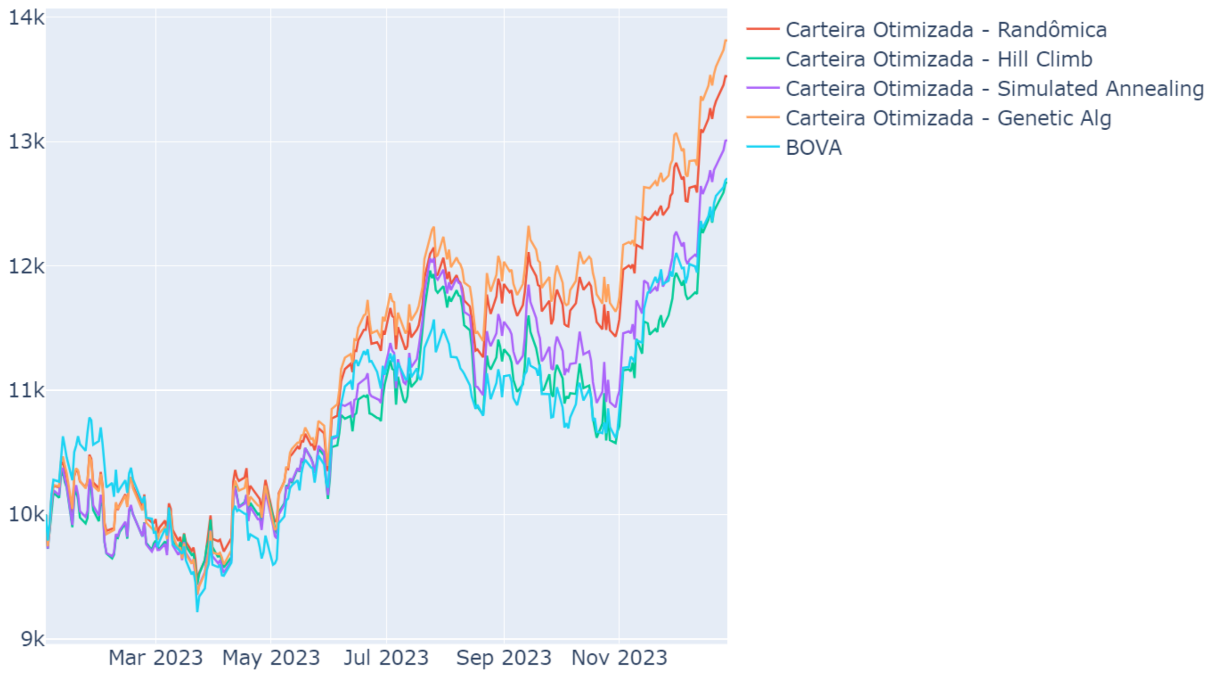
Figura 20 - *"Dataframe" Evolução de Capital*

	Carteira Otimizada - Randômica	Carteira Otimizada - Hill Climb	Carteira Otimizada - Simulated Annealing	Carteira Otimizada - Genetic Alg	BOVA
Date					
2023-01-02	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000
2023-01-03	9771.588827	9725.285862	9727.099257	9745.181292	9790.510187
2023-01-04	9915.796128	9870.730937	9880.334753	9900.783339	9930.819536
2023-01-05	10136.841510	10064.708017	10090.795476	10133.496837	10159.797275
2023-01-06	10233.293537	10174.631136	10195.344473	10237.482150	10279.645603
...	...	...	...	...	...
2023-12-21	13274.713649	12443.173085	12772.940164	13545.038652	12506.090767
2023-12-22	13324.110675	12471.670057	12803.198056	13602.634017	12563.578475
2023-12-26	13456.899002	12591.266328	12930.805429	13738.354164	12632.758196
2023-12-27	13526.855922	12665.043223	13006.922230	13814.204501	12689.272066
2023-12-28	13525.071684	12671.795873	13010.880656	13814.639774	12704.862407

Fonte: Autoria própria.

Conforme vemos no Gráfico 8, os resultados de evolução do capital e, portanto, do retorno das carteiras, superaram os do IBOVESPA, exceto para o algoritmo da Subida de Encosta, consideradas as suas limitações. Isto é esperado, uma vez que o índice IBOVESPA tem uma amplitude muito significativa e absorve muitos dos impactos negativos do mercado, embora apresente retorno positivo para o ano analisado.

Gráfico 8 - *Comparativo - Evolução de Capital*



Fonte: Autoria própria.

As carteiras otimizadas, portanto, apresentaram melhores resultados. Dentre elas, uma carteira otimizada com algoritmo de busca local superou a de otimização randômica (método de Monte Carlo), a carteira otimizada com o Algoritmo Genético. Como indicado, nem sempre um maior índice de Sharpe representará maior retorno da carteira, ele mede o retorno ajustado ao risco. E mesmo tendo um índice menor que as duas outras carteiras otimizadas por algoritmos de busca local, a carteira otimizada com Algoritmo Genético se saiu melhor. Os melhores resultados foram o do Algoritmo Genético, finalizando o ano com capital total de R$ 13.814,63 e do Método de Monte Carlo, que utiliza o poder computacional para gerar resultados aleatórios e escolher o melhor dentre os resultados, finalizando o ano com R$13.525,07. A tabela 4 contém os lucros de todas as carteiras, bem como o índice de Sharpe calculado no momento da otimização com os dados da primeira série de tempo.

No nosso estudo, a ordem dos melhores resultados não coincidiu com a ordem dos índices de Sharpe calculados com os dados da primeira série temporal (2018 a 2022).

Tabela 4 - ***Comparativo dos Lucros das Carteiras***

CARTEIRA	ÍNDICE DE SHARPE (DADOS DE 2022)	LUCRO[22] DA CARTEIRA EM 2023	LUCRO DA CARTEIRA EM 2023 (%)
OTIMIZADA COM ALGORITMO GENÉTICO	0,8763	R$ 3.814,63	38,14
OTIMIZADA RANDOMICAMENTE	0,8625	R$ 3.525,07	35,25
OTIMIZADA COM SIMULATED ANNEALING	1,006	R$ 3.010,88	30,1
OTIMIZADA COM HILL CLIMB	1,009	R$ 12.671,89	26,71
BOVA	-	R$ 2.704,86	27,04

Fonte: Autoria própria.

4.6.1 Cálculo do CAPM dos ativos e do portfólio

O cálculo do retorno esperado por meio do Modelo de Precificação de Ativos Financeiros (CAPM) desempenha um papel crucial na avaliação e gestão de carteiras de investimentos. O CAPM oferece uma abordagem sistemática para estimar o retorno que um investidor deve esperar com base no risco associado a um ativo específico. Assim, também comparamos o Retorno Esperado calculado a partir do CAPM com o retorno verificado anteriormente, para cada uma das carteiras. A Figura

[22] O lucro expresso na Tabela 3 não considera Imposto de Renda e Taxa de Corretagem; sendo, portanto, o lucro bruto da operação.

21 demonstra os passos da nossa aplicação para calcular o CAPM de cada ação individualmente.

Figura 21 - *Cálculo do CAPM dos Ativos*

```
# Calculando os BETAs e ALPHAs de todas as ações
bova_returns = prices_bova.pct_change().dropna()
itar e Depurar (Ctrl+Shift+D)
betas = []
alphas = []
for asset in returns:
    print(asset)
    beta, alpha = np.polyfit(bova_returns, returns[asset], 1)
    betas.append(beta)
    alphas.append(alpha)
```

```
#Calculando o CAPM para todas as ações

risk_free = media_selic
market_risk = bova_returns.mean() * 252

capm_assets = []
for i, asset in enumerate(returns):
    capm_assets.append(risk_free + (betas[i] * (market_risk-risk_free)))

capm_assets
✓ 0.0s
```

Fonte: Autoria própria.

A Figura 22 mostra os passos utilizados na aplicação para o cálculo do CAPM para uma carteira de ações, abaixo:

Figura 22 - *Cálculo do CAPM do Portfólio*

```
# Calculando o CAPM da Carteira

capm_portfolio = np.sum(capm_assets * pesos_hill_climb['weights'])*100
capm_portfolio
✓ 0.0s
```

Fonte: Autoria própria.

A Tabela 5 faz um comparativo entre os retornos que observamos das carteiras e o CAPM calculado para cada carteira, lembrando que o cálculo do CAPM da carteira é feito apenas ponderando cada CAPM individual dos ativos pelo peso do ativo na carteira em questão.

Tabela 5 - ***Comparativo - Retorno Esperado (CAPM) e Retorno Observado***

CARTEIRA	RETORNO ESPERADO CALCULADO PELO CAPM (%)	RETORNO OBSERVADO EM 2023 (%)
OTIMIZADA COM ALGORITMO GENÉTICO	24,43	38,14
OTIMIZADA RANDOMICAMENTE	24,07	35,25
OTIMIZADA COM SIMULATED ANNEALING	24,48	30,1
OTIMIZADA COM HILL CLIMB	25,21	26,71

Fonte: Autoria própria.

Portanto, verifica-se ainda que todas as carteiras superaram o Retorno Esperado calculado pelo CAPM, que é um parâmetro de mínimo retorno aceitável para o investidor, dado o risco envolvido no investimento.

5. CONSIDERAÇÕES FINAIS

Neste trabalho, exploramos a aplicação da Moderna Teoria do Portfólio de Markowitz na otimização de carteiras de investimentos, utilizando ações das 10 empresas mais valiosas da Bolsa de Valores de São Paulo (B3) em 2023. Inicialmente, apresentamos o mercado de capitais brasileiro, seu funcionamento, órgãos reguladores e a importância do mercado de valores mobiliários. Em seguida, discutimos a natureza dos dados dos pregões da bolsa e como esses dados são organizados e divulgados, fundamentais para a análise realizada neste trabalho. O trabalho, outrossim, proporcionou uma compreensão teórica e prática da aplicação da Moderna Teoria do Portfólio de Markowitz, demonstrando a importância da diversificação e otimização de carteiras de investimentos para a maximização do Sharpe Ratio e a busca por uma fronteira eficiente entre risco e retorno.

Para facilitar a aplicação do conhecimento teórico, utilizamos da linguagem Python e dos algoritmos de inteligência artificial para desenvolver uma aplicação que nos permitisse estimar carteiras otimizadas e ainda realizamos um estudo de caso onde simulamos dois momentos de análise: uma de decisão e outra de verificação de desempenho de um investimento. Dessa forma, buscamos implementar uma teoria sólida para otimização de pesos em um investimento. Essa otimização pode ser o diferencial entre um investimento arbitrário e um investimento teoricamente fundamentado. Portanto, a teoria de Markowitz nos serviu como fundamento e as tecnologias empregadas (Python e IA) como facilitadores para esse fim.

No segundo momento, analisamos a evolução do patrimônio resultante da utilização dos pesos ótimos estimados, e comparamos o desempenho das carteiras otimizadas com o desempenho da BOVA (Índice Bovespa). A análise realizada revelou a capacidade de obter estimativas de carteiras ótimas com ganhos significativos, mesmo considerando as limitações inerentes ao risco de mercado.

Contudo, é imperativo ressaltar que, além das considerações proporcionadas pela modelagem computacional, outros fatores cruciais impactam a decisão final na seleção de uma carteira de investimentos, especialmente em cenários de incerteza no mercado financeiro. Aspectos como o cenário econômico global e local, questões políticas relacionadas ao mercado financeiro e o perfil individual de cada investidor não podem ser negligenciados.

A aplicação desenvolvida neste trabalho foi projetada para ser acessível e eficaz, especialmente para usuários leigos. Uma sugestão de contribuição para esta análise é desenvolver para esta aplicação uma interface gráfica. Isso ajudaria os usuários menos familiarizados com códigos a utilizar a aplicação de maneira mais cômoda.

.

REFERÊNCIAS

Araújo, A. F**. A Teoria de Markowitz e a Otimização de Carteiras de Investimentos.** Tese de Doutorado, Universidade de São Paulo, 2010.

Assaf Neto, A, & Guasti Lima, F. **Investimento em Ações**: Guia Teórico e Prático para Investidores. Inside Books, 2008.

Assaf Neto, A. **Finanças Corporativas e Valor**. 8ª Ed. Atlas, 2022.

Assaf Neto, A. **Mercado Financeiro**. 14ª Ed. Atlas, 2018.

Bodie, Z., Kane, A., & Marcus, A. J. **Investments**. McGraw-Hill Education, 2015.

Costa, W. **A história do Python**, 2023. Disponível em: https://www.dio.me/articles/a-historia-do-python.

Acesso em 20/09/2023.

Cerbasi, G. **Investimentos Inteligentes**. Sextante, 2018.

Damodaran, A. **Investment Valuation: Tools and Techniques for Determining the Value of Any Asset**. John Wiley & Sons, 2007.

DE MEDEIROS, G. L. B.; DE MEDEIROS, L. N. P. Ausência de educação financeira no Brasil: O impacto à sociedade e a possibilidade de reversão / Lack of financial education in Brazil: The impact on society and the possibility of reversing. **Brazilian Journal of Development**, [S. l.], v. 7, n. 10, p. 101408–101417, 2021. DOI: 10.34117/bjdv7n10-449. Disponível em: https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/38778.

Acesso em: 17 de setembro de 2023.

Doane, D. P., & Seward, L. E. **Applied Statistics in Business and Economics**. McGraw-Hill Education, 2014.

Fabozzi, F. J. **Modern Portfolio Theory and Risk Management**. John Wiley & Sons, 2022.

Guerra, J. P. **Teoria Moderna de Portfólio - De Markowitz a Estratégias de Alocação de Ativos no Brasil**. Elsevier, 2009.

Hayes, G. **mlrose Documentation**: Release 1.3.0. 2019. Disponível em: https://mlrose.readthedocs.io/_/downloads/en/stable/pdf.

Acesso em: 20/09/2023.

Hilpisch, Y. **Python for Finance**: Analyze Big Financial Data. O'Reilly Media, 2014.

Hull, J. C. **Options, Futures, and Other Derivatives**. Pearson, 2017.

Malkiel, B. G. **A Random Walk Down Wall Street**. W. W. Norton & Company, 2015.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. Técnicas de pesquisa. **São Paulo: Atlas**, 1990.

Markowitz, H. **Portfolio Selection**. The Journal of Finance, 7(1), 77-91,1952.

Markowitz, H. **Portfolio Selection**: Efficient Diversification of Investments. John Wiley & Sons, 1959.

McKinney, W. **Python for Data Analysis:** Data Wrangling with Pandas, Numpy, and Jupyter. 3th Ed. O'Reilly Media, 2022.

Miguel, Franklin & Ramos, Dorel. (2017). **Analysis of PROINFA Power Plants Portfolio from the Perspective of Markowitz**. IEEE Latin America Transactions. 15. 1650-1656. 10.1109/TLA.2017.8015048.

Moura, F. R. **Carteiras Eficientes e Ingênuas**: uma análise comparativa com o uso do modelo de Markowitz. Monografia. Departamento de Economia. Universidade Federal de Sergipe. 2009.

Python Package Index. **Documentações das bibliotecas**. 2023. Disponível em: https://pypi.org/.

Acesso em: 20//09/2023.

Pinheiro, J. **Mercado de Capitais**. 9ª - Ed. Atlas, 2019.

ROSOCHA, L.; VERNEROVA, S.; VERNER, R. **Medical staff scheduling using simulated annealing**. Quality Innovation Prosperity, 2015.

Russell, S. J., & Norvig, P. **Artificial Intelligence**: A Modern Approach. Pearson, 2020.

Santos, A. R., & Tessari, T. L. **Matriz de covariância condicional no cálculo de risco de mercado: estimação e seleção**. Dissertação de Mestrado, Universidade Federal de Minas Gerais, 2012.

SHARDA, R.; DELEN, D.; TURBAN, E.; **Business Intelligence e Análise de Dados para Gestão do Negócio**. 4 ed. Porto Alegre - RS: 2019.

Sharpe, William F. (1966). **Mutual fund performance**. Journal of Business, 39(1), 119-138.

Walliman, N. **Métodos de Pesquisa**. 1 ed. São Paulo: Editora Saraiva, 2014.

Ward, M. et al. "Empirical testing of the capm on the jse". **Investment Analysts Journal**, vol. 41, no. 76, 2012, p. 1-12. Disponível em: https://doi.org/10.1080/10293523.2012.11082546

Acesso em: 24/09/2023.

World Exchange. **Market Statistics – September 2023**. Disponível em: https://focus.world-exchanges.org/issue/september-2023/market-statistics. Acesso em: 17 de setembro de 2023.

ZANINI, F. A. M.; FIGUEIREDO, A. C.. AS TEORIAS DE CARTEIRA DE MARKOWITZ E DE SHARPE: UMA APLICAÇÃO NO MERCADO BRASILEIRO DE AÇÕES ENTRE JULHO/95 E JUNHO/2000. **RAM. Revista de Administração Mackenzie**, v. 6, n. 2, p. 38–65, abr. 2005.