

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS,
ENGENHARIA DE PRODUÇÃO E SERVIÇO SOCIAL

EDEN DE OLIVEIRA SOUZA

CIÊNCIA DE DADOS APLICADA AO INVESTIMENTO EM
AÇÕES COMO SUPORTE À APOSENTADORIA PÚBLICA

ITUIUTABA
2025

EDEN DE OLIVEIRA SOUZA

CIÊNCIA DE DADOS APLICADA AO INVESTIMENTO EM AÇÕES COMO SUPORTE À APOSENTADORIA PÚBLICA

Trabalho de Conclusão de Curso apresentado à Faculdade de Administração, Ciências Contábeis, Engenharia de Produção e Serviço Social da Universidade Federal de Uberlândia, como requisito parcial para obtenção do título de bacharel em Engenharia de Produção.

Orientador: Prof. Dr. Eugênio Pacceli Costa

ITUIUTABA
2025

CIÊNCIA DE DADOS APLICADA AO INVESTIMENTO EM AÇÕES COMO SUPORTE À APOSENTADORIA PÚBLICA

Trabalho de Conclusão de Curso apresentado à Faculdade de Administração, Ciências Contábeis, Engenharia de Produção e Serviço Social da Universidade Federal de Uberlândia, como requisito parcial para obtenção do título de bacharel em Engenharia de Produção.

Ituiutaba, 24 de abril de 2025.

Aos meus pais, pela base que me deram. À minha namorada, pelo apoio constante e companheirismo. Aos meus avós e tio, por fazerem parte da minha história e incentivo. Este trabalho leva em si um pouco de cada um de vocês.

AGRADECIMENTOS

Registro neste espaço meus sinceros agradecimentos a todos que, de alguma forma, contribuíram para a realização deste trabalho e para a minha jornada de crescimento pessoal e acadêmico.

Primeiramente, agradeço profundamente aos meus pais, Osmar e Kenia, que foram pilares fundamentais nesta jornada. Pelo incentivo constante à minha evolução, pela dedicação incansável e pelos muitos sacrifícios realizados para que eu pudesse alcançar o objetivo de concluir o curso superior em uma universidade federal. Alcançar a graduação em Engenharia e ser o primeiro da nossa família a conquistar esse título em uma instituição federal é um marco que carrego com muito orgulho. Os desafios ao longo desse percurso foram incontáveis, mas sem o apoio, o amor e a confiança de vocês, eu não teria chegado até aqui.

Ao olhar para trás, percebo claramente a transformação que vivi ao longo dessa jornada. Cada desafio, aprendizado e conquista foi essencial para minha formação, tanto pessoal quanto profissional. A experiência na Universidade Federal de Uberlândia foi decisiva, marcando um ponto de inflexão significativo em minha vida. Foi nesse ambiente que tive a oportunidade de crescer e me desenvolver, superando limites e me moldando para o futuro.

Aos docentes do curso de Engenharia de Produção, minha eterna gratidão. O conhecimento compartilhado nas aulas, as experiências enriquecedoras, as pesquisas e publicações realizadas durante o curso contribuíram significativamente para a formação sólida que recebi. Sou imensamente grato por cada ensinamento e por todas as oportunidades de aprendizado que me foram proporcionadas, permitindo-me alcançar níveis de desenvolvimento que muitas vezes, me surpreendem.

Agradeço, de forma especial, ao Prof. Dr. Eugênio Pacceli Costa, cuja orientação foi essencial para minha trajetória acadêmica. Em meio aos desafios da Engenharia de Produção, sua condução me ajudou a definir um caminho claro. Sou grato pela confiança, dedicação e pelo valioso conhecimento compartilhado, que foram fundamentais para minha formação profissional.

Cada um de vocês desempenhou um papel fundamental na pessoa que me tornei e na conclusão deste trabalho. Este Trabalho não é apenas o resultado do meu esforço, mas, sem dúvida, é fruto da contribuição de todos que fizeram parte desta jornada. Esta conquista não é exclusivamente minha, mas de todos nós. Agradeço, do fundo do coração, por todo o apoio, dedicação e por estarem ao meu lado neste momento tão significativo da minha vida.

“A sorte é o que acontece quando a preparação
encontra a oportunidade.”
(Seneca)

RESUMO

Diante da tendência de reversão da pirâmide etária brasileira, torna-se necessário considerar alternativas para complementar ou até mesmo substituir a renda dos beneficiários da previdência pública, em caso de um eventual colapso do sistema. Esta pesquisa tem como objetivo verificar a viabilidade de estratégias de construção de portfólios de investimento em ações, como alternativa ao sistema previdenciário público, visando mitigar os impactos de um possível colapso, com base na análise de dados e em métricas de desempenho financeiro, como retorno esperado e risco associado. O portfólio foi estruturado com base em métodos de seleção de ações fundamentados nas estratégias de renomados investidores do mercado financeiros: Benjamin Graham, Décio Bazin, Luiz Barsi e Joel Greenblatt. A análise dos dados foi realizada com o uso da linguagem Python, permitindo a automação dos processos de extração, transformação e carregamento (ETL). Para a comparação entre os ativos, foi aplicada a Teoria Moderna do Portfólio de Markowitz, que considera o retorno esperado, risco individual, correlação e exposição entre os ativos, para estimar o risco e retorno da carteira. Os resultados demonstram que todos os métodos analisados se mostraram viáveis para a complementação da renda no longo prazo, mesmo diante de oscilações no curto prazo. O método Graham destacou-se por apresentar o maior retorno e o menor risco, sendo o único que, segundo o Índice de Sharpe, apresentou uma relação risco-retorno superior a do ativo livre de risco. O método Barsi, gerou o maior risco. O método Bazin proporcionou o maior volume de dividendos no último ano do portfólio. Conclui-se que, sob uma perspectiva de análise fundamentalista e considerando a estratégia de longo prazo, os métodos estudados representam alternativas viáveis para a formação de uma renda complementar à aposentadoria pública, podendo gerar, inclusive, uma renda mensal média superior a 20% do salário-mínimo atual, além de permitir a realização de lucros conforme a necessidade do investidor.

Palavras-chave: Ciência de dados; Aposentadoria Complementar; Investimento em ações; Linguagem Python; Seleção de ações.

ABSTRACT

Given the trend of the inversion of Brazil's demographic pyramid, it becomes essential to consider alternatives to complement or even replace the income provided by the public pension system, in the event of its potential collapse. This study aims to assess the feasibility of equity investment portfolio strategies as an alternative to the public pension system, with the goal of mitigating the impacts of such a scenario based on data analysis and financial performance metrics, including expected return and associated risk. The investment portfolio was structured using stock selection methods inspired by the strategies of renowned financial market investors: Benjamin Graham, Décio Bazin, Luiz Barsi, and Joel Greenblatt. Data analysis was performed using the Python programming language, enabling the automation of extraction, transformation, and loading (ETL) processes. To compare the selected assets, Markowitz's Modern Portfolio Theory was applied, considering expected returns, individual asset risk, correlations, and exposure between assets to estimate the portfolio's overall risk and return. The results demonstrate that all the strategies analyzed proved viable for generating complementary income in the long term, despite short-term fluctuations. The Graham strategy stood out for presenting the highest return and the lowest risk, being the only one that, according to the Sharpe Ratio, displayed a risk-return ratio superior to that of a risk-free asset. The Barsi strategy yielded the highest risk, while the Bazin strategy provided the greatest volume of dividends in the portfolio's final year. In conclusion, from a fundamental analysis perspective and considering a long-term investment strategy, the studied methods represent viable alternatives for building a complementary income stream to the public pension system. These strategies may even generate an average monthly income exceeding 20% of the current minimum wage, while also allowing profit realization according to the investor's needs.

Keywords: Data science; Complementary retirement; Stock investment; Python programming; Stock picking

LISTA DE FIGURAS

Figura 1	Distribuição da população em conjunto de idades no ano de 1970	17
Figura 2	Distribuição da população em conjunto de idades no ano de 2000	18
Figura 3	Distribuição da população em conjunto de idades no ano de 2022	18
Figura 4	Projeção da distribuição da população em conjunto de idades em 2040	19
Figura 5	Projeção da distribuição da população em conjunto de idades em 2060	19
Figura 6	Risco Sistemático x Risco Diversificável	25
Figura 7	Procedimento do desenvolvimento da pesquisa	38
Figura 8	Preço teto corrigidos por ano de teste para o método Barsi	48
Figura 9	Preço teto corrigidos por ano de teste para o método Graham	48
Figura 10	Preço teto corrigidos por ano de teste para o método Bazin	49
Figura 11	Selic mensais nos últimos 25 anos	55
Figura 12	Taxa mensais da Selic acrescido do valor um	55
Figura 13	Relatório para acompanhar performance dos métodos	66
Figura 14	Relatório para acompanhar indicadores de cada método de investimento	67

LISTA DE TABELAS

Tabela 1	Carteira aplicando método Barsi	42
Tabela 2	Carteira aplicando método Bazin	44
Tabela 3	Carteira aplicando método Graham	45
Tabela 4	Carteira aplicando método Greenblatt	46
Tabela 5	Salários-mínimos utilizados na pesquisa	47
Tabela 6	Taxa Selic acumulada por ano	56
Tabela 7	Retorno e porcentagem de cada ativo para o método Barsi	57
Tabela 8	Retorno e porcentagem de cada ativo para o método Bazin	57
Tabela 9	Retorno e porcentagem de cada ativo para o método Graham	58
Tabela 10	Retorno e porcentagem de cada ativo para o método Greenblatt	58
Tabela 11	IS para cada método de investimento	63

LISTA DE SIGLAS E ABREVIATURAS

ACID	Atomicidade, Consistência, Isolamento e Durabilidade
API	<i>Application Programming interface</i>
BESST	Bancos, Energia, Seguros, Saneamento e Telecomunicação
BVPS	<i>Book Value Per Share</i>
CAGR	<i>Compound Annual Growth Rate</i>
CAP	Consistência, disponibilidade e Tolerância
CORR	Correlação
CPF	Cadastro de Pessoa Física
DPNE	<i>Debt Per Net Equity</i>
EBIT	<i>Earnings Before Interest and taxes</i>
ERP	<i>Enterprise Resource Planning</i>
ETF	<i>Exchange-Traded Funds</i>
ETL	<i>Extract, Transform and Load</i>
EPS	<i>Earnings Per Share</i>
EV	<i>Enterprize Value</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
INSS	Instituto Nacional de Seguro Social
IS	Índice Sharpe
IV	Valor intrínseco
JSCP	Juro Sobre o Capital Próprio
KDD	<i>Knowledge Discovery in database</i>
LGPD	Lei geral de proteção de dados
MD ³	<i>Make Data-DRIVEN Decisions</i>
NoSQL	<i>Not Only SQL</i>
PE	<i>Price to Earnings</i>
ROE	<i>Return on Equit</i>
SELIC	Sistema Especial de Liquidação e Custódia
SGBD	Sistema de Gerenciamento de dados
SQL	<i>Structured Query Language</i>

SUMÁRIO

1 INTRODUÇÃO	14
1.1 CONTEXTUALIZAÇÃO.....	14
1.2 OBJETIVOS DE PESQUISA	15
1.2.1 <i>Objetivo Geral</i>	15
1.2.2 <i>Objetivos específicos</i>	15
1.3 JUSTIFICATIVA	15
1.4 DELIMITAÇÃO DO TRABALHO	16
1.5 ESTRUTURA DO TRABALHO	16
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1. O PROBLEMA DA PREVIDÊNCIA PÚBLICA	17
2.2 INVESTIMENTO E DIVERSIFICAÇÃO DE PORTFÓLIO.....	20
2.3 INVESTIMENTO EM RENDA VARIÁVEL.....	20
2.4 INDICADORES DE ANÁLISE DE INVESTIMENTOS	21
2.5 ANÁLISE DE INVESTIMENTO	23
2.5.1 <i>Portfolio Selection – Harry Markowitz</i>	23
2.5.2 <i>Fórmula mágica de Greenblatt de seleção de ativos</i>	27
2.5.3 <i>Fórmula de Ben Graham</i>	27
2.5.4 <i>Método Bazin</i>	28
2.5.5 <i>Método Barsi</i>	29
2.6 DADOS, INFORMAÇÃO E CONHECIMENTO	29
2.7 PROCESSO DE ETL	31
2.8 CIÊNCIA DE DADOS.....	33
2.8.1 <i>Linguagem de programação</i>	35
2.8.2 <i>Bibliotecas e API Python</i>	35
2.8.3 <i>Business Intelligence</i>	36
3 METODOLOGIA	37
3.1 CARACTERIZAÇÃO DA PESQUISA.....	37
3.1.1 <i>Natureza de pesquisa</i>	37
3.1.2 <i>Procedimento de pesquisa</i>	37
3.2 ETAPAS DE PESQUISA.....	38
4 RESULTADOS E DISCUSSÃO	41
4.1 COLETA DE DADOS, APLICAÇÃO DOS MÉTODOS E MONTAGEM DA CARTEIRA	41
4.1.1 <i>Método Barsi</i>	41
4.1.2 <i>Método Bazin</i>	42
4.1.3 <i>Método Graham</i>	44
4.1.4 <i>Método Greenblatt</i>	45
4.2 SIMULAÇÃO DAS CARTEIRAS.....	46
4.2.1 <i>Definição da quantia disponível para investimento</i>	46
4.2.2 <i>Correção dos valores</i>	47
4.3 SIMULAÇÃO E ALOCAÇÃO DOS ATIVOS.....	49
4.3.1 <i>Simulação método Barsi</i>	49
4.3.2 <i>Simulação método Bazin</i>	51
4.3.3 <i>Simulação método Graham</i>	52

4.3.4	<i>Simulação método Greenblatt</i>	53
4.4	ANÁLISE DA SIMULAÇÃO	55
4.4.1	<i>Cálculo da Selic média e seu desvio padrão</i>	55
4.4.2	<i>Cálculo do retorno do portfólio</i>	56
4.4.3	<i>Cálculo do risco do portfólio</i>	59
4.4.4	<i>Índice de viabilidade do retorno em relação ao risco para cada método de investimento e indicadores de rentabilidade</i>	63
4.5	ANÁLISE DE RESULTADOS	67
5	CONCLUSÕES	68
	REFERÊNCIAS	71

1 INTRODUÇÃO

A presente seção tem como objetivo apresentar o contexto do cenário que fundamenta a pesquisa desenvolvida, destacando os elementos essenciais que motivaram sua realização. Serão discutidos o objetivo geral do estudo, as justificativas que sustentam sua relevância, as delimitações impostas ao escopo do trabalho e, por fim, a forma como o texto está estruturado.

1.1 Contextualização

De acordo com Domingues, Ferreira e Neves (2024), com a expectativa de vida aumentando, e a taxa de natalidade retraindo no Brasil, há uma tendência que se ocorra um déficit na previdência pública ao passar dos anos. Ainda complementam que existe um teto para a base de cálculo das contribuições para a Previdência Social, assim, caso algum trabalhador esteja fora deste teto, quando aposentado receberá menos que está acostumado, gerando, assim, uma queda no padrão de vida habitual.

Porém, a situação tende a piorar com o passar do tempo, pois o Brasil está sofrendo um processo de envelhecimento de sua população, fazendo com o que o número de pessoas que se beneficiam deste programa social tende a crescer, enquanto o número de contribuintes esteja em queda (Domingues, Ferreira e Neves, 2024). Sendo necessários novas formas para se complementar a renda no futuro, sendo o mercado de capitais uma alternativa.

O mercado de capitais tem como objetivo interligar pessoas superavitárias com deficitárias, caracterizando, assim, a fonte de financiamento do setor privado/consumo. Exercendo um papel importante no processo econômico do Brasil (Oliveira, 2024). Possibilitando que o investidor esteja diretamente ligado a negócios de diversos setores, sem a necessidade de grandes aportes financeiros (Santos, 2022).

Porém, segundo Rodrigues (2022), existem variadas modalidades de investimentos, dificultando a escolha da melhor estratégia, que seja rentável e faça o investidor se sentir confortável. Como exemplo, tem-se o mercado de ações com diferentes tipos de companhias, e escolher a “melhor” ação não é uma tarefa simples. Contudo, com o desenvolvimento da internet e técnicas estatística, em geral o público passou a ter mais acesso à informação o que contribui com o processo de escolha de ativos.

Neste âmbito, é importante destacar que o processo de tomar decisões dependerá de conhecimento, habilidade e a atitude, além de depender de fatores psicológicos, culturais, políticos e financeiro que cada um tenha acesso (Basilio, Freitas e Kampffe, 2018).

Segundo Grus (2016), quem possui o conhecimento em ferramentas de *Data Science* consegue extrair conhecimento de dados desorganizados, construindo *insights* que podem guiar para melhores estratégias, permitindo adotar a postura de MD³ (*Make data-driven Decisions*) (Afsharian, 2024).

Diante desse contexto, a proposta desta pesquisa é estudar a viabilidade de uma alternativa para complementar a renda dos indivíduos em caso de colapso da Previdência Social, por meio da análise de estratégias de investimento no mercado de capitais.

1.2 Objetivos de pesquisa

Nesta seção são apresentados o objetivo geral da pesquisa e os objetivos específicos que orientam este estudo

1.2.1 Objetivo Geral

Verificar a viabilidade de estratégias de construção de portfólios de investimento em ações, como alternativa ao sistema previdenciário público, visando mitigar os impactos de um possível colapso, com base na análise de dados e em métricas de desempenho financeiro, como retorno esperado e risco associado.

1.2.2 Objetivos específicos

Os objetivos específicos são:

- Construir um portfólio de ações seguindo as estratégias de quatro investidores relevantes, considerados gurus do mercado financeiro e de capitais;
- Analisar a relação retorno e risco com a aplicação das estratégias de investimento, tendo como base de avaliação à teoria *Portfolio Selection*;
- Comparar fluxos de caixa gerados por dividendos em cada estratégia.
- Analisar a viabilidade das estratégias de investimento como complemento da renda previdenciária e desenvolver um relatório visual para se acompanhar os ativos.

1.3 Justificativa

Este estudo, justifica-se pela possibilidade de aplicar ferramentas de ciência de dados para realizar uma análise aprofundada de grandes conjuntos de dados financeiros. Isso permite embasar decisões de investimento de longo prazo com maior precisão, para a construção de

portfólios que podem complementar ou até mesmo substituir, a previdência pública, em caso de colapso no futuro

1.4 Delimitação do Trabalho

Este estudo delimita-se à aplicação de ferramentas de análise de dados para avaliar o desempenho de quatro estratégias de investimento, comparando seus retornos e riscos ao longo de 25 anos, por meio de um *backtest*, simulando um investidor que aplicou mensalmente o valor equivalente à contribuição obrigatória do INSS (Instituto Nacional do Seguro Social). Diferentemente de estudos anteriores (Basilio, Freitas e Kampffe 2018; De Oliveira, De Souza, Bendere *et.al.*, 2023; Domingues e Ferreira, 2024; Gelatti e Do Nascimento Afonso, 2024; Kleinert, 2024; Morais, 2024; Oliveira, 2019; Oliveira, 2024; Santos, 2022), que analisam estratégias de investimento em períodos mais curtos com foco comparação de retorno absoluto baseado somente em valores iniciais aplicados e cesta de ações mais arbitrárias, esta pesquisa investiga sua viabilidade como alternativa ou complemento à previdência social. O estudo considera um horizonte mais longo, permitindo avaliar se as flutuações do mercado se tornam menos relevantes no longuíssimo prazo. O período de 25 anos foi escolhido devido à limitação de disponibilidade histórica dos dados de preços dos ativos, sendo este o intervalo máximo acessível. Ainda que inferior aos 35 anos atualmente exigidos para a aposentadoria por tempo de contribuição, trata-se do recorte temporal mais próximo possível dentro das condições técnicas viáveis para a realização do estudo.

1.5 Estrutura do trabalho

Este trabalho implementa ferramentas de análise de dados, para auxiliar na tomada de decisão de escolha de ativos, que irão compor um portfólio. Será utilizada a linguagem de programação Python, com a incorporação de bibliotecas para auxiliar na execução do processo de ETL (*Extract, Transform and Load*). O estudo se limitará ao mercado financeiro e de capitais brasileiro, utilizando exclusivamente ações (*tickers*, que se referem aos códigos de negociação das ações) como ativos nas estratégias analisadas.

2 FUNDAMENTAÇÃO TEÓRICA

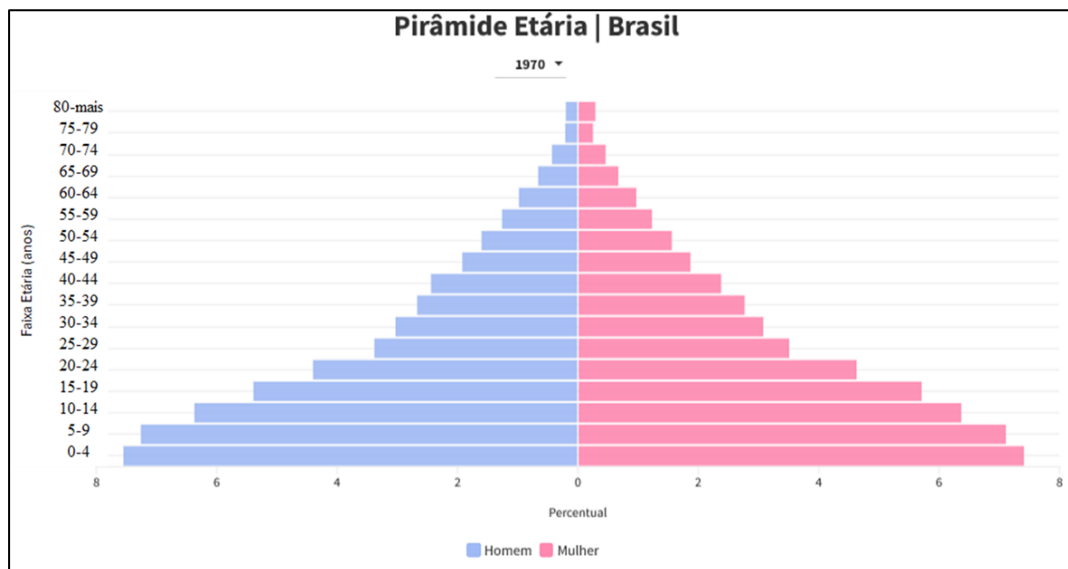
Neste bloco, são apresentados e fundamentados conceitos essenciais relacionados ao mercado financeiro, incluindo estratégias de investimento, bem como termos amplamente utilizados em ciência de dados.

2.1. O problema da previdência pública

De acordo com Ministério da Previdência Social (2024), os regimes públicos de previdência têm caráter obrigatório e mensal para todos os que exercem atividades remuneradas. Para os contribuintes com carteira assinada os descontos são realizados diretamente da folha de pagamento. O Ministério complementa ainda que o INSS atua como um seguro social, no qual os trabalhadores ativos são responsáveis por financiar a aposentadoria dos beneficiários.

Figura 1, mostra a distribuição de pessoas no ano de 1970 por idade.

Figura 1 - Distribuição da população em conjunto de idades no ano de 1970

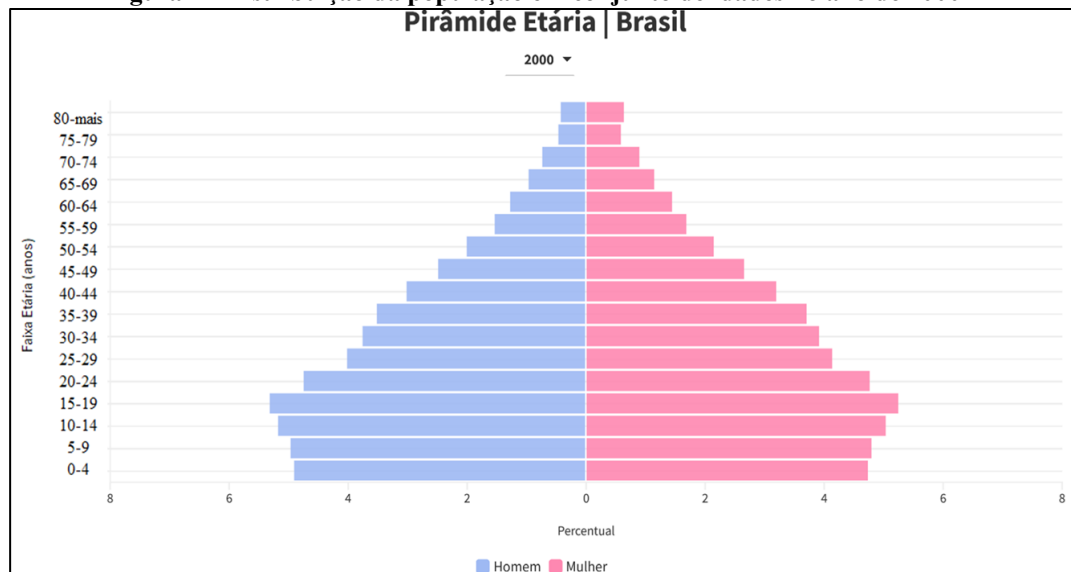


Fonte: Instituto Brasileiro de Geografia e Estatística (IBGE), Censo Demográfico (1970)

Pode-se observar que a maior parte da população “estava” abaixo dos 30 anos, o que indica uma situação favorável para os beneficiários do seguro social com idade superior a 60 anos, uma vez que há um maior número de trabalhadores ativos contribuindo para o sistema.

A seguir, a Figura 2, é mostrado a distribuição de pessoas no ano de 2000 por idade.

Figura 2 - Distribuição da população em conjunto de idades no ano de 2000

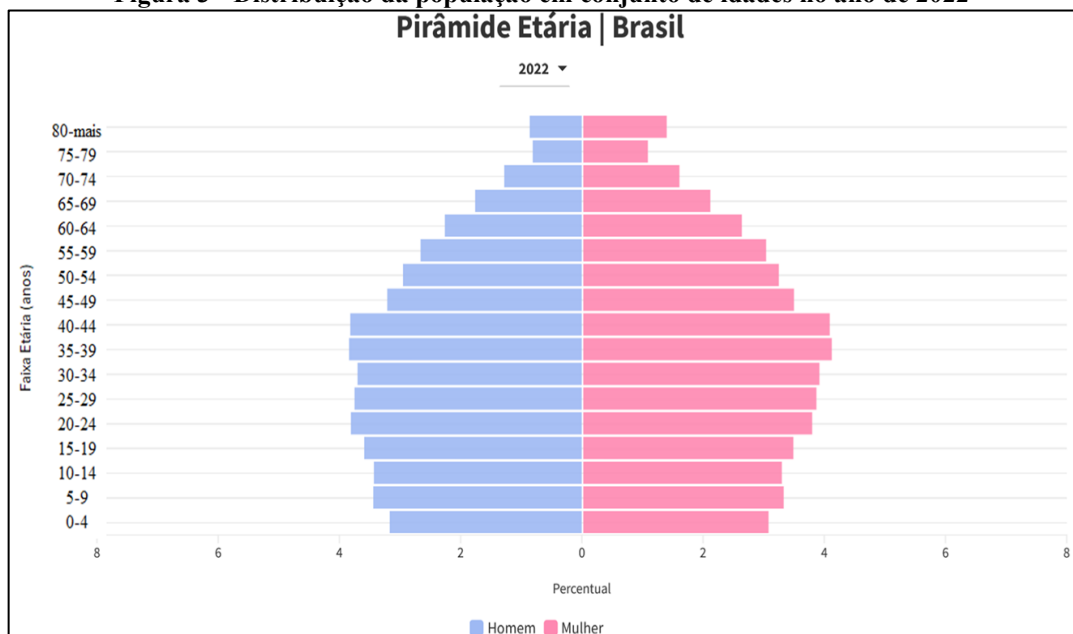


Fonte: Censo Demográfico – IBGE (2000)

No ano de 2000, verifica-se uma redução da população ativa e um aumento dos beneficiários, dando a entender que a população está envelhecendo.

Figura 3, mostra a distribuição de pessoas no ano de 2022 por idade:

Figura 3 - Distribuição da população em conjunto de idades no ano de 2022



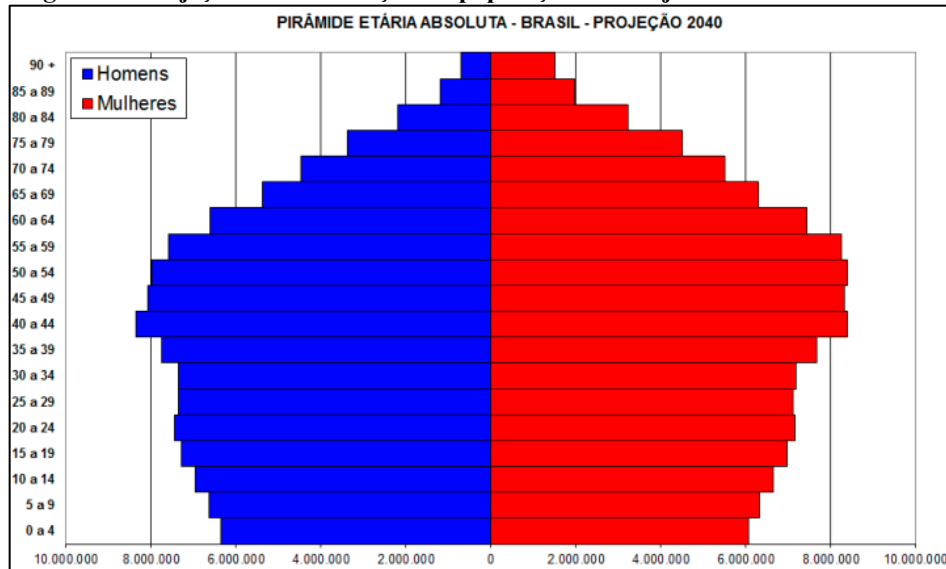
Fonte: Censo Demográfico – IBGE (2022)

Na Figura 3 se nota a inversão nas taxas de natalidades, seguindo um padrão de inversão de pirâmide. O que pode desequilibrar economicamente a balança entre aposentados e

trabalhadores formais implicando em maior déficit futuro da previdência social pública brasileira, já que as futuras gerações de trabalhadores terão mais aposentados para subsidiar.

O IBGE elaborou projeções demográficas para os anos de 2040 e 2060, representadas nas figuras 4 e 5, que ilustra a distribuição estimada da população em cada ano respectivamente.

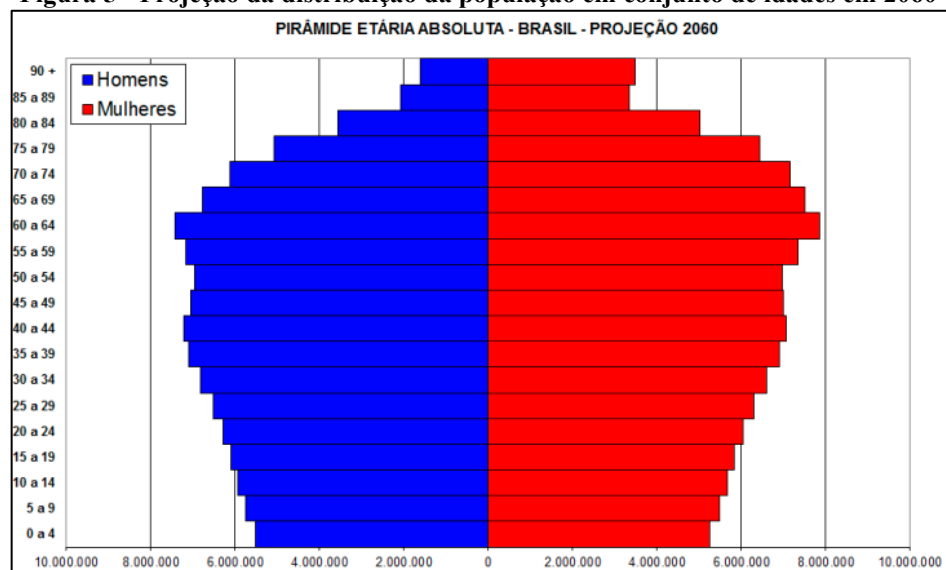
Figura 4 - Projeção da distribuição da população em conjunto de idades em 2040



Fonte: IBGE (2018)

A projeção mostra o possível acréscimo no número de indivíduos para os anos de 50 a 60, aumentando a distribuição no centro da pirâmide.

Figura 5 - Projeção da distribuição da população em conjunto de idades em 2060



Fonte: IBGE (2018)

A projeção indica que os grupos etários mais representativos estão migrando para a extremidade superior da figura, enquanto extremidade inferior continua a reduzir. Ou seja, de acordo com a projeção, está se aproximando o momento em que haverá menos pessoas trabalhando do que aposentados, o que pode levar ao colapso do sistema previdenciário

2.2 Investimento e diversificação de portfólio

Segundo Oliveira (2019), investir envolve o desenvolvimento de estratégias e o estabelecimento de metas para precificar uma carteira de ativos, com o objetivo de gerar resultados financeiros a partir de um risco calculado, que é mitigado por meio da diversificação.

Segundo Basilio, Freitas e Kampffe (2018), o termo da diversificação de investimentos vem sendo estudado há décadas. Devido a procura das pessoas em encontrar algo que faça crescer as suas riquezas produzindo uma sensação de segurança. Para isso se busca juntar alguns ativos e formar um portfólio que o risco agregado seja inferior ao individual dos ativos. Realizar a alocação eficiente tem como objetivo maximizar o retorno, e isso tende a resultar em uma distribuição mais equilibrada e segura dos pesos entre os ativos, o que aumenta a probabilidade de um desempenho consistente e minimiza os riscos associados. Considera-se que uma carteira de investimentos é segura quando a probabilidade de o investidor perder todo o capital – denominada nível de ruína – mantém-se abaixo de um limite previamente definido, de acordo com o perfil de risco do investidor (Mittal, Bhattacharya e Mandal, 2022).

2.3 Investimento em renda variável

Recentemente, o mercado financeiro nacional passou a oferecer uma maior variedade de investimentos, atraindo o interesse de investidores locais e internacionais. Pode-se notar isso pelo aumento significativo no número de Cadastro de Pessoa Físicas (CPF's) que se registraram na bolsa atualmente, mostrando o aumento de interesse e participação de investidores no mercado de Renda Variável e Fixa (Morais, 2024).

Segundo Moraes (2024), renda variável são aqueles produtos cujos rendimentos não se sabe ou não se pode prever quais serão, pois, esses necessitam de que alguma coisa que no futuro venha a acontecer. Devido a essa incerteza, esses tipos de ativos podem gerar maiores ganhos, mas atrelado a isso se vem o risco de perdas maiores. Sendo o oposto do que seria a renda Fixa, que tendem a não sofrer flutuações.

Os tipos de investimentos em Renda Variável segundo Moraes (2024), são:

- **Ações:** títulos que representam uma parte do capital social de alguma empresa;

- **Fundo de investimento em ações:** são fundos que necessitem que 67% do seu patrimônio líquido estejam em ações. Este produto é para pessoas que desejam investir em empresas, mas não possuem conhecimento ou tempo para realizar uma análise mais aprofundada de cada negócio e terceiriza a decisão de se investir;
- **Commodities:** são produtos básicos que são produzidos em alta escala e comercializados;
- **Derivativos:** seu valor se baseia em outro ativo e são usados para garantir contra as variações de seus valores.

Segundo De Oliveira, De Souza e Bender *et al.* (2023), o aumento de brasileiros na Bolsa de Valores, está provocando uma transformação no mercado de Renda Variável. Os ativos que mais têm se destacado, devido a sua performance, são as ações. Os autores afirmam que no terceiro trimestre de 2021, as empresas aumentaram o pagamento de dividendos, devido às preferências por distribuir partes dos lucros (dividendos), ao invés de fazer novos investimentos. Além disso, mais 90% das empresas mapeadas globalmente mantiveram, ou aumentaram, as suas distribuições.

2.4 Indicadores de análise de investimentos

Segundo Rodrigues (2022), *Valuation* é a metodologia para ajudar a calcular o valor das companhias. Assim, os empresários podem procurar investimentos ou até encontrar o preço justo do seu negócio quando os colocar à venda. Já os investidores, podem projetar qual os valores das frações das companhias chegarão no futuro, estimando o retorno que se terá ao se investir na companhia.

Indicadores utilizados para fazer a precificação de ativos segundo Rodrigues (2022):

- **Relação Dívida pelo patrimônio líquido (*Debt Per Net Equity – DPNE*):** determina o percentual que a dívida da companhia representa sobre o patrimônio líquido da empresa, sendo calculada pela equação (1), a seguir:

$$DPNE = \frac{\text{Dívida Líquida}}{\text{Patrimônio Líquido}} * 100 \quad (1)$$

- **Retorno sobre capital investido (*Return on Equity – ROE*):** relaciona o lucro da empresa com o patrimônio líquido. sendo calculada pela equação (2).

$$ROE = \frac{Lucro_{líquido}}{Quantidade_{ações} * preço_{ações}} * 100 \quad (2)$$

- **Lucro Por Ação (*Earnings Per Share – EPS*):** indicador que mostra o lucro para cada ação que a empresa possui. Esta é uma das mais importantes métricas para determinar o valor justo de uma empresa, sendo calculada pela equação (3).

$$EPS = \frac{Lucro\ Líquido}{Quantidade\ de\ ações} \quad (3)$$

- **Relação Preço Lucro (*Price to Earnings – PE*):** indicador que relaciona o valor de mercado de uma ação com a quantidade de lucro que a empresa gera, uma média boa considerada pelo mercado é um P/L < 15 (Domingues, Ferreira e Neves, 2024), sendo calculada pela equação (4).

$$PE = \frac{Preço\ por\ Ação}{EPS} \quad (4)$$

- **Valor Patrimonial por Ação (*Book Value Per Share – BVPS*):** indica o valor patrimonial da empresa, proporcional para cada ação que o negócio possui. Uma média boa considerada pelo mercado é um VPA < 1,5 (Domingues, Ferreira e Neves, 2024), sendo calculada pela equação (5).

$$BVPS = \frac{Patrimonio\ Líquido}{Quantidade\ de\ ações} \quad (5)$$

- **Lucros Antes de Juros e Impostos (*Earnings Before Interest and taxes – EBIT*):** lucro operacional da empresa. Este indicador, é usado para analisar o potencial de fluxo de caixa que a companhia pode alcançar, sendo um indicador usado para empresas com poucas dívidas, que é divulgado pelas empresas.
- ***Dividend Yield*:** indicador que mede a lucratividade que um investimento gerou. Mostrando a proporcionalidade entre os dividendos pagos pela empresa e o valor de cada ação cotada no mercado, sendo calculada pela equação (6).

$$Dividend\ Yield\% = \frac{Dividendos\ por\ ação}{Preço\ por\ ação} \quad (6)$$

- **Taxa de Distribuição (*Payout*):** razão entre os proventos pagos e o lucro líquido que a empresa alcançou no último ano (Santiago Nina, 2021). Informando assim a taxa que a empresa retorna de seus lucros para os seus acionistas. Não podendo comprometer a capacidade de investimento da empresa, sendo calculado pela equação (7).

$$Payout = \frac{\text{Dividendo por ação}}{EPS} * 100 \quad (7)$$

- **SELIC (Sistema Especial de Liquidação e Custódia):** é a taxa que determina o custo dos empréstimos realizados entre os bancos e com o Banco Central. Quando há mais dinheiro saindo do que entrando nas instituições financeiras, existe o risco de elas fecharem o dia com saldo negativo. Para evitar isso, os bancos recorrem a empréstimos entre si ou com o Banco Central, pagando juros conforme essa taxa (Rodrigues, 2022).

2.5 Análise de investimento

Nesta seção são apresentadas estratégias de investimentos, sendo elas a teoria de Graham, Bazin, Greenblatt e Barsi, assim como métodos para otimização de alocação de ativos financeiros utilizando o método *Portfolio Selection* criado por Markowitz.

2.5.1 *Portfolio Selection* – Harry Markowitz

Segundo Oliveira (2024), a teoria considera que a diversificação não se limitava a reduzir o risco dos ativos individualmente, mas levar em consideração como eles interagem entre si, ao serem colocados em uma carteira de investimento, então, a otimização de portfólio resultaria em uma fronteira eficiente entre o risco e o retorno esperado, e as carteiras que estivessem nessa fronteira eficiente, o investidor poderia escolher a carteira ótima a depender do seu perfil sendo a de mais ou menos risco (considera risco o desvio padrão dos retornos), mais ou menos retorno.

A Teoria Moderna do Portfólio considera que o risco presente em uma carteira de investimento apresenta dois componentes: Origem própria do ativo (podendo ser chamado de o risco do negócio ao qual a empresa está inserida, por exemplo) e o outro é de origem sistêmica. Sendo assim, o primeiro está relacionado com as características intrínsecas do ativo, podendo ser a qualidade, tipo e a relação que se tem com outro ativo da carteira, já o segundo, está

relacionado com o evento do sistema econômico, político e social como um todo. Porém, enquanto este último não se consegue influenciar, o primeiro pode tentar mitigá-lo através da diversificação (Santiago Nina, 2021).

Segundo Moraes (2024), para se fazer uma seleção de ativos são necessárias três fases de estudos:

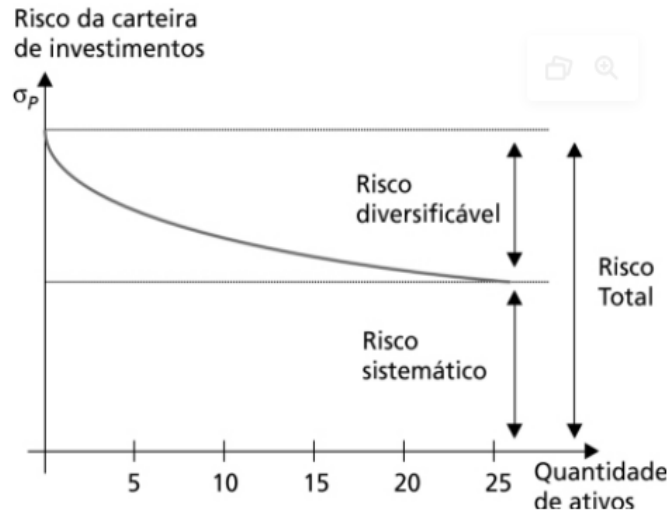
- **Análise dos títulos:** analisar o retorno esperado de cada ativo que se deseja colocar no portfólio, determinado, assim, uma Taxa Mínima de Atratividade (TMA), considerando, assim, que essa taxa possa compensar o risco do investimento e trazer retornos maiores que em uma operação livre de risco;
- **Análise de carteiras:** envolve projetar o retorno e o risco da seleção que se colocou no portfólio, utilizando ferramentas técnicas para realizar a validação destes dados;
- **Seleção de carteira:** alocar a melhor combinação de ativos, levando em consideração o apetite ao risco de cada portfólio desejado. Dentre as várias opções possíveis, o melhor portfólio é aquele que mais se adapta ao grau de satisfação do investidor.

Para se realizar a minimização do risco de um portfólio, Moraes (2024) afirma que a Teoria das Carteiras deve seguir alguns pressupostos:

1. O investidor tende a maximizar o retorno para um nível aceitável de risco;
2. Diante ativos de mesmo risco, deve-se escolher o de maior retorno;
3. Diante de ativos de mesmo retorno deve-se escolher o de menor risco;
4. Como todos os ativos estão expostos aos mesmos custos de transação e impostos, eles são desconsiderados;
5. Há uma taxa isenta de risco que o investidor pode utilizar para levantar fundos (SELIC).

Na Figura 6, a seguir, pode-se ver a relação de risco diversificável à medida que se aumenta os ativos em carteira:

Figura 6 - Risco Sistemático x Risco Diversificável



Fonte: Assaf Neto (2021)

Segundo Assaf Neto (2021), conforme se amplia a quantidade de ativos, o risco total decresce em razão da diminuição do risco diversificável (não sistemático). Porém, este processo é limitado ao risco sistemático. Ainda conclui, que a partir de certo número de ativos o risco se estabiliza ou se tem pouca variação, resultando apenas a parte sistemática. O que de acordo com a imagem se alcança próximo aos 10 ativos.

Ainda de acordo com Assaf Neto (2021), o desvio-padrão mede o grau de variação absoluta dos retornos em relação a sua média. Por isso, o risco de um ativo isolado é diferente do risco que ele representa quando inserido em uma carteira. O risco de um ativo em um portfólio é avaliado pela contribuição que ele oferece ao risco total da carteira, o que pode ser determinado por meio de uma equação específica (8).

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n W_i W_j CORR_{i,j} \sigma_i \sigma_j \quad (8)$$

Em que o W_i e W_j são as participações de cada ativo na carteira, $CORR_{i,j}$ é a correlação de cada ativo e σ_j e σ_i são os desvios padrões dos retornos de cada ativo, a fórmula de variância do portfólio calcula a contribuição de cada par de ativos usando a multiplicação dos pesos, dos desvios padrão e da correlação, e pode ser representada de forma matricial para portfólios com muitos ativos.

De acordo com Gelatti e Do Nascimento Afonso (2024), a correlação (CORR) mede como os ativos se comportam em relação ao outro, variando a $-1 < CORR_{x,y} < 1$. Se a correlação for igual a 1, significa que há uma correlação positiva perfeita, indicando que as variáveis (x,

y) se movem exatamente na mesma direção. Já uma correlação de -1 representa uma correlação negativa perfeita, o que implica que os ativos se movem em direções opostas. Quando a correlação é 0, não há relação linear entre as variáveis. No entanto, na prática, as correlações raramente são perfeitas.

Valores próximos a 1 indicam uma forte correlação positiva, ou seja, quando um ativo sobe, o outro tende a subir também, embora não necessariamente com a mesma intensidade. Valores próximos de -1 indicam uma forte correlação negativa, ou seja, quando um ativo cai, o outro tende a subir. Já valores próximos de 0 indicam correlação baixa ou nula, significando que os comportamentos dos ativos não estão relacionados.

De acordo com Assaf Neto (2021), o retorno de um ativo pode ser calculado por meio do retorno logarítmico, que representa a variação percentual dos preços ao longo do tempo, conforme demonstrado na equação (9)

$$\text{Taxa de retorno} = \ln \left(\frac{P_t}{P_{t-1}} \right) \quad (9)$$

Em que:

P_t é o preço do ativo em um tempo t;

P_{t-1} é o preço do ativo imediatamente anterior ao tempo t.

Para se calcular o Retorno do investimento em percentual se usa a equação (10).

$$\text{Retorno}_{\text{percentual}} = e^{\text{taxa}_{\text{retorno}}} \quad (10)$$

Fonte: Assaf Neto (2021)

Assim como o risco que depende da participação de cada ativo terá na carteira, o retorno esperado do portfólio também necessita esta consideração (Assaf Neto, 2021). De acordo com a equação (11).

$$E(R) = \sum_{i=1}^n W_n \times R_n \quad (11)$$

Em que:

$E(R)$ = Retorno esperado do portfólio;

W_n = Porcentagem do ativo na carteira;

R_n = Retorno do ativo que será colocado na carteira.

Assaf Neto (2021) complementa que o retorno ($E(R)$) de qualquer investimento deve ser maior que um ativo livre de risco, para validar o risco corrido ao se investir em outros tipos de ativos, logo o **Índice Sharpe (IS)** representa a relação entre o prêmio pago pelo risco assumido e o risco do investimento. Sendo calculado pela equação (12).

$$IS = \frac{E(R) - R_{RF}}{\sigma_p} \quad (12)$$

Em que:

R_{RF} = Retorno da Renda Fixa (SELIC) de menor risco disponível;

σ_p = Risco do portfólio.

2.5.2 Fórmula mágica de Greenblatt de seleção de ativos

De acordo com De Oliveira, De Souza, Bender *et al.* (2023), a “fórmula” consiste em achar empresa com alto valor (diferente de preço) e fundamentos sólidos, porém, que estejam sendo negociadas a preços descontados. O método cria um *ranking* e quanto maior for o índice de Greenblatt, mais descontada está a empresa. Seguindo a equação (13).

$$\text{GREENBLATT} = \text{Pontuação } \frac{EV}{EBIT} + \text{Pontuação ROE} \quad (13)$$

Assim a Pontuação EV (*Enterprize Value* [Valor da firma = Valor de mercado + Dívida Líquida]) ordenando o $\frac{EV}{EBIT}$ das empresas do maior para o menor, enumerando 1 para o maior e prosseguindo com a sequência nos demais. Enquanto a Pontuação ROE é alcançada ordenando o ROE do menor para o maior, do qual o menor terá pontuação 1 e assim sucessivamente (De Oliveira; De Souza; Bender *et al.*, 2023).

2.5.3 Fórmula de Ben Graham

Rodrigues (2022) diz que o método de Graham foi desenvolvido para realizar o *valuation* das empresas baseado no valor intrínseco das ações. O valor intrínseco pode ser entendido como o preço justo da ação. Sendo assim, o preço justo nem sempre será o mesmo que o de mercado, pois este último oscila. A metodologia na precificação do ativo se baseia no valor da empresa e não o preço que está sendo vendida.

Após calcular o valor justo da ação, o investidor tem em suas mãos uma ferramenta para ajudar a montar seu próprio portfólio. Para isso, escolhe ações que estão abaixo do seu preço

justo. Assim, o valor intrínseco (IV) pode ser calculado de acordo com equação (14) (RODRIGUES, 2022).

$$IV = \sqrt{BVPS_{esperado} \times EPS_{esperado} \times BVPS_{Real} \times EPS_{Real}} \quad (14)$$

Em que:

$BVPS_{esperado}$ = mostra o valor máximo que se espera pagar, sobre o patrimônio líquido.

$EPS_{esperado}$ = demonstra o valor máximo que se espera pagar, sobre o lucro

$BVPS_{Real}$ = representa quanto a empresa está sendo negociada em relação ao seu patrimônio líquido total.

EPS_{Real} = indica quanto a empresa está sendo negociada em relação ao seu lucro.

Porém, $BVPS_{esperado}$ considerado adequado pelo mercado é, no máximo 15, enquanto um $EPS_{esperado}$ aceitável é de, no máximo 1,5. Substituindo esses valores na equação (14):

$$IV = \sqrt{15 \times 1,5 \times BVPS_{Real} \times EPS_{Real}}$$

$$IV = \sqrt{22,5 \times BVPS_{Real} \times EPS_{Real}}$$

O valor de 15 significa que se espera pagar pela empresa no máximo 15 vezes o seu lucro, e 50% (ou 0,5) a mais do seu valor patrimonial.

Além de também necessitar filtros seguindo os seguintes aspectos adaptado com base no trabalho de Oliveira (2019):

- **Tamanho mínimo da empresa:** receita maior ou igual de R\$300 milhões de reais, no ano anterior ao qual será rodado o teste (2023);
- **Situação financeira estável:** empresa com a liquidez corrente ($\frac{Ativo}{Passivo}$ *Circulante*) maior ou igual a 1;
- **Lucros estáveis:** sem reportar prejuízos nos últimos cinco anos.

2.5.4 Método Bazin

Domingues, Ferreira e Neves (2024), o Método Bazin considera a média de distribuição de dividendos nos últimos 3 anos, mesmo isso não garantindo a permanência desta distribuição, mas concede uma segurança a mais para o investidor. *Payout* de no máximo 50%, para garantir

que a empresa é saudável financeiramente e conseguirá manter seus investimentos, DPNE menor que 1,75% e *Dividend Yield* maior que 6% (Oliveira, 2019).

2.5.5 Método Barsi

Método Barsi filtra empresas que estão nos setores seguindo o acrônimo BESST (Bancos, Energia, Seguros, Saneamento e Telecomunicação), o qual considera que empresas presentes nestes setores são mais propensas a serem seguras e estáveis, pois, oferecem produtos/serviços indispensáveis para a sociedade (De Oliveira; De Souza e Bender, 2023). Segundo Domingues, Ferreira e Neves (2024), o método consiste em investir em empresas que estejam sendo negociadas abaixo do $\frac{\text{Preço}}{\text{BVPS}}$. Além de se estimular um preço teto de compra de cada ativo, do qual o valor máximo que se aceite pagar em uma empresa, quando seu *Dividend yield* seja de no mínimo 6%.

2.6 Dados, informação e conhecimento

De acordo com Martins, Junior e Bortolitti (2023), com a evolução dos meios de comunicação os volumes de dados e a tecnologia estão em ascensão, ocasionando um deságio de gestão desta quantidade de elementos, porém, não é comum que se tenha condições para tratá-los de maneira apropriada os transformando em informação afim de usá-los eficientemente criando conhecimento.

De acordo com Da Silva, Peres e Boscaroli (2017), o dado é a circunstância de algo, o fato ou um valor documentado, antes de ser agregado ou relacionado. Quando um sentido é adicionado ou se agrega ele se torna a informação. O conhecimento vem da reflexão, buscando contextualizar a informação, tornando assim consciente e capaz de tomar decisões. Estes dados podem ser identificados como dados estruturados e não estruturados:

- **Dados estruturados:** são resultados de um processo de geração, observação ou medição. Sendo normalmente armazenados em tabelas. Ex: sistema de controle de caixa ou pagamentos;
- **Dados não estruturados:** são os que não possuem uma estrutura padrão ou fixa. Ex.: vídeos, música e mensagens em redes sociais.

Estes dados são encontrados de diferentes maneiras sendo elas: (Marquesone, 2016)

- **Dados internos:** são aquelas do qual as empresas são donas. Sendo eles:
 - Dados de sistemas de gerenciamento (Ex.: ERP - *Enterprise Resource Plannin*);
 - Arquivos (Ex.: Documentos escaneados ou notas fiscais);

- Documentos gerados por colaboradores (Ex.: .pdf ou .xlsx);
- Sensores sendo dados provenientes de medidores ou câmeras;
- Registro de logs seriam localização de veículos ou dados de servidores.
- **Dados da web:** são dados que podem ser coletados de fontes externas (Obs.: Respeitando os direitos de seus detentores e sua vontade em compartilhar, seguindo a LGPD – Lei geral de proteção de dados). Sendo eles:
 - Dados de domínio público: dados disponibilizados pelo governo, ou empresas privadas (empresas negociadas em bolsa de valores tem obrigação em compartilhar informações relacionadas as suas situações fiscais) o compartilhamento de informações internas;
 - Dados de sites de terceiros: sendo os que se encontram em *marketplaces* ou *blogs*;
 - Dados de mídias sociais: sendo os dados referentes a Instagram ou Youtube.

O acesso a esse conjunto massivo de dados está transformando todas as áreas de atuação, consequência do aumento grandioso na necessidade de coleta, armazenamento e recuperação desses dados, superando as capacidades de gerenciamento, sendo definido como *big data* (Sant’Ana, 2016).

De acordo com Najafabadi, Villsnustre, Khosgoftaar *et al.* (2015), *big data* geralmente refere-se aos dados que excedem o armazenamento, processamento e capacidade computacional de banco de dados convencionais e técnicas de análise de dados tradicionais. Ele surgiu pelo aumento da capacidade de armazenamento e crescimento do poder de processamento computacional, e a viabilidade de crescer o volume de dados, o que ocasionou uma maior quantidade de dados do que as empresas conseguiam processar.

Sendo assim, o *big data* necessita que os profissionais possuam competências diversas nas empresas, como extrair os dados em um momento e coordenar ações para os recuperar, criando visualizações para realizar análises criando *insights* para tomada de decisão baseada em dados (Rautenberg e Do Carmo, 2019).

Segundo Chang e Grady (2019), para ser considerado como *Big Data*, é necessário que possua o que se nomeou como 4Vs, que são:

- Volume: conjunto longo e extenso de dados para serem analisados;
- Variedade: representa que para fazer qualquer análise é necessário carregar arquivos de diferentes armazéns e possui diferentes tipos de dados;
- Velocidade: um elevado fluxo de dados, devido ao alto processado de conjuntos de dados em um tempo reduzido;

- Variabilidade: significa que o dado pode mudar a sua estrutura a qualquer momento. Sendo necessário pensarem mediadas para escalar esse processado (Ex.: *Cloud Computing*).

O *Big Data Analytics* é um processo de se descobrir potenciais *insights* dos dados.

Tendo para isso análises a seguir tipos de análises que são: (Tan, 2018)

- Análise Descritiva: refere-se a fato que já aconteceu;
- Análise Diagnóstica: refere-se a entender o porquê algo está acontecendo (Marquesone, 2016);
- Análise Preditiva: se refere ao que pode ou provavelmente vai acontecer;
- Análise Prescritiva: se tem como base, o que pode ser feito para que se alcance resultados ótimos.

Tan (2018), complementa que o *Big Data Analytics* possibilita que os gestores tomem decisões baseadas em evidências ao invés de intuição.

2.7 Processo de ETL

De acordo com Oracle (2024), o processo de *ETL* é um processo em que os gestores e empresas orientados a tomar decisões baseado dados, usam para direcionar suas pesquisas coletando os dados de várias fontes e os reunindo para criar relatórios, os analisando e tomando decisões. O *ETL data pipelines* providencia a fundação da análise de dados, baseado na regra do negócio, ele limpa e organiza os dados direcionando para o objetivo de inteligência corporativa, extraindo dados dos sistemas, os limpando para agregar valor, qualidade, estabilidade e consistência e por fim os carregando no lugar desejado para a análise (IBM, 2024). Seguindo assim as etapas:

- *Extract*: de acordo com Long (2024), o processo de extração envolve a coleta de dados, que consiste na busca de dados em qualquer localidade, sensores ou até utilizando uma API - *Application Programming interface* (Uma interface de programação que armazena instruções para a execução de tarefas, automatizando processo de coleta por exemplo), que comunica duas interfaces a fim de se coletar dados promovendo conexões específicas para realizar essa coleta. Devido ao volume que os dados possuem se faz necessário uma busca detalhada, definindo assim o termo *Data Mining*, do qual são definidos processos automatizados para se explorar de maneira analítica um grande conjunto de dados, com a finalidade de tentar encontrar padrões, sendo importantes para

embasar a compreensão da informação gerando assim o conhecimento (Da Silva, Peres e Bortolotti, 2017).

- *Transform*: a etapa de processamento de dados é fundamental para garantir a qualidade da análise e visualização subsequentes (Long, 2024). Segundo Marquesone (2016), as bases de dados frequentemente apresentam problemas estruturais, tornando essencial o tratamento adequado das informações. O conceito ‘*garbage in, garbage out*’ ressalta a importância da estruturação e limpeza dos dados para minimizar erros e evitar distorções nos resultados. Ainda segundo o autor, a limpeza dos dados não se limita à remoção de registros problemáticos, mas envolve a aplicação de técnicas que reduzem vieses e garantem maior precisão nas decisões.
- *Load*: de acordo com a IBM (2024), os dados transformados são armazenados em um *Data Warehouse* ou *Data Lake*, a dependendo de seu formato. No ambiente corporativo, há um investimento significativo em *hardwares* e *softwares* que sejam dedicados ao armazenamento e gerenciamento de dados. Para isso, utilizam-se diversos Sistemas de Gerenciamento de dados (SGBDs), que podem operar com dados estruturados, por meio da linguagem *Structured Query Language* (SQL) ou com dados não estruturados e semiestruturados, por meio de soluções baseadas em *Not Only SQL* (NoSQL) (Oliveira, Melo, Santos *et al.*, 2018).
 - Banco de dados relacional: de acordo com Oracle (2024), o modelo relacional garante a integridade dos dados por meio de suas estruturas lógicas, armazenando as informações de maneira que elas se relacionam entre si. Nesse modelo, cada linha de uma tabela possui um identificador único, chamado de *primary key*. Para estabelecer relacionamentos com outras tabelas, utiliza-se a *foreign key*, que é usada para fazer consultas na linguagem SQL. As colunas de dados representam atributos, e cada registro contém apenas um valor por coluna. Esse modelo segue o princípio ACID (Atomicidade, Consistência, Isolamento e Durabilidade) que assegura a integridade e confiabilidade das transações.
 - Atomicidade: para que uma transação seja realizada banco de dados, é preciso que ela seja completada, não aceitando resultados parciais;
 - Consistência: garantem que os dados no estado que se encontra sem sofrer alterações;
 - Isolamento: a transação se mantém invisível para outros acessos até ser completada;

- Durabilidade: garantem que as alterações sejam permanentes após serem completadas.
- Banco de dados não relacionais: de acordo com IBM (2024), significa que estes bancos de dados aceitam formatos mais flexíveis sobre formatos que se encaixem melhor nos seus dados. Ao contrário do Banco de dados relacional, este possui uma maior escalabilidade verticalmente, não sendo necessário altos investimentos em *hardwares*, seguindo o modelo CAP (Consistência, Disponibilidade e Tolerância):
 - Consistência: cada consulta retorna o resultado mais recente ou um erro;
 - Disponibilidade: garante (prioriza) que cada resultado não retorne um erro;
 - Tolerância: alguma demora ou perdas entre os nós não interrompem a operação do sistema. No entanto, isso pode afetar o comportamento do sistema dependendo das prioridades estabelecidas, tendo sistemas que priorizam mais a consistência do que a disponibilidade ou o contrário.

2.8 Ciência de dados

De acordo com Rautenberg (2019), *Data Science* é o método pelo qual se utiliza *softwares* para transformar dados em informações, resultando tomada de decisões direcionadas a dados. Esta área está em expansão no mundo, impulsionada pela transformação digital e pelo avanço das tecnologias de automação. A quantidade massiva de dados catalogadas a cada dia torna necessário a existência de especialistas em manipulações e gestão de *insights* de grandes bases de dados (Chimidt e Ernee, 2020).

Ciência de dados está conectada com análise de *Big Data*, se referindo a gestão e execução de todo o ciclo dos dados, incluindo todo o processo de análise. No qual seu objetivo é resolver alguma necessidade do negócio, seguindo o fluxo da produção de conhecimento (Chang e Grady, 2019).

De acordo com Chang e Grady (2019), em adição aos 4Vs do *Big Data*, outros termos necessários para os cientistas de dados garantirem a acurácia de suas análises são:

- Veracidade: refere a assertividade dos dados, logo, é necessário que se realize a limpeza dos dados para evitar o *garbage-in, garbage-out*, que significa a entrada de dados errôneos, impactando a saída do processo.

- Validade: refere-se à adequação do seu conjunto de dados para o local em que ele será necessário. Pois, além de acurados, os dados necessitam ser pertinentes para as hipóteses que se deseja testar.
- Volatilidade: é a tendência de os dados mudam as suas estruturas com o passar do tempo. Logo, é fundamental compreender que essas mudanças ocorreram e como contorná-las.
- Visualização: É um passo importante para se aplicar a ciência de dados para permitir que humanos entendam e analisem os resultados de maneira mais simplificada. Podendo ser do tipo:
 - Visualização Exploratória: se refere às técnicas usadas para navegar no sistema para se compreender as distribuições dos dados.
 - Visualização avaliativa: refere-se à visualização que permite avaliar e entender a performance e assertividade de uma análise em particular ou método de *machine learning*;
 - Visualização Explicativa: diz respeito a uma apresentação explicativa, tornando dados complexos em uma maneira mais simples de se entender para tomar decisões.
- Valor: é a forma como se é medida o ganho que foi atingido, proveniente do resultado de um processo de *data science* implementado.

De acordo com Da Silva, Peres e Boscarioli (2017), o processo de *Knowledge Discovery in database (KDD)* se inicia com a organização das bases de dados da área em que se tem o interesse em analisar, passando por etapas de processamento/limpeza. Nesse processo, o conhecimento será a base para responder às perguntas levantadas anteriormente (Pauleen e Wang, 2017).

Em um ambiente com alta taxa de mudanças, é necessário tomar decisões rápidas e informadas, adotando métodos mais ágeis e flexíveis para coletar, analisar e apresentar os dados (Dos Santos, Gonzaga, Strauss *et al.*, 2023).

Surge a necessidade de realizar o *MD³*, se utilizam ferramentas de computação para extrair informações úteis de um amontoado de dados, a fim de criar visualizações e fornecer suporte à tomada de decisão (Chen e Bai, 2023).

2.8.1 Linguagem de programação

De acordo com McKinney (2022) e Faes (2018), desde o seu surgimento o Python se tornou uma das linguagens interpretadas bastante populares. É uma linguagem bem formatada que pode ser usada para resolução de problemas do mundo real, sendo formatada para ser fácil de se entender e usar.

Faes (2018), afirma que esta linguagem possui uma sintaxe simples, além de suportar inúmeras bibliotecas de terceiros o tornando uma opção mais fácil de ler e escrever seus *scripts* (códigos), do que outras linguagens, o que torna a manutenção da vida útil do código mais simples, quando não há necessidade de fazer por sua conta caso já o tenha.

2.8.2 Bibliotecas e API Python

Nesta seção, são expostas as bibliotecas e API em Python que foram utilizadas neste trabalho, suas aplicações e características.

- **Pandas:** permite trabalhar com dados estruturados de maneira intuitiva, utilizando objetos como o *DataFrame*, uma estrutura orientada a colunas, e o *Series*, um *array*(vetor) unidimensional nomeado. Ele facilita operações aritméticas, tratamento de dados faltantes e a execução de operações relacionais, semelhantes às vistas em SGBDs (McKinney, 2022).
- **Numpy:** de acordo com McKinney (2022), essa biblioteca fornece uma estruturação dos dados e algoritmos, sendo mais eficiente no trabalho com *arrays* multidimensionais para realizar operações matemáticas. Um *array* é uma estrutura projetada para trabalhar de maneira mais eficiente com grandes conjuntos de dados, além da possibilidade de ser criado a partir de dados com estruturas diferentes (Behrman, 2023).
- **Plotly.express:** É um sistema de alto nível de plotagem de gráficos, de forma simples, declarando o atributo a serem expresso a partir da coluna da base de dados e escolhendo o tipo de gráfico desejado, além de possuir uma plotagem flexível e interativa (Dabbas, 2021).
- **Yfinance:** API contém um vasto conjunto de dados do histórico de preços das ações e algumas informações financeiras de empresas. Este módulo está ganhando bastante popularidade na comunidade Python, embora não seja endossado oficialmente pelo Yahoo e adotado com frequência por muitos usuários, para obter principalmente preços de ações e criptomoedas (Mandal, 2023).

- Streamlit: de acordo com Richards (2021), é uma estrutura de aplicação web (*web application framework*) que pode ser usada para o compartilhamento de resultados analíticos, a construção de experiências interativas complexas e a ilustração de novos modelos de *machine learning*, além de facilitar o processo de *deploy* (implementação online). O *Streamlit* tem com o objetivo aumentar a interação com os dados, acelerar o desenvolvimento e facilitar o compartilhamento das informações.

2.8.3 Business Intelligence

Sharda, Delen e Turban (2019), o termo *Business Intelligence* combina arquitetura, base de dados e ferramentas analíticas, permitindo o acesso interativo aos dados, inclusive em tempo real. Isso facilita a manipulação de informações e oferece aos tomadores de decisão análises focadas, ajudando-os a tomar decisões mais embasadas e baseadas em dados

Então, os responsáveis por tomarem decisões podem utilizar *software* de *Business Intelligence* para que as leituras dos dados sejam feitas de maneira mais eficiente e fácil. Fazendo com que os tomadores de decisões possam as realizar com o maior acesso a informações úteis (De Souza, 2022).

3 METODOLOGIA

Neste tópico, será definida a caracterização da pesquisa, sua natureza e a forma como foi subdividida para nortear o trabalho e facilitar o acompanhamento.

3.1 Caracterização da pesquisa

Nesta seção, será abordada a caracterização da pesquisa, destacando sua natureza e os procedimentos adotados. A natureza da pesquisa será discutida, considerando sua classificação em relação ao objetivo, abordagem e método. Além disso, serão apresentados os procedimentos de pesquisa utilizados para coletar e analisar os dados, a fim de fornecer uma compreensão clara sobre o processo investigativo empregado no estudo.

3.1.1 Natureza de pesquisa

De acordo Gil (2017), o presente estudo se classifica como uma pesquisa aplicada, voltada para a aquisição de conhecimento a partir da análise de uma situação específica. O objetivo é avaliar estratégias de investimento para a construção de uma carteira de ativos. A pesquisa se concentra na análise de dados históricos, com foco em identificar quais estratégias de investimento têm se mostrado mais eficientes em termos de retorno e risco.

Quanto à abordagem, trata-se de uma pesquisa quantitativa, pois utiliza processamento eletrônico de dados organizados em tabelas, facilitando os cálculos estatísticos. O processo é organizado e sequenciado de forma lógica à medida que a pesquisa se desenvolve (Gil, 2002).

Quanto aos objetivos, de acordo com a definição de Gil (2017), a pesquisa pode ser classificada como descritiva, pois busca estudar as características de um grupo, sendo no presente trabalho estudar as particularidades de cada método e os comparar.

3.1.2 Procedimento de pesquisa

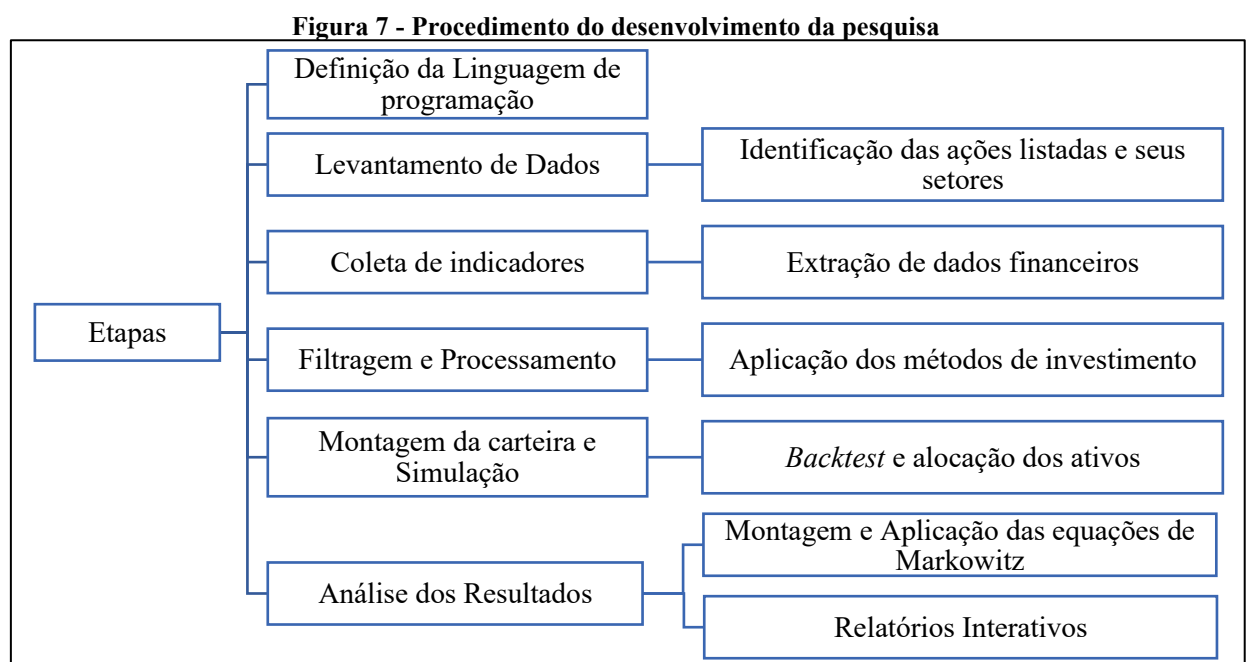
No que diz respeito ao procedimento de pesquisa, o foco adotado nesta pesquisa é a metodologia documental. Essa escolha é fundamentada pelo objetivo de a pesquisa demonstrar mais clareza, ser específico e a um cenário, tendo como objetivo o fornecimento de uma resposta definitiva para o problema proposto (Gil, 2017).

Pesquisas documentais realizadas na área das Ciências Econômicas valem-se de dados quantitativos, disponibilizados por meio de tabelas e banco de dados. Para a realização do processo analítico, é necessário o uso de procedimentos estatísticos, como correlação, testes de

hipóteses regressão entre outros (Gil, 2017). Sendo os dados coletados do próprio mercado financeiro através de uma API.

3.2 Etapas de Pesquisa

A presente pesquisa foi subdividida em 6 etapas, as quais estão estruturadas seguindo o modelo de Kleinert (2024), e adaptada para o estudo em questão: Definição da linguagem de programação, levantamento de dados, coleta de indicadores, Filtragem e Processamento, Montagem da Carteira e Simulação, Análise de Resultados, conforme Figura 7:



Fonte: Adaptado de Kleinert (2024)

A primeira etapa foi a definição da linguagem de programação a ser utilizada, sendo escolhido o Python para este estudo, do qual é uma das linguagens mais usadas e populares do mundo na atualidade, reconhecida por sua versatilidade, poder e facilidade de aprendizado. É amplamente empregado em diversas áreas, como desenvolvimento web, ciência e análise de dados e automação (Python, 2025), que são os focos desta pesquisa.

Ao escolher a linguagem, é necessário obter os *Tickers* de cada empresa, que são os códigos usados para identificá-las na bolsa de valores. Para obter essas informações, foi utilizada a biblioteca *Investpy* (Pypi, 2025), um pacote que realiza *Web scrapping*, ou seja, a extração automatizada de dados de sites. No caso, o *Investpy* coleta informações do site

investing.com, permitindo acessar os códigos das empresas listadas na bolsa de valores brasileira

A seguir, são aplicados na API do Yahoo Finance, chamada yfinance, permitindo o acesso a diversos indicadores financeiros, como setores, histórico de preços, dividendos, valor patrimonial (foi utilizado o valor contábil), quantidade de ações, lucro, dívida bruta, caixa, receita, liquidez corrente e EBIT.

Com os indicadores foi possível iniciar a quarta etapa, de acordo com cada método foi filtrado quais ações respeitavam a estratégia e teria a capacidade de compor uma carteira de investimento. Após a filtragem, as empresas que se mostraram capazes perante as estratégias de investimentos montou-se uma carteira de dez ativos.

Foi realizado o *backtest*, no qual, mensalmente, 7,5% do salário-mínimo vigente em cada ano foi investido nos 10 ativos selecionados, ao longo do período de 2000 a 2024. O percentual adotado justifica-se por corresponder à contribuição mínima obrigatória ao INSS na atualidade, permitindo uma comparação entre investimento no mercado de capitais e modelo previdenciário tradicional.

Adicionalmente, conforme evidenciado na Figura 6 Assaf Neto (2021), a ampliação do número de ativos na carteira para além de próximo de 10 ativos não reduz de maneira significativa o risco, pois ao adicionar ativos se reduz o risco não sistemático, que corresponde aos riscos específicos de cada empresa. Em contrapartida, o risco sistemático, relacionado às flutuações inerentes ao mercado como um todo, permanece presente independentemente da quantidade de ativos incluídos na carteira. Dessa forma, a ampliação excessiva da carteira torna-se contraproducente para uma análise fundamentalista dos ativos, uma vez que dificulta o acompanhamento e análise de cada empresa adquirida, sem trazer benefícios em termos de mitigação de risco

A seleção dos ativos seguiu critérios metodológicos específicos, considerando que, dependendo da abordagem adotada, a compra de determinada ação somente ocorreria caso já estivesse listada na bolsa de valores e seu preço inferior ao preço-teto de compra estipulado pelo método. Para garantir um critério adequada de comparação temporal, esse preço foi ajustado utilizando a inflação média do período, assegurando sua equivalência ao longo do tempo

Com a carteira idealizada e os ativos comprados foi calculado o retorno esperado e risco do portfólio para cada carteira segundo o método de Harry Markowitz, que se utiliza do retorno de cada ativo utilizando a biblioteca *Numpy* para realizar cálculos matemáticos de desvio padrão, correlação e logaritmo, com os resultados alcançado foi montado um relatório utilizando o

streamlit e plotly.express. O *Streamlit* será o responsável por desenvolver uma aplicação web para poder acessar as visualizações dos dados de maneira estruturada aplicando os conceitos de *business intelligence* para direcionar a melhor tomada de decisão.

4 RESULTADOS E DISCUSSÃO

Nesta seção do trabalho, são discutidos os resultados da pesquisa, com o objetivo de identificar o método que apresentou o melhor desempenho ao longo do tempo e aquele que proporcionou a melhor relação entre risco e retorno. O algoritmo descrito a seguir, estruturado de forma natural, descreve o processo adotado para a análise e a seleção das ações na construção da carteira de investimentos.

Algoritmo de Seleção de Ações:

1. Coleta e Preparação dos Dados a depender do método, segundo a seção 2.5.2 até a 2.5.3;
2. Cálculo dos Indicadores Fundamentais segundo as equações da seção 2.4;
3. Aplicação dos Métodos de Seleção de Ações;
4. Ordenação e Priorização das Ações a depender do método, nos dados alcançados no passo 1;
5. Correção pela Inflação e Simulação de Compras;
6. Avaliação da Performance da Carteira.

4.1 Coleta de dados, aplicação dos métodos e montagem da carteira

Nesta seção, é detalhado a aplicação dos métodos utilizados para a montagem de cada carteira, bem como o processo de coleta, limpeza e estruturação dos dados das empresas.

4.1.1 Método Barsi

Para a aplicação do método Barsi, foram catalogadas 390 *tickers*, de empresas na bolsa de valores. O primeiro filtro considerou apenas empresas de setores perenes, classificados como BESST. Como a API utilizada acessa um site estrangeiro, esses setores são denominados como Serviços Financeiro, Utilidades e Serviço de Telecomunicação. Após essa filtragem, o número de *tickers* foi reduzido para 83.

Para garantir a representatividade da amostra e minimizar distorções nos resultados, foram excluídas empresas cujo *dividend yield* médio excedesse a 15% nos últimos cinco anos, pois valores acima desse limite podem indicar distribuições anormais, venda de ativos ou descapitalização. Como o método adotado exige que os dividendos sejam sustentáveis a partir do fluxo de caixa operacional, a aplicação desse critério permitiu a remoção de *outliers*, resultando em uma amostra final de 73 *tickers*.

O próximo critério aplicado foi a exclusão de empresas com *dividend yield* médio inferior a 6% nos últimos cinco anos, uma vez que o método estabelece esse valor como pagamento mínimo para viabilizar o investimento. Para garantir uma informação mais precisa e conservadora do desempenho dos ativos, optou-se por utilizar a média dos indicadores, reduzindo a amostra final para 38 *tickers*.

Por fim, foram selecionadas empresas com preço por valor patrimonial ($\frac{Preço}{BVPS}$) médio dos últimos cinco anos inferiores a 1, reduzindo a amostra para 27 *tickers*. Optou-se por utilizar a média dos últimos cinco anos em razão a limitação da API empregada na coleta de dados, que disponibiliza apenas os valores patrimoniais referentes a esse período. Seguindo o método Barsi, priorizaram-se ativos com o $\frac{Preço}{BVPS}$ inferior a 1, indicando possível subvalorização pelo mercado, porém a ordenação realizada levou em consideração o maior índice de *dividend yield*. As dez empresas com os menores valores desse indicador foram escolhidas. Para evitar exposição excessiva, manteve-se apenas a primeira ocorrência de cada empresa com múltiplas ações listadas. Além disso, as ações CMIG4 E GGBR4 foram desconsideradas na montagem da carteira, pois, ao comparar os gráficos disponíveis no site base com os dados fornecidos pela API, foi identificada uma inconsistência nos valores no período de análise. Os resultados estão na Tabela 1.

Tabela 1 - Carteira aplicando método Barsi

Ação	Empresa
BMG4	Banco BMG
BAZA3	Banco da Amazônia
PINE3	Banco Pine
BRSR6	Banco do Estado do Rio Grande do Sul
BMIN4	Banco Mercantil de Investimentos
CSMG3	Companhia de Saneamento de Minas Gerais (COPASA)
TAEE3	Transmissora Aliança de Energia Elétrica (TAESA)
BBAS3	Banco do Brasil
ABCB4	Banco ABC Brasil
CLSC4	Centrais Elétricas de Santa Catarina

Fonte: Autoria Própria (2025)

4.1.2 Método Bazin

Para a aplicação do método Bazin, foram catalogados os mesmos 390 *tickers*, de empresas na bolsa de valores, como a estratégia de Bazin visa investir em empresas pouco

endividadas, o indicador DPNE não pode ser aplicada as empresas do setor financeiro, pois, o seu modelo de negócio gera uma obrigação de pagamento a cada novo dinheiro que entra na empresa, logo para a aplicação do método foi desconsiderado este setor, resultando em 292 *tickers*.

Para aprimorar a coleta dos dados, foram removidas as ações com preço e quantidade de ações negociadas zerados no site, o que pode indicar inconsistências ou deslistagem das empresas da bolsa. Isso resultou em 254 *tickers*.

É necessário realizar a limpeza de empresas com lucro negativo nos últimos 5 anos, para a aplicação do filtro de *payout* máximo, e com isso também reduziu os *tickers* para 196.

O método conduz os investimentos somente a empresas que destinam de até 50% no máximo dos seus lucros como dividendos para os seus acionistas, visando manter a possibilidade de crescimento e investimento, buscando a permanência e competitividade no mercado. Ocorrendo assim uma redução para 147 *tickers*.

Por fim, a estratégia visa investir em empresas dos quais possuam uma DPNE menor que 1,75. Com isso, se obteve um total de 117 *tickers*.

Para remover os *outliers*, assim como no método anterior, que também considera o pagamento de dividendos, foram selecionadas as empresas com *dividend yield* superior a 6% e inferior ou igual a 15%, resultando em 39 *tickers*. A carteira foi então montada ordenando as empresas do maior para o menor *dividend yield*. Caso uma mesma empresa tenha mais de uma ação listada, apenas a primeira ocorrência foi mantida para evitar exposição excessiva. Além disso, as ações CMIG4 E GGBR4 foram desconsideradas na montagem da carteira, pois, ao comparar os gráficos disponíveis no site base com os dados fornecidos pela API, foi identificada uma inconsistência nos valores no período de análise. O resultado da carteira está apresentado na Tabela 2.

Tabela 2 - Carteira aplicando método Bazin

Ação	Empresa
ROMI3	Indústrias ROMI
GOAU3	Metalúrgica Gerdau
WHRL4	Whirlpool
UNIP6	Unipar Carbocloro
TASA4	Taurus Armas
CMIG3	Companhia Energética de Minas Gerais (CEMIG)
TAEE3	Transmissora Aliança de Energia Elétrica (TAESA)
CLSC4	Centrais Elétricas de Santa Catarina (CELESC)
CGRA3	Grazziotini
CPLE6	Companhia Paranaense de Energia (COPEL)

Fonte: Autoria Própria (2025)

4.1.3 Método Graham

Para aplicar o método Graham, foram considerados os mesmos 390 *tickers*, de empresas na bolsa de valores. A estratégia de Graham busca investir em empresas descontadas, ou seja, cujas ações estão sendo negociadas abaixo de seu valor intrínseco. Como primeiro filtro, foram selecionadas apenas empresas estruturadas, com receita mínima de R\$ 300 milhões, reduzindo o número de *tickers* para 233.

A estratégia, exige a verificação de boa situação financeira da empresa, esta situação é adquirida acompanhando a liquidez corrente ($\frac{\text{Ativo Circulante}}{\text{Passivo Circulante}}$), o resultado sendo 1, a empresa é aceita no método. O que resultou na redução para 167 ações.

O próximo filtro considera a consistência dos lucros, incluindo apenas empresas que não registraram prejuízo nos últimos três anos, reduzindo o total para 127 ações. Além disso, as ações CMIG4 e GGBR4 foram excluídas da carteira devido a inconsistência nos dados entre a API utilizada e os registros oficiais da bolsa de valores durante o período de análise.

Por fim, aplica-se a equação IV, selecionando as 10 ações com a menor razão entre o preço médio e o valor intrínseco, ou seja, as mais descontadas. Essas ações estão apresentadas na Tabela 3. Caso uma mesma empresa possua diferentes tipos de ações listadas, apenas a primeira ocorrência é mantida, descartando as demais para evitar exposição excessiva

Tabela 3 - Carteira aplicando método Graham

Ação	Empresa
STBP4	Santos Brasil Participações
KLBN4	Klabin
WHRL3	Whirlpool
REDE3	Rede Energia Participações (ENERGISA)
BEEF3	Minerva Foods
CEAB3	CEA Modas
TECN3	Grupo Tecnos
RADL3	Raia Drogasil
VAMO3	Vamos Locação de Caminhões. Máquinas e Equip
POMO4	Marcopolo

Fonte: Autoria Própria (2025)

4.1.4 Método Greenblatt

Para este método, o primeiro filtro foi eliminar empresas do setor financeiro, pois o indicador “dívida líquida” presente no EV não pode ser aplicado a empresas deste setor, já que o “vir a ser” destas empresas necessitam trabalhar alavancadas para executar suas funções. E com isso, reduziu o número de ações de 390 para 292.

Com essa seleção em mãos, foram excluídas as empresas que apresentaram lucro médio negativo nos últimos três anos, reduzindo o total para 255 ações. Como o método envolve divisões de indicadores, como EV e ROE, valores negativos poderiam distorcer os resultados. Por isso foram removidas as empresas com patrimônio líquido (valor contábil multiplicado pelo número de ações) e EBIT negativos, resultando em 178 *tickers*.

Para eliminar os *outliers*, todas as empresas com ROE superior a 20 foram descartadas, pois valores acima desse limite podem indicar eventos atípicos. O método visa investir em empresas com desempenho estável e constante. Essa filtragem resultou na redução para 90 ações.

Por fim, é aplicada a equação da fórmula mágica de Greenblatt, e as 10 ações que possuírem a menor soma dos 2 indicadores serão as selecionadas, sendo as mais descontadas e com mais perspectiva de ganho segundo o método, caso a mesma empresa seja selecionada com ações diferentes a primeira ocorrência será mantida e a próxima será desconsiderada para evitar uma extra exposição, é mostrada na Tabela 4 a seleção.

Tabela 4 - Carteira aplicando método Greenblatt

Ação	Empresa
SAPR4	Companhia de Saneamento do Paraná (SANEPAR)
TAE4	Transmissora Aliança de Energia Elétrica (TAESA)
EMAE4	Empresa Metropolitana Águas Energia (EMAE)
CEDO4	Companhia de Fiação Tecidos Cedro Cachoeira
PTNT4	Pettenati Industria Textil
AHEB3	São Paulo Turismo
POSI3	Positivo
GEPA4	Rio Paranaapanema Energia
DOHL4	Dohler
PETR4	Petrobras

Fonte: Autoria Própria (2025).

4.2 Simulação das carteiras

Nesta seção, é apresentado como foi definido o valor disponível para investimento, o valor investido e como foi corrigido certos valores no tempo.

4.2.1 Definição da quantia disponível para investimento

Como a pesquisa, tem foco na aposentadoria, investigou a possibilidade de complementar por meio de investimentos em ações. Para simular esse cenário, foi estabelecida uma contribuição de 7,5%, equivalente à contribuição mínima exigida para o INSS na atualidade, que garante a aposentadoria dos brasileiros. Essa porcentagem foi aplicada sobre o salário-mínimo mensal. O período de teste abrangeu de janeiro de 2000 a dezembro de 2024, utilizando os valores do salário-mínimo de cada ano, conforme apresentado na Tabela 5.

Tabela 5 - Salários-mínimos utilizados na pesquisa

Ano	Salário (R\$)	Ano	Salário (R\$)	Ano	Salário (R\$)
2000	151,00	2009	465,00	2018	954,00
2001	180,00	2010	510,00	2019	998,00
2002	200,00	2011	540,00	2020	1.039,00
2003	240,00	2012	622,00	2021	1.100,00
2004	260,00	2013	678,00	2022	1.212,00
2005	300,00	2014	724,00	2023	1.302,
2006	350,00	2015	788,00	2024	1.412,00
2007	380,00	2016	880,00		
2008	415,00	2017	937,00		

Fonte: Mix vale (2024)

4.2.2 Correção dos valores

Os métodos Barsi, Bazin e Graham só realizam a compra de ativos quando é definido um preço teto ou valor intrínseco. Se o preço de negociação estiver abaixo desse valor, a compra será efetuada. Portanto, foram realizadas compras de ações com preços compatíveis para cada período analisado, garantindo que os preços teto e o valor intrínseco correspondessem aos valores de cada época, com a devida correção. As aquisições foram realizadas na quantidade de uma ação por vez, até que o saldo disponível fosse totalmente utilizado. Para que a compra ocorresse, era necessário que o preço de negociação da ação fosse inferior ao seu preço teto ou valor intrínseco, conforme definido pelo método aplicado

De acordo com Banco Central (2024), a inflação média nos últimos 25 anos foi de R\$100,00 no ano 2000 serem proporcionais a R\$447,39 hoje, de acordo com Reis (2019), é possível utilizar a equação do CAGR (*Compound Annual Growth Rate*) para convergir a inflação acumulada nos últimos 25 anos na inflação média aplicando a equação 16:

$$Taxa_{média} = \left(\frac{Valor_{final}}{Valor_{inicial}} \right)^{\frac{1}{tempo}} \quad (16)$$

Fonte: Reis (2019)

Totalizando assim, uma taxa anual média de 6,18%. Para corrigir os ativos será aplicada a equação 17. Ressalta-se que a taxa média de inflação foi utilizada devido à impossibilidade de acessar as taxas de inflação mensal ou anual ao longo de todo o período.

$$Preço_{corrigido} = \frac{Preço}{(1+inflação_{média})^{tempo}} \quad (17)$$

Fonte: Assaf Neto (2022)

Sendo assim, utilizando esta taxa de desvalorização para o preço teto do método Barsi, obteve-se os valores apresentados na Figura 8. Ressalta-se que foi necessário realizar a correção dos preços tetos ao longo do tempo, devido à limitação dos dados financeiros disponíveis, que abrangem apenas os últimos cinco anos. Como a análise considera um período de 25 anos, tornou-se necessário projetar os preços tetos para os anos anteriores com base nessa taxa de desvalorização, a fim de garantir a coerência histórica dos critérios de compra utilizados

Figura 8 - Preço teto corrigidos por ano de teste para o método Barsi

Ticker	Preço_Teto	2000	2001	2002	2003	2004	2005	2006	2007	...	2015	2016	2017	2018	2019	2020	2021	2022	2023
CSMG3.SA	25.281498	5.65	5.99	6.37	6.76	7.18	7.62	8.09	8.59	...	13.88	14.74	15.65	16.62	17.64	18.73	19.89	21.12	22.42
TAE3.SA	17.481064	3.90	4.15	4.40	4.67	4.96	5.27	5.59	5.94	...	9.60	10.19	10.82	11.49	12.20	12.95	13.75	14.60	15.51
BMGB4.SA	4.654881	1.04	1.10	1.17	1.24	1.32	1.40	1.49	1.58	...	2.56	2.71	2.88	3.06	3.25	3.45	3.66	3.89	4.13
PINE4.SA	4.234167	0.95	1.00	1.07	1.13	1.20	1.28	1.36	1.44	...	2.32	2.47	2.62	2.78	2.95	3.14	3.33	3.54	3.76
BMIN4.SA	23.860069	5.33	5.66	6.01	6.38	6.77	7.19	7.64	8.11	...	13.10	13.91	14.77	15.68	16.65	17.68	18.77	19.93	21.16
BAZA3.SA	73.751606	16.47	17.49	18.57	19.72	20.94	22.23	23.60	25.06	...	40.49	42.99	45.65	48.47	51.47	54.65	58.02	61.61	65.42
BRSR6.SA	14.548931	3.25	3.45	3.66	3.89	4.13	4.39	4.66	4.94	...	7.99	8.48	9.01	9.56	10.15	10.78	11.45	12.15	12.90
BBAS3.SA	24.496683	5.47	5.81	6.17	6.55	6.95	7.38	7.84	8.32	...	13.45	14.28	15.16	16.10	17.09	18.15	19.27	20.46	21.73
ABC84.SA	20.881658	4.66	4.95	5.26	5.58	5.93	6.29	6.68	7.10	...	11.46	12.17	12.92	13.72	14.57	15.47	16.43	17.44	18.52
CLSC4.SA	65.709417	14.67	15.58	16.54	17.57	18.65	19.80	21.03	22.33	...	36.07	38.30	40.67	43.18	45.85	48.69	51.70	54.89	58.28

Fonte: Autoria Própria (2025)

Na primeira coluna temos os nomes das ações, na coluna o preço teto calculado para o ano de 2024, e as demais colunas para cada ano.

Agora para o preço teto do método Graham os valores são apresentados na Figura 9.

Figura 9 - Preço teto corrigidos por ano de teste para o método Graham

Ticker	Valor_Intrinseco	2000	2001	2002	2003	2004	2005	2006	2007	...	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
STBP3.SA	56.34	12.58	13.36	14.19	15.06	15.99	16.98	18.03	19.14	...	30.93	32.84	34.87	37.03	39.32	41.74	44.32	47.06	49.97	53.06
KLB4.SA	18.40	4.11	4.36	4.63	4.92	5.22	5.55	5.89	6.25	...	10.10	10.73	11.39	12.09	12.84	13.63	14.48	15.37	16.32	17.33
WHRL3.SA	20.05	4.48	4.75	5.05	5.36	5.69	6.04	6.42	6.81	...	11.01	11.69	12.41	13.18	13.99	14.86	15.77	16.75	17.78	18.88
REDE3.SA	22.38	5.00	5.31	5.63	5.98	6.35	6.75	7.16	7.60	...	12.29	13.05	13.85	14.71	15.62	16.58	17.61	18.70	19.85	21.08
BEEF3.SA	39.18	8.75	9.29	9.86	10.47	11.12	11.81	12.54	13.31	...	21.51	22.84	24.25	25.75	27.34	29.03	30.82	32.73	34.75	36.90
CEAB3.SA	23.26	5.19	5.52	5.86	6.22	6.60	7.01	7.44	7.90	...	12.77	13.56	14.40	15.29	16.23	17.23	18.30	19.43	20.63	21.91
VAMO3.SA	32.79	7.32	7.78	8.26	8.77	9.31	9.88	10.49	11.14	...	18.00	19.11	20.30	21.55	22.88	24.30	25.80	27.39	29.08	30.88
POMO4.SA	14.35	3.20	3.40	3.61	3.84	4.07	4.33	4.59	4.88	...	7.88	8.37	8.88	9.43	10.01	10.63	11.29	11.99	12.73	13.51
RADL3.SA	77.15	17.23	18.29	19.42	20.63	21.90	23.25	24.69	26.22	...	42.36	44.97	47.75	50.70	53.84	57.16	60.70	64.45	68.43	72.66
TECN3.SA	9.60	2.14	2.28	2.42	2.57	2.73	2.89	3.07	3.26	...	5.27	5.60	5.94	6.31	6.70	7.11	7.55	8.02	8.52	9.04

Fonte: Autoria Própria (2025)

Na primeira coluna, temos os nomes das ações, na coluna o preço teto calculado para o ano de 2024, e as demais colunas para cada ano.

Agora para o preço teto do método de Bazin, os valores apresentados na Figura 10.

Figura 10 - Preço teto corrigidos por ano de teste para o método Bazin

ticker	Preco_teto	2000	2001	2002	2003	2004	2005	2006	2007	...	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
GOAU3.SA	16.500000	3.68	3.91	4.15	4.41	4.68	4.97	5.28	5.61	...	9.06	9.62	10.21	10.84	11.51	12.23	12.98	13.78	14.64	15.54
WHRL4.SA	9.867478	2.20	2.34	2.48	2.64	2.80	2.97	3.16	3.35	...	5.42	5.75	6.11	6.48	6.89	7.31	7.76	8.24	8.75	9.29
TAE3.SA	18.841792	4.21	4.47	4.74	5.04	5.35	5.68	6.03	6.40	...	10.34	10.98	11.66	12.38	13.15	13.96	14.82	15.74	16.71	17.75
CMIG3.SA	15.827761	3.53	3.75	3.99	4.23	4.49	4.77	5.07	5.38	...	8.69	9.23	9.80	10.40	11.04	11.73	12.45	13.22	14.04	14.91
GGBR3.SA	21.974206	4.91	5.21	5.53	5.87	6.24	6.62	7.03	7.47	...	12.06	12.81	13.60	14.44	15.33	16.28	17.29	18.36	19.49	20.70
CGRA3.SA	36.513019	8.15	8.66	9.19	9.76	10.36	11.01	11.69	12.41	...	20.05	21.28	22.60	24.00	25.48	27.05	28.73	30.50	32.39	34.39
CPLE6.SA	9.855283	2.20	2.34	2.48	2.63	2.80	2.97	3.15	3.35	...	5.41	5.74	6.10	6.48	6.88	7.30	7.75	8.23	8.74	9.28
UNIP6.SA	77.600442	17.33	18.40	19.54	20.75	22.03	23.39	24.83	26.37	...	42.60	45.24	48.03	51.00	54.15	57.50	61.05	64.82	68.83	73.08
PAT4.SA	48.545556	10.84	11.51	12.22	12.98	13.78	14.63	15.54	16.50	...	26.65	28.30	30.05	31.90	33.88	35.97	38.19	40.55	43.06	45.72
ROMI3.SA	15.327931	3.42	3.63	3.86	4.10	4.35	4.62	4.91	5.21	...	8.42	8.94	9.49	10.07	10.70	11.36	12.06	12.80	13.60	14.44

Fonte: Autoria Própria (2025)

Na primeira coluna temos os nomes das ações, na coluna o preço teto calculado para o ano de 2024, e as demais colunas para cada ano.

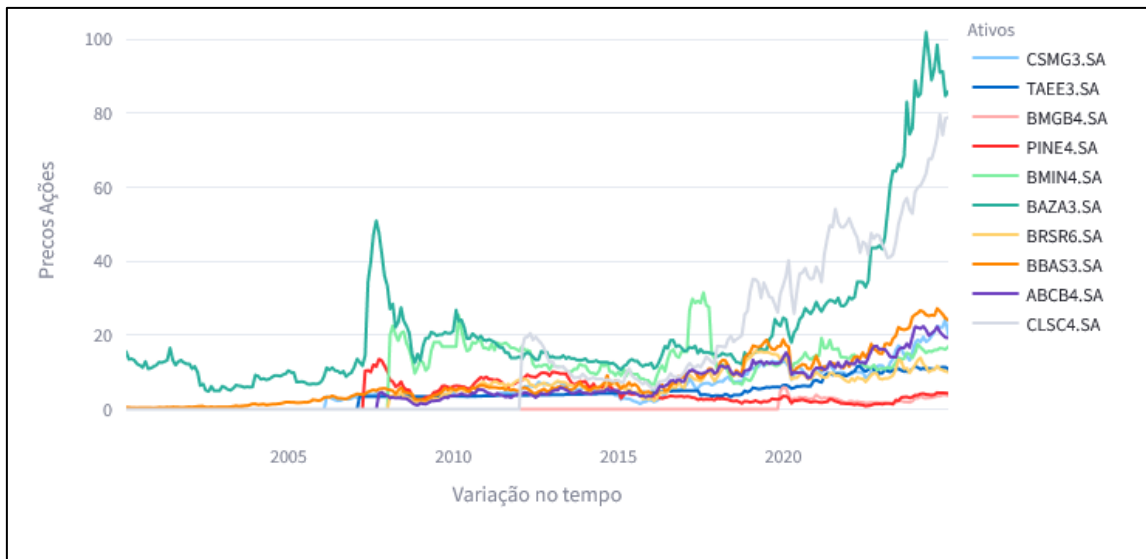
4.3 Simulação e alocação dos ativos

Nesta seção, são apresentadas as simulações de compra dos ativos, detalhando a alocação das posições e a performance de cada ativo ao longo do período analisado.

4.3.1 Simulação método Barsi

A compra foi feita de maneira automática, em que tendo recurso para comprar e o preço teto sendo maior que o preço efetivo, seria comprado uma unidade e assim por diante, até não se ter mais recursos para compra ou não possuir mais ações viáveis para se adquirir. No gráfico 1, é apresentado a performance dos ativos durante o período analisado. Se ação não estiver disponível em uma determinada data, o gráfico permanecerá estático até sua abertura de capital na bolsa.

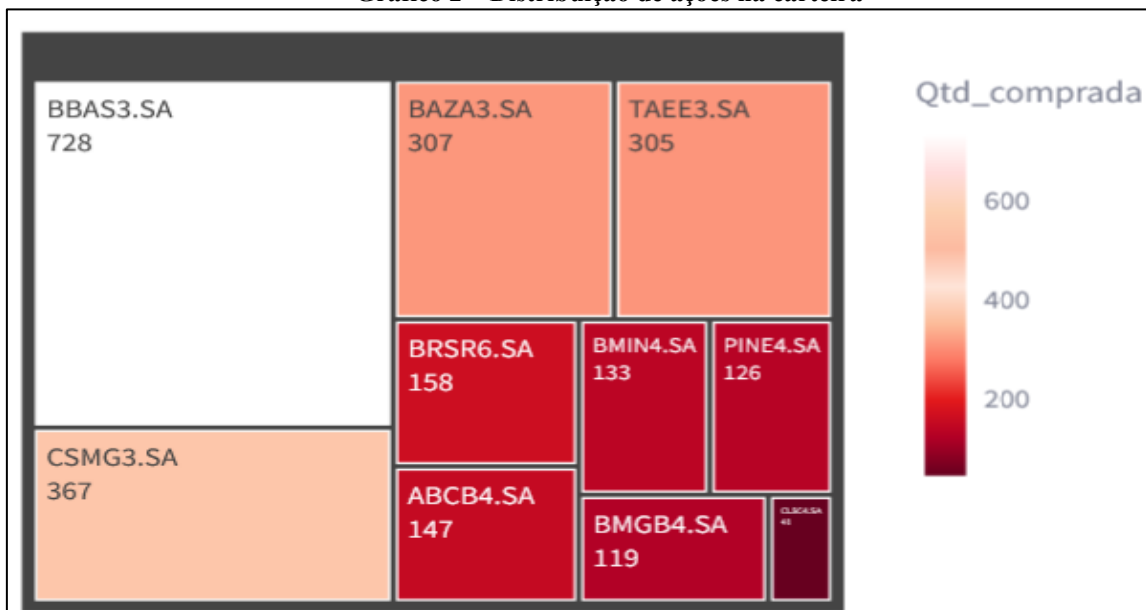
Gráfico 1 – Performance dos ativos selecionados para o método



Fonte: Autoria Própria (2025)

Apesar das variações ao longo dos anos, inclusive com períodos de quedas expressivas, observa-se que a tendência geral dos ativos foi de valorização no longo prazo. Alguns papéis apresentaram um comportamento mais cíclico, com oscilações mais frequentes, porém ainda assim evidenciaram uma trajetória ascendente ao serem analisados em horizontes mais amplos já a distribuição de quantidade de ações para cada ativo, no fim da simulação, será apresentada no gráfico 2.

Gráfico 2 – Distribuição de ações na carteira



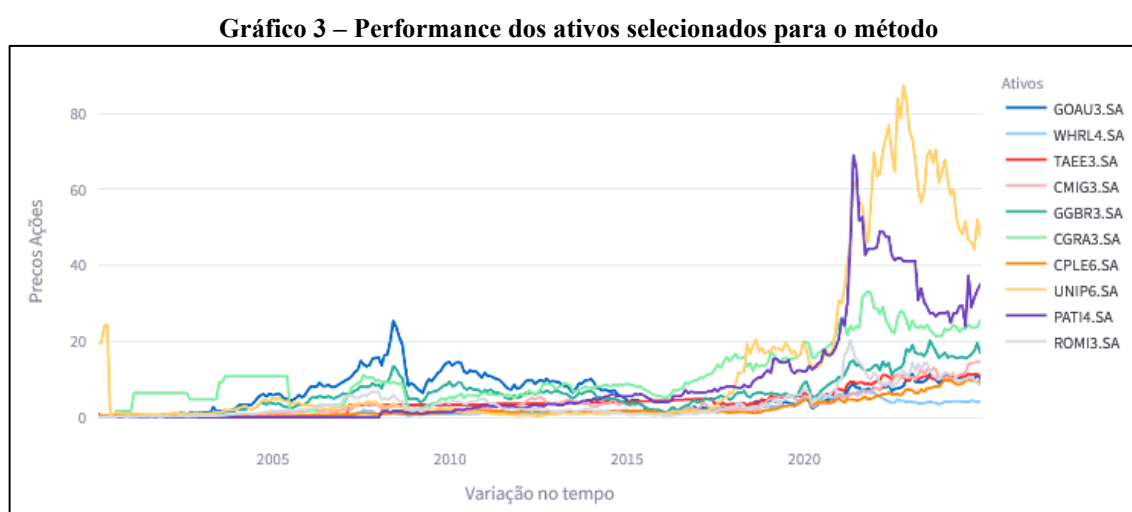
Fonte: Autoria Própria (2025)

As maiores posições em termos de quantidade de ações foram BBAS3, CSMG3, BAZA3 e TAEE3, sendo duas do setor financeiro e uma do setor de saneamento e um do setor

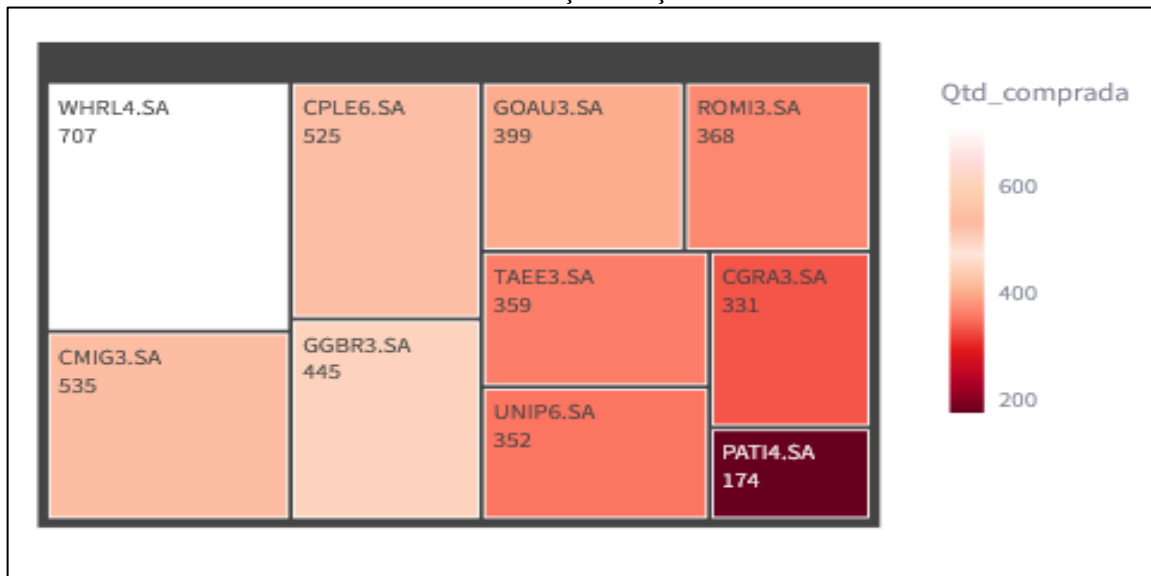
de energia. A posição com menor posição foi a CLSC4. Observa-se uma considerável exposição ao setor bancário, o que pode ser explicado pela lógica do método adotado: embora haja um filtro setorial, o critério de seleção é baseado principalmente no desconto em relação ao valor patrimonial e na remuneração ao acionista. Dessa forma, empresas do setor financeiro que apresentam bons fundamentos e estão subvalorizadas tendem a ser mais recorrentes na carteira resultante da aplicação desse modelo,

4.3.2 Simulação método Bazin

No gráfico 3, é apresentado a performance dos ativos durante o período analisado. Se ação não estiver disponível em uma determinada data, o gráfico permanecerá estático até sua abertura de capital na bolsa, também sendo realizado a compra de maneira automática.



Esse método apresentou maior assimetria, com picos acentuados nos preços das ações. Algumas delas se comportaram de maneira mais cíclica, mas ainda assim demonstraram uma tendência progressiva nos valores das empresas. Já a distribuição da quantidade de ações por ativo ao final da simulação será apresentada no Gráfico 4.

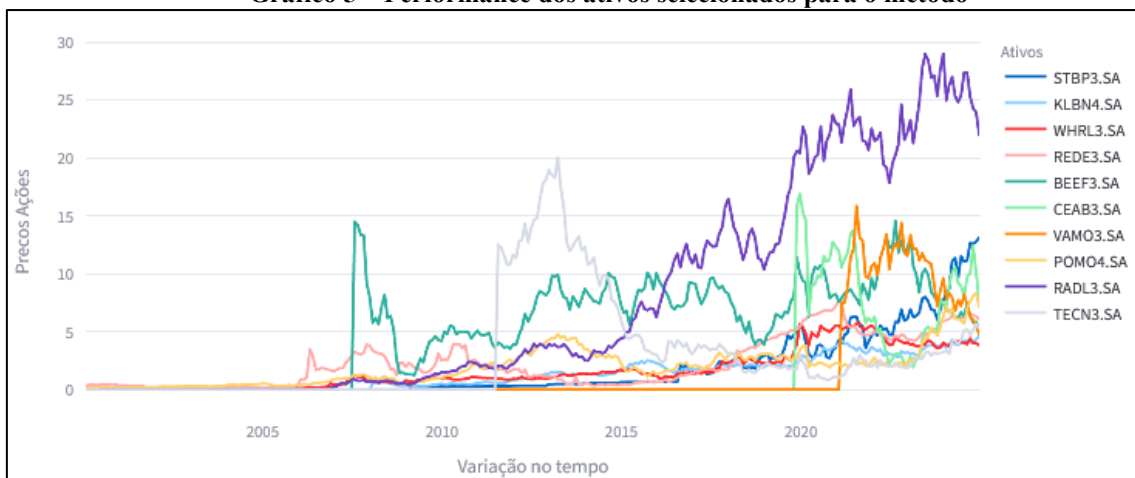
Gráfico 4 – Distribuição de ações na carteira

Fonte: Autoria Própria (2025)

As maiores posições em termos de quantidade de ações foram WHRL4 setores industriais. As empresas com menores aportes foi a PATI4, do setor de siderurgia.

4.3.3 Simulação método Graham

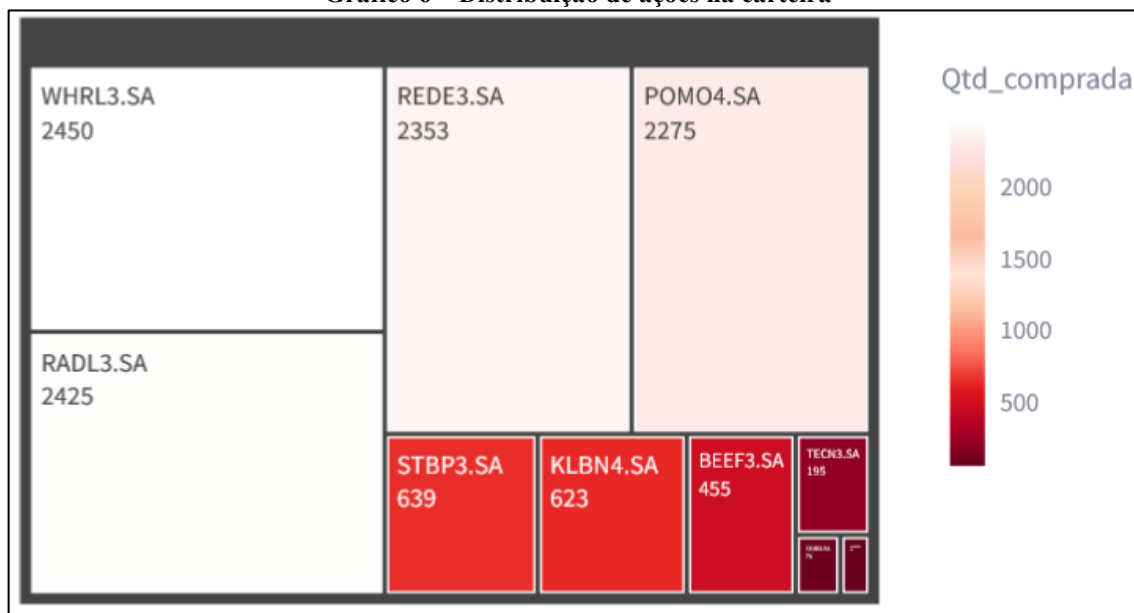
No gráfico 5, é apresentado a performance dos ativos durante o período analisado. Se ação não estiver disponível em uma determinada data, o gráfico permanecerá estático até sua abertura de capital na bolsa, também sendo realizado a compra de maneira automática.

Gráfico 5 – Performance dos ativos selecionados para o método

Fonte: Autoria Própria (2025)

Para esse método, ainda é possível notar assimetrias, porém pode ver uma tendência mais a valorização dos ativos, ou seja, a multiplicação de capital. A seguir, teve-se a distribuição de quantidade de ações para cada ativo, no fim da simulação, será apresentada no gráfico 6.

Gráfico 6 – Distribuição de ações na carteira

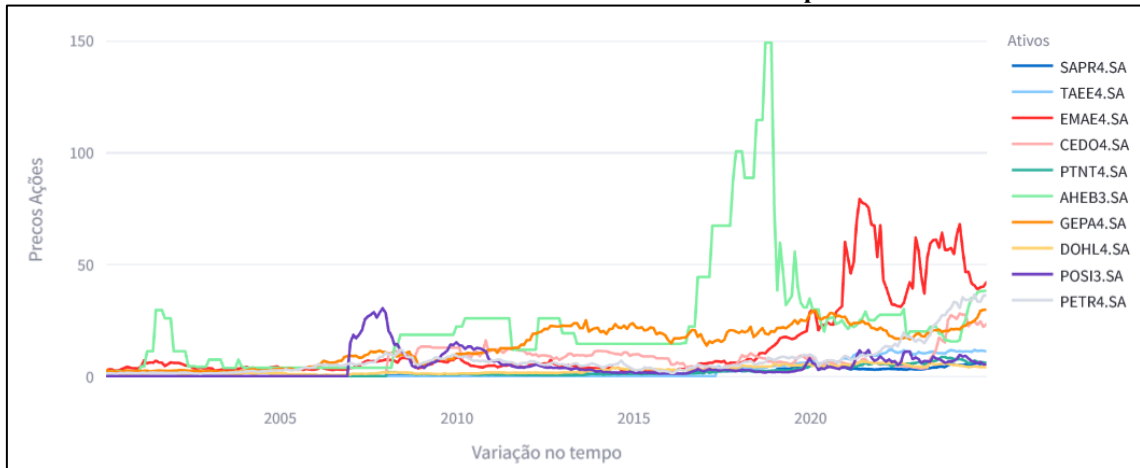


Fonte: Autoria Própria (2025)

As maiores posições em quantidade de ações foram WHRL4, RADL3, REDE3 e POMO4, representando, respectivamente, os setores industrial, de saúde, de energia e de consumo. As empresas com menor aporte foram CEAB3 e VAMO3, pertencentes aos setores de energia e transporte, respectivamente

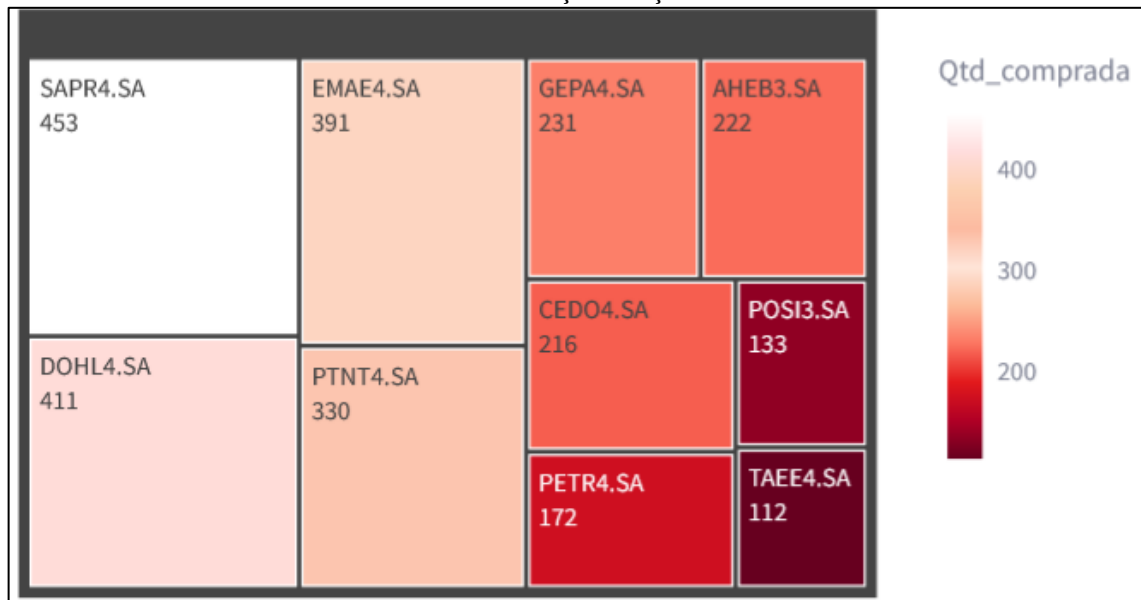
4.3.4 Simulação método Greenblatt

No Gráfico 5, é apresentado a performance dos ativos durante o período analisado se ação não estiver disponível em uma determinada data, o gráfico permanecerá estático até sua abertura de capital na bolsa, também, sendo realizado a compra de maneira automática.

Gráfico 7 – Performance dos ativos selecionados para o método

Fonte: Autoria Própria (2025)

Neste método, é possível observar assimetrias pontuais, no entanto, os ativos apresentam tendência de valorização ao longo do tempo. Ainda assim, nota-se que o comportamento dos ativos é mais cíclico. A seguir, será apresentada a distribuição da quantidade de ações por ativo ao final da simulação, conforme ilustrado no Gráfico 8.

Gráfico 8 – Distribuição de ações na carteira

Fonte: Autoria Própria (2025)

As maiores posições em quantidade de ações foram SAPR4, EMAE4 e DOHL4, representando dois do setor de saneamento e um do setor de energia. A empresa com menor aporte foi TAE4, pertencente ao setor elétrico.

4.4 Análise da simulação

Nesta seção é montado um relatório com visualizações para interpretar os dados alcançados com as aplicações das estratégias de investimento.

4.4.1 Cálculo da Selic média e seu desvio padrão

Para realizar o cálculo da verificação se os retornos dos investimentos valem o risco, é necessária uma taxa livre de risco para realizar a comparação. Primeiramente, foi coletado informações das Selic mensais nos últimos 25 anos alcançando a Figura 11.

Figura 11 - Selic mensais nos últimos 25 anos

Mês	2000	2001	2002	2003	2004	2005	2006	2007	2008	...	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Janeiro	0.0146	0.0127	0.0153	0.0197	0.0127	0.0138	0.0143	0.0108	0.0093	...	0.0094	0.0106	0.0109	0.0058	0.0054	0.0038	0.0015	0.0073	0.0112	0.0097
Fevereiro	0.0145	0.0102	0.0125	0.0183	0.0108	0.0122	0.0115	0.0087	0.0080	...	0.0082	0.0100	0.0087	0.0047	0.0049	0.0029	0.0013	0.0076	0.0092	0.0080
Março	0.0145	0.0126	0.0137	0.0178	0.0138	0.0153	0.0142	0.0105	0.0084	...	0.0104	0.0116	0.0105	0.0053	0.0047	0.0034	0.0020	0.0093	0.0117	0.0083
Abril	0.0130	0.0119	0.0148	0.0187	0.0118	0.0141	0.0108	0.0094	0.0090	...	0.0095	0.0106	0.0079	0.0052	0.0052	0.0028	0.0021	0.0083	0.0092	0.0089
Mai	0.0149	0.0134	0.0141	0.0197	0.0123	0.0150	0.0128	0.0103	0.0088	...	0.0099	0.0111	0.0093	0.0052	0.0054	0.0024	0.0027	0.0103	0.0112	0.0083
Junho	0.0139	0.0127	0.0133	0.0186	0.0123	0.0159	0.0118	0.0091	0.0096	...	0.0107	0.0116	0.0081	0.0052	0.0047	0.0021	0.0031	0.0102	0.0107	0.0079
Julho	0.0131	0.0150	0.0154	0.0208	0.0129	0.0151	0.0117	0.0097	0.0107	...	0.0118	0.0111	0.0080	0.0054	0.0057	0.0019	0.0036	0.0103	0.0107	0.0091
Agosto	0.0141	0.0160	0.0144	0.0177	0.0129	0.0166	0.0126	0.0099	0.0102	...	0.0111	0.0122	0.0080	0.0057	0.0050	0.0016	0.0043	0.0117	0.0114	0.0087
Setembro	0.0122	0.0132	0.0138	0.0168	0.0125	0.0150	0.0106	0.0080	0.0110	...	0.0111	0.0111	0.0064	0.0047	0.0046	0.0016	0.0044	0.0107	0.0097	0.0084
Outubro	0.0129	0.0153	0.0165	0.0164	0.0121	0.0141	0.0109	0.0093	0.0118	...	0.0111	0.0105	0.0064	0.0054	0.0048	0.0016	0.0049	0.0102	0.0100	0.0093
Novembro	0.0122	0.0139	0.0154	0.0134	0.0125	0.0138	0.0102	0.0084	0.0102	...	0.0106	0.0104	0.0057	0.0049	0.0038	0.0015	0.0059	0.0102	0.0092	0.0079
Dezembro	0.0120	0.0139	0.0174	0.0137	0.0148	0.0147	0.0099	0.0084	0.0112	...	0.0116	0.0112	0.0054	0.0049	0.0037	0.0016	0.0077	0.0112	0.0089	0.0093

Fonte: Receita Federal (2025)

Para calcular a taxa acumulada se utilizou da equação (18), definindo a taxa SELIC, acumulada anualmente, depois a sua média nos últimos 25 anos e o desvio padrão que ela sofreu.

$$Taxa_{acumulada} = [\prod_{i=1}^{12} 1 + r_i] - 1 \quad (18)$$

Fonte: Assaf Neto (2022)

Inicialmente foi somado um a todos os valores da tabela alcançando a figura 12:

Figura 12 - Taxa mensais da Selic acrescido do valor um

Mês	2000	2001	2002	2003	2004	2005	2006	2007	2008	...	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Janeiro	1.0146	1.0127	1.0153	1.0197	1.0127	1.0138	1.0143	1.0108	1.0093	...	1.0094	1.0106	1.0109	1.0058	1.0054	1.0038	1.0015	1.0073	1.0112	1.0097
Fevereiro	1.0145	1.0102	1.0125	1.0183	1.0108	1.0122	1.0115	1.0087	1.0080	...	1.0082	1.0100	1.0087	1.0047	1.0049	1.0029	1.0013	1.0076	1.0092	1.0080
Março	1.0145	1.0126	1.0137	1.0178	1.0138	1.0153	1.0142	1.0105	1.0084	...	1.0104	1.0116	1.0105	1.0053	1.0047	1.0034	1.0020	1.0093	1.0117	1.0083
Abril	1.0130	1.0119	1.0148	1.0187	1.0118	1.0141	1.0108	1.0094	1.0090	...	1.0095	1.0106	1.0079	1.0052	1.0052	1.0028	1.0021	1.0083	1.0092	1.0089
Mai	1.0149	1.0134	1.0141	1.0197	1.0123	1.0150	1.0128	1.0103	1.0088	...	1.0099	1.0111	1.0093	1.0052	1.0054	1.0024	1.0027	1.0103	1.0112	1.0083
Junho	1.0139	1.0127	1.0133	1.0186	1.0123	1.0159	1.0118	1.0091	1.0096	...	1.0107	1.0116	1.0081	1.0052	1.0047	1.0021	1.0031	1.0102	1.0107	1.0079
Julho	1.0131	1.0150	1.0154	1.0208	1.0129	1.0151	1.0117	1.0097	1.0107	...	1.0118	1.0111	1.0080	1.0054	1.0057	1.0019	1.0036	1.0103	1.0107	1.0091
Agosto	1.0141	1.0160	1.0144	1.0177	1.0129	1.0166	1.0126	1.0099	1.0102	...	1.0111	1.0122	1.0080	1.0057	1.0050	1.0016	1.0043	1.0117	1.0114	1.0087
Setembro	1.0122	1.0132	1.0138	1.0168	1.0125	1.0150	1.0106	1.0080	1.0110	...	1.0111	1.0111	1.0064	1.0047	1.0046	1.0016	1.0044	1.0107	1.0097	1.0084
Outubro	1.0129	1.0153	1.0165	1.0164	1.0121	1.0141	1.0109	1.0093	1.0118	...	1.0111	1.0105	1.0064	1.0054	1.0048	1.0016	1.0049	1.0102	1.0100	1.0093
Novembro	1.0122	1.0139	1.0154	1.0134	1.0125	1.0138	1.0102	1.0084	1.0102	...	1.0106	1.0104	1.0057	1.0049	1.0038	1.0015	1.0059	1.0102	1.0092	1.0079
Dezembro	1.0120	1.0139	1.0174	1.0137	1.0148	1.0147	1.0099	1.0084	1.0112	...	1.0116	1.0112	1.0054	1.0049	1.0037	1.0016	1.0077	1.0112	1.0089	1.0093

Fonte: Autoria Própria (2025)

Em sequência, foi realizado o produto da sequência de todas as colunas e subtraindo o valor um, assim como na equação, alcançando a Tabela 6 para cada ano:

Tabela 6 - Taxa Selic acumulada por ano

Ano	SELIC (%)	Ano	SELIC (%)	Ano	SELIC (%)
2000	17,45	2009	9,92	2018	6,42
2001	17,31	2010	9,78	2019	5,95
2002	19,16	2011	11,62	2020	2,75
2003	23,33	2012	8,48	2021	4,44
2004	16,24	2013	8,21	2022	12,38
2005	19,04	2014	10,91	2023	13,03
2006	15,08	2015	13,29	2024	10,89
2007	11,85	2016	14,03		
2008	12,48	2017	9,96		

Fonte: Autoria Própria (2025)

Com esses dados, é calculado o desvio padrão transformando os valores acumulado em um *array* de 1 dimensão, para realizar este cálculo, foi utilizado a biblioteca NumPy, devido a eficiência da biblioteca em realizar operações vetorizadas. Alcançando o resultado de 4,79% de desvio padrão e 12,16% de retorno médio ao ano.

4.4.2 Cálculo do retorno do portfólio

Será aplicada a equação do retorno logarítmico para o cálculo do retorno anualizado, com uso da biblioteca NumPy. Do qual, inicialmente, é realizado o logaritmo da divisão do $preço_t$ pelo $preço_{t-1}$ e somar essas variações dos retornos, como é demonstrado na equação (9). Com esse resultado, é aplicado a exponencial e subtraído um do valor para ter o retorno em porcentagem, levando em consideração o tempo que cada empresa ficou aberta na bolsa. Para o cálculo do retorno anual médio é aplicado a equação do CAGR.

A seguir é calculado a porcentagem de cada ativo na carteira, de acordo com a equação (19):

$$w_{ativoi} = \frac{preco_{ativoi} * Qtd_{ativoi}}{\sum_{i=1}^{10} preco_i * Qtd_{ativoi}} \quad (19)$$

Com os dados de percentual de alocação e retorno de cada ativo, foram obtidos os resultados referentes a cada posição no método Barsi, conforme apresentados na Tabela 7. Ao total foi disponibilizado para investir R\$ 14.973,12 durante todo o período.

Tabela 7 - Retorno e porcentagem de cada ativo para o método Barsi

Ação	Retorno anual (%)	Porcentagem Carteira (%)
CSMG3	9,97	11,38
TAEE3	6,74	05,07
BMGB4	-5,67	0,66
PINE3	-4,85	0,82
BMIN4	-1,02	3,42
BAZA3	7,08	40,19
BRSR6	6,22	2,41
BBAS3	17,12	26,80
ABCB4	9,11	4,32
CLSC4	12,91	4,93

Fonte: Autoria Própria (2025)

Este método investiu efetivamente um valor de **R\$ 14.970,49**, sobrando uma quantia de **R\$ 2,63** para investir. Para o método Bazin será apresentado na Tabela 8:

Tabela 8 - Retorno e porcentagem de cada ativo para o método Bazin

Ação	Retorno anual (%)	Porcentagem Carteira (%)
ROMI3	17,50	4,99
GOAU3	10,09	5,40
WHRL4	11,21	4,38
UNIP6	4,05	36,73
TASA4	4,76	1,65
CMIG3	14,56	10,60
TAEE3	6,82	5,79
CLSC4	12,91	10,96
CGRA3	11,69	12,66
CPL6	12,60	6,84

Fonte: Autoria Própria (2025)

Este método investiu efetivamente um valor de **R\$ 14.968,99**, sobrando uma quantia de **R\$ 4,13** para investir.

Para o método Graham será apresentado na Tabela 9:

Tabela 9 – Retorno e porcentagem de cada ativo para o método Graham

Ação	Retorno anual (%)	Porcentagem Carteira (%)
STBP3	28,17	7,34
KLBN4	13,21	2,57
WHRL3	13,66	8,38
REDE3	11,96	12,74
BEEF3	-5,81	1,92
CEAB3	-12,99	0,52
TECN3	11,78	4,49
VAMO3	-10,81	0,22
ROMO4	16,52	14,91
RADL3	40,22	46,91

Fonte: Autoria Própria (2025)

Este método investiu efetivamente um valor de **R\$ 14.972,70**, sobrando uma quantia de **R\$ 0,42** para investir.

Para o método Greenblatt será apresentado na Tabela 10:

Tabela 10 - Retorno e porcentagem de cada ativo para o método Greenblatt

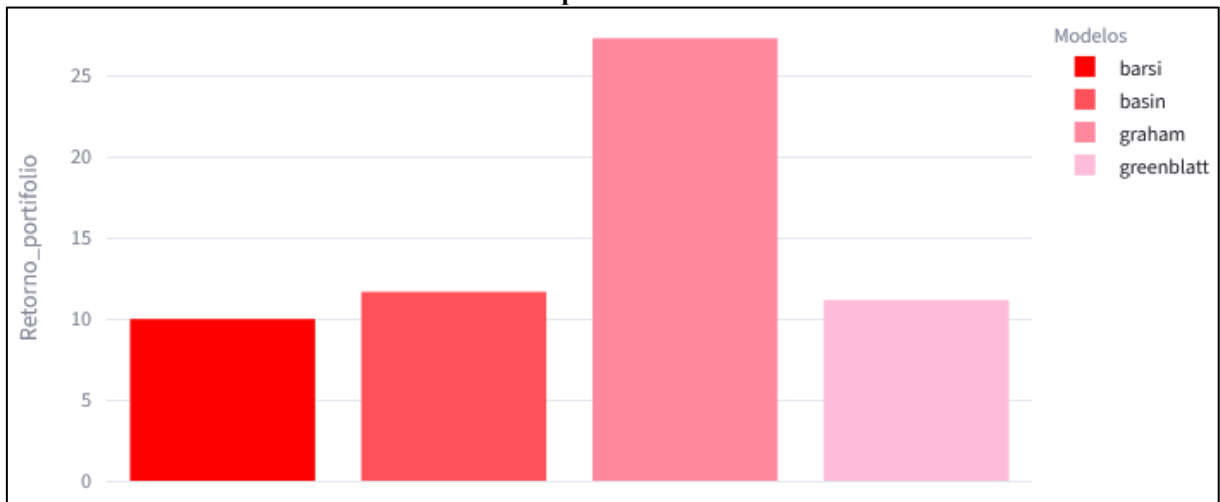
Ação	Retorno anual (%)	Porcentagem Carteira (%)
SAPR4	17,60	4,84
TAE4	12,78	2,43
EMAE4	12,13	32,14
CEDO4	4,92	9,85
PTNT4	7,28	3,88
AHEB3	12,06	16,62
POSI3	-5,70	1,35
GEPA4	10,28	13,45
DOHL4	6,37	3,27
PETR4	2,00	12,18

Fonte: Autoria Própria (2025)

Com este método, foi investido um valor de **R\$ 14.972,64**, sobrando uma quantia de **R\$ 0,48** para investir.

Com os resultados individuais apresentadas foi calculado o retorno esperado para cada método, apresentado no Gráfico 9 a seguir:

Gráfico 9 – Retorno esperado dos métodos analisados



Fonte: Autoria Própria (2025)

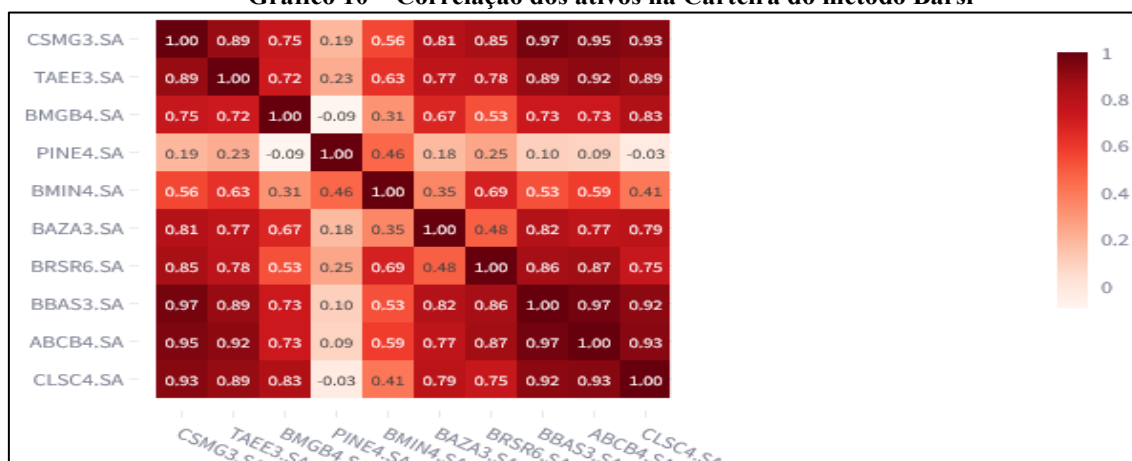
O método que teve o maior retorno esperado foi o Graham alcançando o valor de 27% aproximadamente, menor foi o Barsi com 9,97%, em seguida foi Bazin com 11,66% e Greenblatt 11,14%.

4.4.3 Cálculo do risco do portfólio

No modelo desenvolvido por Markowitz, o risco é representado pelo desvio-padrão dos retornos dos ativos, bem como pelas correlações entre eles. Correlações próximas de 1 indicam uma relação positiva, ou seja, quando um ativo sobre o outro também tende a subir, embora não necessariamente na mesma intensidade.

Para o cálculo do risco do portfólio, foram obtidas as correlações entre os ativos de cada método, utilizando a biblioteca Pandas para análise dos *dataframes*. No caso do método Barsi, a correlação entre os ativos está apresentada no Gráfico 10.

Gráfico 10 – Correlação dos ativos na Carteira do método Barsi

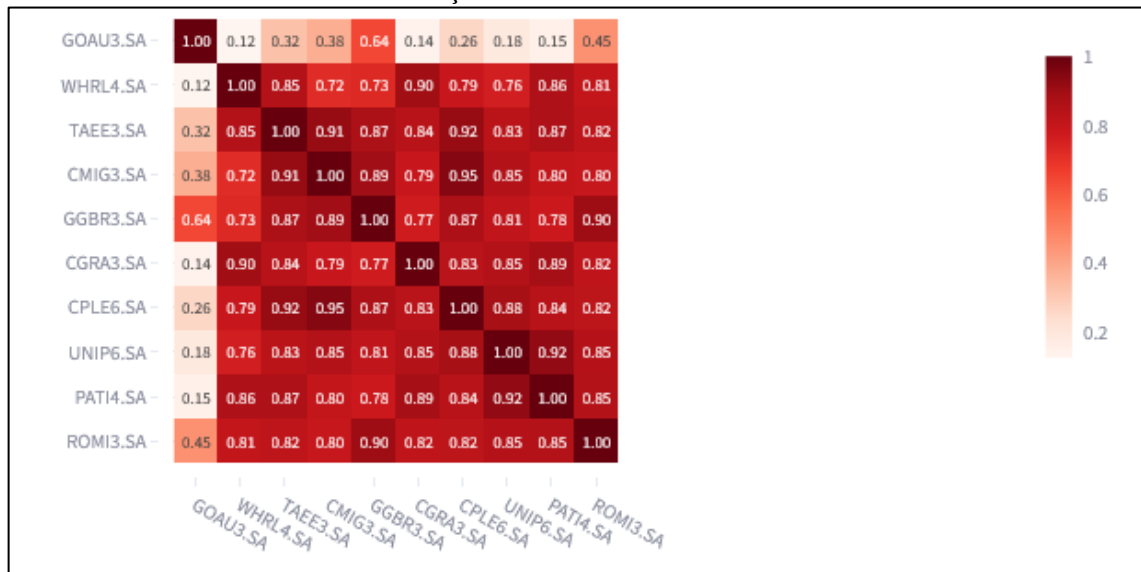


Fonte: Autoria Própria (2025)

O gráfico de correlação apresenta o grau de relação entre os ativos, comparando cada ativo representado na linha com aquele localizado na coluna. A coloração das células indica a intensidade da correlação: quanto mais escuro o tom de vermelho, mais próxima de 1 é a correlação, sugerindo uma forte relação positiva entre os ativos. Isso significa que ambos tendem a apresentar variações semelhantes, embora não necessariamente na mesma magnitude. Em contrapartida, cores mais próximas do vermelho claro indicam correlação fraca ou inexistente, sinalizando que os ativos apresentam movimentos que não se relacionam entre si.

É possível ver, que como o método foca em certos tipos de setores, existem correlações fortes e médias em algumas empresas, exemplo sendo quando se cruza a ação ABCB4 na linha e a SAPR4 na coluna se alcança uma correlação de 0,95, o que representa uma relação forte entre esses dois ativos, enquanto PINE4 e BMGB4 possui correlação de 0, não apresentando correlação entre si. Na Gráfico 11, será apresentada a correlações do método Bazin.

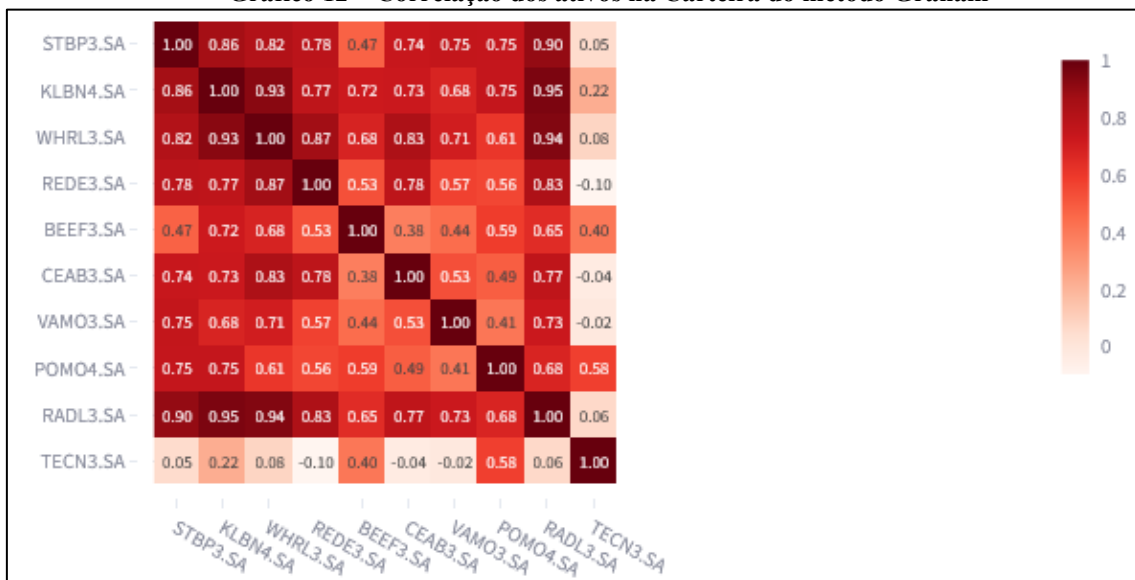
Gráfico 11 – Correlação dos ativos na Carteira do método Bazin



Fonte: Autoria Própria (2025)

Neste método, existe também uma correlação forte em alguns ativos da carteira, porém, em maior quantidade. Na Gráfico 12, será apresentado a correlação dos ativos para o método Graham.

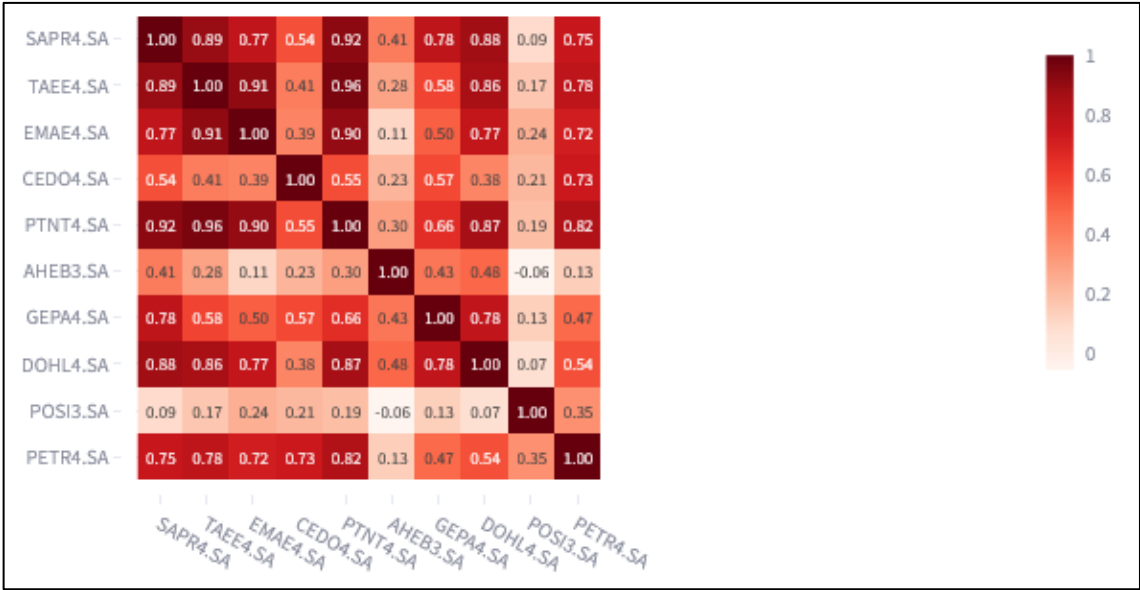
Gráfico 12 – Correlação dos ativos na Carteira do método Graham



Fonte: Autoria Própria (2025)

Já neste método, observa-se uma forte correlação entre alguns ativos, no entanto há variações significativas entre os demais, incluindo correlações negativas e próximas de zero. Na Gráfico 13, é apresentada a correlação dos ativos para o método de Greenblatt.

Gráfico 13 – Correlação dos ativos na Carteira do método Greenblatt

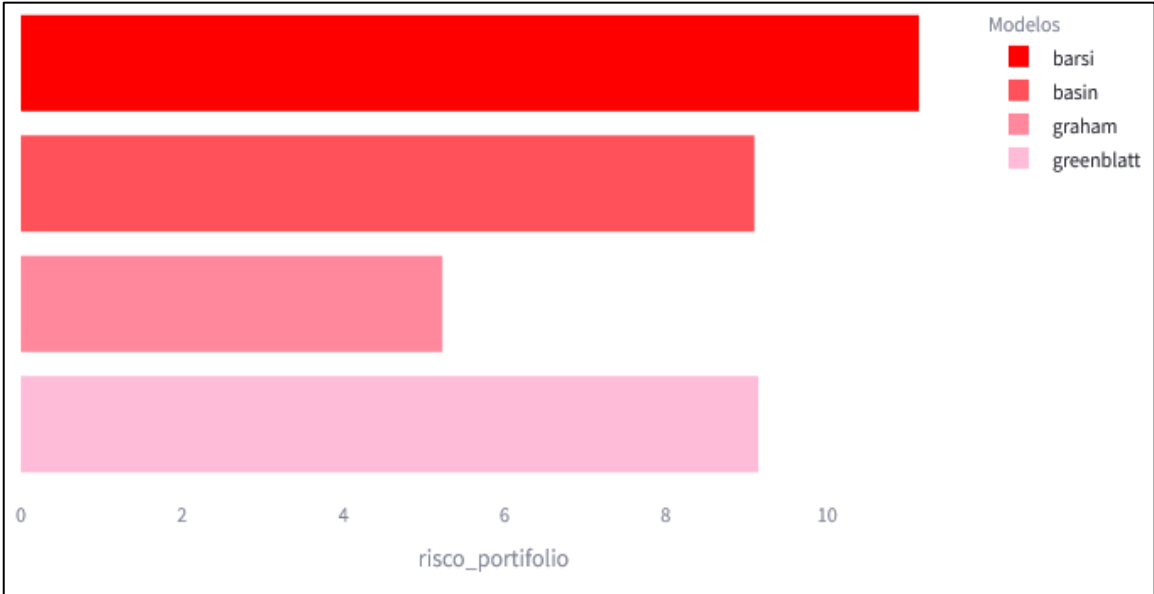


Fonte: Autoria Própria (2025)

Já neste método, é possível notar que se reduziu a correlação dos ativos, tendo mais valores intermediários e com correlações mais baixas do que nos outros métodos.

Com as correlações, porcentagem de cada ativo na carteira e risco individuais foi calculado o risco do portfólio para os métodos apresentados no Gráfico 14.

Gráfico 14 - Risco do portfólio para os métodos analisados



Fonte: Autoria Própria (2025)

O risco de um portfólio é determinado pela combinação da alocação percentual em cada ativo, da correlação entre esses ativos e do desvio padrão dos seus preços. Essa relação é

expressa na equação (8) do risco do portfólio, que mede a variação esperada dos retornos da carteira. Em outras palavras, quanto maior a volatilidade dos ativos, a correlação entre eles e a concentração nos papéis, maior será o risco do portfólio.

Seguindo a hipótese de que todo investidor tende a buscar o maior retorno com o menor risco admissível (Assaf Neto, 2021), um portfólio será considerado mais eficiente quando apresentar um desvio padrão menor, indicando menor volatilidade, e um retorno esperado maior. Dessa forma, a diversificação estratégica dos ativos, reduzindo a correlação entre eles e equilibrando a exposição, pode contribuir para a construção de um portfólio mais atrativo e equilibrado.

Então o método que produz o investimento com o maior risco é o Barsi com um valor de 11,14%, tendo Bazin e Greenblatt praticamente com riscos similares aproximadamente 9,09% e o método Graham com o menor risco sendo o valor de 5,22%.

4.4.4 Índice de viabilidade do retorno em relação ao risco para cada método de investimento e indicadores de rentabilidade

Nesta seção, é aplicada a equação do IS, em que se utilizará os valores de retorno de cada método, uma taxa livre de risco, que no caso será a SELIC, que mesmo ela apresentando uma alta variação nos anos, todo investimento realizado nela não ocasiona perda de capital, aplicando se obtém a Tabela 11, mostrando o índice para cada método.

Tabela 11 - IS para cada método de investimento

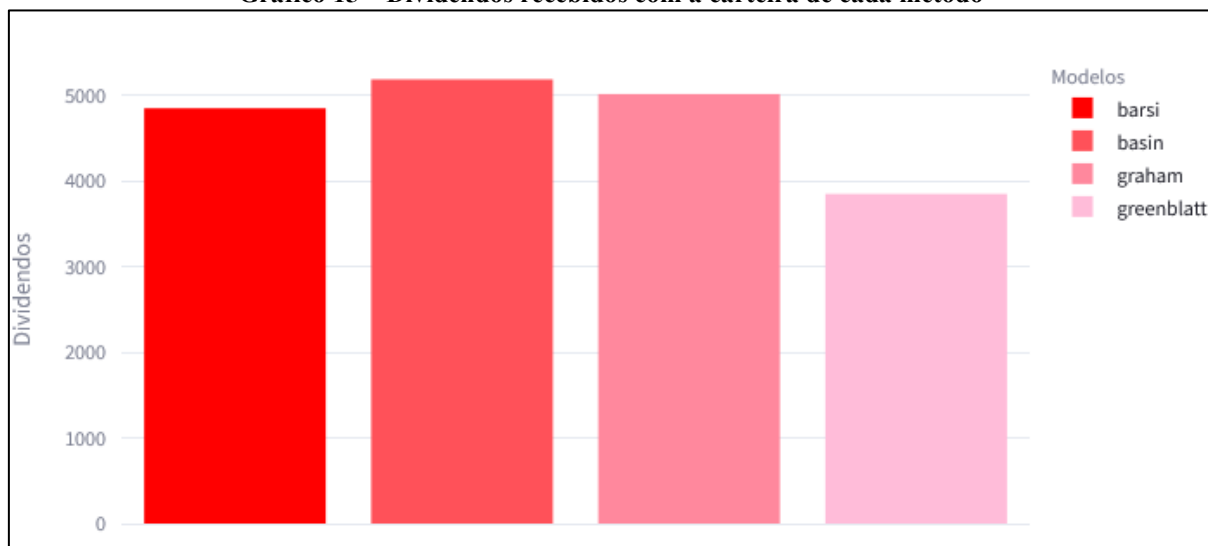
Método de investimentos	Índice Sharpe alcançado
Barsi	-0,20
Bazin	-0,06
Graham	2,90
Greenblatt	-0,13

Fonte: Autoria Própria (2025)

IS maior que 1, significa que o investimento é atrativo ao se comparar com a taxa livre de risco, pois, para cada unidade de risco extra se obtém uma unidade de retorno, para valores maiores que zero e menor que 1, o investimento não compensa o risco extra corrido. E para valores menores que 1, o retorno é menor que a renda fixa o que não compensa correr risco para ganhar menor retorno que o livre de risco. Logo, em relação ao IS somente o método Graham possui retorno considerável.

Em relação ao último ano (2024), foi catalogado o valor que cada carteira geraria de dividendos para o seu proprietário, caso tivessem essas posições formadas neste ano, apresentados no Gráfico 15.

Gráfico 15 – Dividendos recebidos com a carteira de cada método

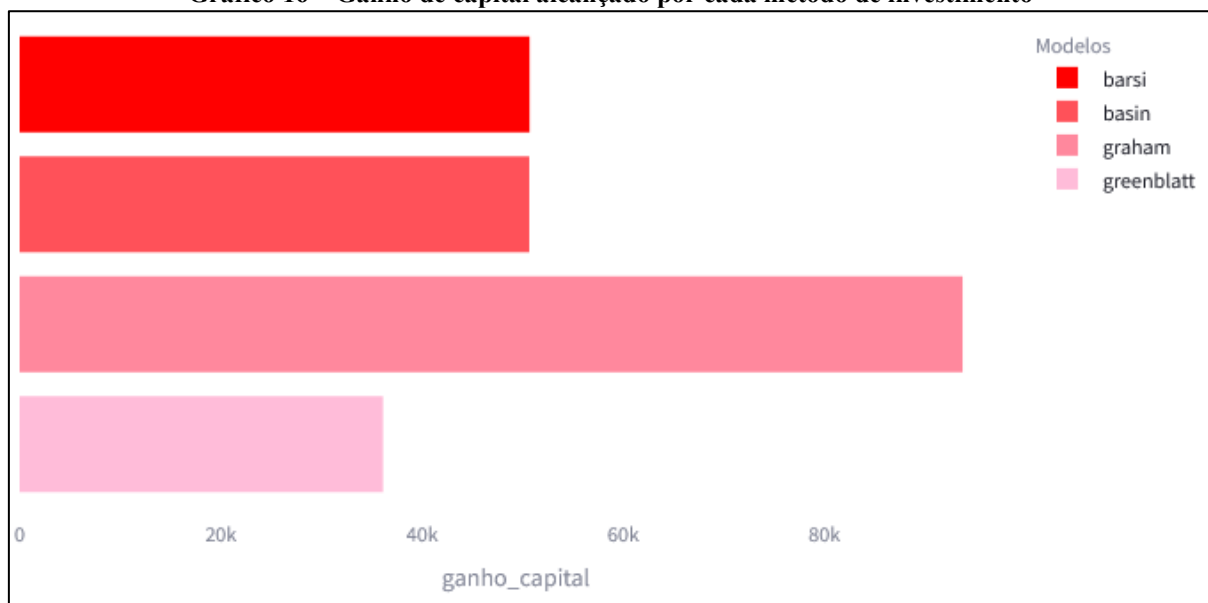


Fonte: Autoria Própria (2025)

O método que gerou a maior distribuição de dividendos foi Bazin, totalizando R\$ 5.183,00, no último ano. No entanto, devido à origem estrangeira da API utilizada, a única forma de remuneração direta aos acionista considerada na análise foi o pagamento de dividendos. No Brasil, existe um incentivo fiscal para o pagamento de juros sobre capital próprio (JSCP), uma vez que essa modalidade de distribuição é contabilizada como passivo, reduzindo a base tributável da empresa e, consequentemente, o montante de impostos devidos. Essas características permitem que as empresas possam distribuir quantias superiores aos acionistas.

Dessa forma, a ausência da consideração do JSCP pode subestimar o real retorno obtido pelos investidores que seguem abordagens voltadas à geração de dividendos no mercado brasileiro. Para fins de análise do ganho de capital de cada método foi elaborado o Gráfico 16.

Gráfico 16 – Ganho de capital alcançado por cada método de investimento



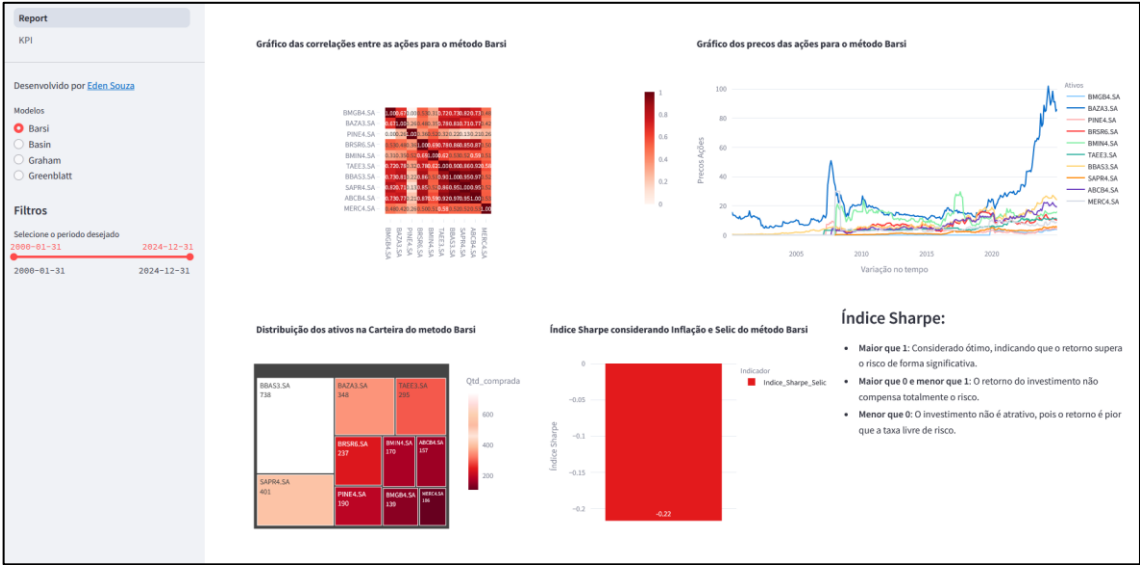
Fonte: Autoria Própria (2025)

O método que apresentou o maior ganho de capital foi o Graham, resultando em um montante de R\$ 93.729,69. Em comparação, caso o valor investido tivesse sido alocado em dívida pública por meio da SELIC, o retorno esperado seria de R\$ 27.158,32. Esse valor é inferior ao obtido por qualquer um dos métodos testados, evidenciando a existência de um custo de oportunidade associado à escolha do investimento.

Ressalta-se que, por limitações na disponibilidade dos dados, não foi possível considerar os dividendos juntamente com os aportes mensais ao longo de todo o período analisado. A API utilizada na obtenção das informações históricas disponibiliza os proventos apenas dos últimos três anos, o que inviabilizou a inclusão consistente dos dividendos na simulação de longo prazo. Dessa forma, os resultados apresentados referem-se exclusivamente ao ganho de capital obtido com a valorização das ações, não refletindo, portanto, o retorno total que seria alcançado com o reinvestimento ou o acúmulo dos proventos distribuídos.

A análise de dados referentes aos métodos de investimento foi realizada por meio do processo de ETL, utilizando a linguagem Python. Os dados foram posteriormente implementados e disponibilizados para visualização por meio de uma aplicação em ambiente local (*localhost*), conforme ilustrado nas figuras 13 e 14. Essa aplicação pode ser migrada para uma plataforma web, integrando-se ao próprio site do *framework* utilizado.

Figura 13 - Relatório para acompanhar performance dos métodos

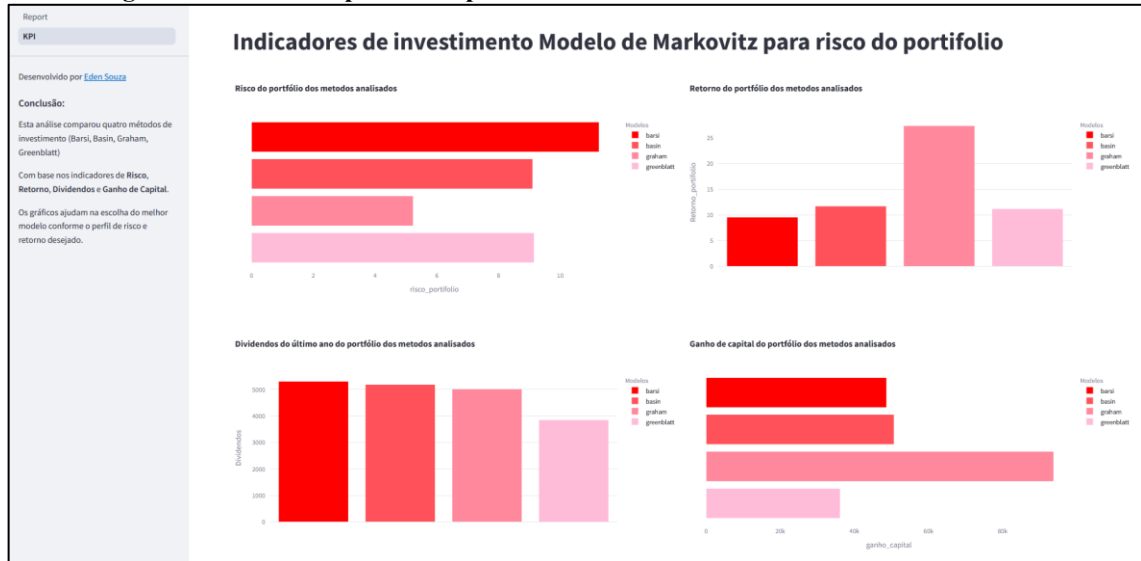


Fonte: Autoria Própria (2025)

Além dos indicadores que apresentam o histórico e o estado de cada método, a ferramenta dispõe de funcionalidades para filtragem, permitindo uma análise mais precisa. Por meio de um botão de seleção única, é possível escolher um método específico, enquanto um controle deslizante (*slider*) possibilita a filtragem de períodos específicos dos históricos das ações. Adicionalmente, a interface interativa permite a seleção de um período específico ao clicar e arrastar, além da opção de filtrar um único ativo em qualquer uma das visualizações disponíveis, proporcionando maior flexibilidade na análise dos dados.

A visualização apresentada tem como objetivo auxiliar o investidor na identificação dos ativos selecionados por cada método, bem como na análise da correlação entre esses ativos e da distribuição dos aportes ao longo do tempo. Além disso, é introduzido o IS, que avalia se o retorno adicional obtido pelo método justifica o risco assumido em comparação a um ativo livre de risco. Vale ressaltar que o IS não indica qual ativo apresenta maior retorno esperado em relação ao livre de risco, mas sim o retorno excedente compensa a volatilidade do investimento.

Figura 14 - Relatório para acompanhar indicadores de cada método de investimento



Fonte: Autoria Própria (2025)

Tendo as mesmas características que a outra aba, mostrando indicadores e formatos diferentes, comparando cada método em questão de risco, retorno, ganho de capital e dividendos, estas visualizações têm como objetivo exemplificar como seria a análise dos indicadores ao se aplicar cada método, se tratando de uma análise fundamentalista os métodos devem ser reaplicados no momento dos investimentos como acompanhamento e melhora do método específico, devido ao seu caráter dinâmico.

4.5 Análise de resultados

De acordo com as equações utilizadas, o método que gerou o maior retorno foi o de Graham, além de apresentar o menor risco, contrariando a premissa amplamente difundida de que, quanto maior o risco, maior o retorno. Por outro lado, o método de Barsi apresentou o menor retorno e risco, sendo, no entanto, o mais indicado para a geração de fluxo de caixa via dividendos.

O maior fluxo de caixa proveniente de dividendos foi gerado pelo método de Bazin, seguido de perto pelo de Graham. Isso era esperado, uma vez que esses métodos priorizam a remuneração por proventos em vez da valorização do capital. Vale destacar que, ao utilizar a API Yfinance, identificou-se que a plataforma considera como remuneração ao acionista apenas os pagamentos de dividendos, o que é comum em mercados internacionais. No entanto, no Brasil, existe também o pagamento de juros sobre capital próprio (JSCP), uma forma de remuneração que envolve a retenção de 15% de imposto de renda na fonte. Esse modelo é

vantajoso para as empresas, pois reduz o montante de imposto devido. Por isso, diversas companhias, especialmente dos setores BESST, optam por essa modalidade de pagamento.

Durante o *backtest*, os quatro métodos analisados apresentaram retornos consideráveis sobre o capital investido. Ainda que apenas o método de Graham tenha sido classificado como eficiente pelo IS, todas as estratégias superaram o retorno da taxa SELIC no período analisado. Isso demonstra que a exposição ao mercado de ações pode ser mais vantajosa, no longo prazo, do que adotar uma postura conservadora de rentista.

Com base na análise realizada, verificou-se que a aplicação das estratégias possibilita a geração de uma renda mensal entre R\$ 320,57 e R\$ 431,92, sem a necessidade de venda dos ativos. Esse montante equivale a cerca de 22,70% a 30,59% do salário-mínimo vigente em 2024. Considerando a possibilidade de manter os aportes pelos próximos dez anos – tempo necessário para alcançar os 35 anos de contribuição e aposentadoria – e o reinvestimento dos proventos recebidos, há um potencial relevante de crescimento patrimonial e da renda passiva mensal. Adicionalmente, em situações excepcionais, os ativos podem ser vendidos, aproveitando-se da isenção de imposto de renda para alienações mensais de R\$ 20.000,00, bem como do benefício da isenção sobre dividendos.

Sendo assim, o método de Graham se destacou por gerar o maior retorno médio, com o menor risco, além de se demonstrar viável de acordo com o IS. Ele apresentou um retorno considerável sobre o capital, distribuição de dividendos semelhante aos outros métodos e superou a inflação média calculada, o que demonstra crescimento patrimonial. Com base nos testes e análises realizados nesta pesquisa, conclui-se que a estratégia de Graham é a mais viável para investidores com foco previdenciário e de longo prazo.

Porém, para garantir a padronização e coerência na análise, a inflação acumulada dos 25 anos foi convertida em uma taxa média anual, o que permitiu a correção uniforme dos preços dos ativos. No entanto, essa média não considera as variações anuais da inflação, impedindo a avaliação de cenários com taxas mais altas ou mais baixas (otimista e pessimista), que impactariam diretamente no preço teto de compra, para melhor capturar as incertezas do ambiente econômico.

5 CONCLUSÕES

O presente trabalho propôs a construção de um portfólio de ações com a aplicação de quatro estratégias de investimento, analisando os fundamentos dos ativos conforme as restrições de cada método. A viabilidade do investimento como forma de complementar a previdência pública foi avaliada por meio da relação entre retorno e risco de cada portfólio. A

pesquisa adotou uma metodologia aplicada, com abordagem descritiva e procedimento documental

O objetivo foi aplicar ferramentas de análise de dados ao mercado financeiro, demonstrando seu potencial e apresentando alternativas baseadas em estratégias amplamente conhecidas. Ressalta-se que nenhum dos métodos deve ser seguido sem estudos aprofundados, e que os resultados apresentados não constituem recomendação de investimento.

Utilizando as equações de retorno-risco, pode-se comparar os ativos em relação a sua porcentagem de valor absoluto em cada carteira, assim como seu retorno e risco individuais e assim formar retornos e risco médio para cada portfólio e assim constatar os métodos que obtiveram maiores retornos e riscos.

Aplicando o Python, pode-se realizar uma análise completa de indicadores, empresas, algoritmos de maneira automática para os métodos e ao se construir o código foi possível aplicar em todos os métodos fazendo poucas mudanças, o que mostra o poder de escalonamento nas análises financeiras ao utilizar esta ferramenta.

O relatório proposto e desenvolvido possibilitou analisar o comportamento de cada ativo de maneira flexível, interativa e facilitada. O que mostra o potencial do *Business Intelligence* para a tomada de decisões estratégicas. Além de como utilizar uma ferramenta *open source* e de código aberto promove uma flexibilidade na extração e carregamento em diferentes estruturas de *dataset*, além do tamanho da personalização individual possibilitada, ao se comparar com outros *softwares* de *BI* existentes.

Conclui-se que, sob a perspectiva de análise fundamentalista e considerando uma estratégia de longo prazo, os métodos de seleção de ações estudados representam alternativas viáveis para a formação de uma renda complementar à aposentadoria pública. Os resultados demonstram que uma carteira devidamente balanceada entre diferentes ativos e níveis de exposição pode alcançar retornos superiores aos oferecidos por ativos livres de risco, ao mesmo tempo em que pode reduzir a volatilidade da carteira mitigando assim os riscos associados. Dessa forma, a alocação estratégica de recursos em investimentos que geram dividendos demonstra-se uma alternativa viável e eficiente para a complementação da renda previdenciária, proporcionando maior segurança financeira no longo prazo

Além disso, a utilização de critérios fundamentados na análise de indicadores econômicos e financeiros, conforme proposto pelos métodos de Graham, Bazin, Barsi e Greenblatt, mostrou-se eficaz para identificar ativos com bom potencial de valorização e geração de renda passiva. Estas estratégias, quando aplicadas de forma disciplinada e com visão de longo prazo, podem contribuir significativamente para a construção de uma aposentadoria

mais segura e sustentável, especialmente em cenários de incerteza quanto à manutenção do atual sistema previdenciário público.

Como pesquisa futura derivada deste trabalho recomenda-se a aplicação de pesquisa operacional, executando assim a equação da carteira de mínima variação alcançando a carteira ótima segundo o Markowitz, se buscará a porcentagem de cada ativo para se obter o risco de portfólio mínimo, além de comparar se a carteira e compras realizadas de maneira automática e constante são um exemplo de carteira eficiente.

Outra possibilidade relevante é simular cenários com foco no valor futuro dos investimentos, considerando metas de longo, como a aposentadoria. Projetar a evolução patrimonial com aportes ao longo de 35 anos pode orientar decisões mais assertivas, permitindo comparar métodos de seleção de ativos e analisar estratégias sob perspectiva otimista, pessimista e média, tornando o planejamento mais realista e alinhado aos objetivos do investidor.

REFERÊNCIAS

- AFSHARIAN, Mohsen. **Data science essentials in business administration: A multidisciplinary perspective**. Decision Analytics Journal, 2024.
- ASSAF NETO, Alexandre. **Matemática Financeira e suas aplicações** 15. ed. 2. reimpr. Barueri, SP: Atlas, Instituto Assaf, 2022.
- ASSAF NETO, Alexandre. **Mercado financeiro**. 15. ed. 2. reimpr. Barueri, SP: Atlas, Instituto Assaf, 2021.
- BANCO CENTRAL. **Calculadora do Cidadão**. Disponível em: <<https://www3.bcb.gov.br/CALCIDADAO/publico/exibirFormCorrecaoValores.do?method=exibirFormCorrecaoValores>>. Acessado em: 20 de dezembro de 2024.
- BASILIO, Marcio Pereira; FREITAS, Jéssica Galdino de; KAMPFFE, Milton George Fonseca; REGO, Ricardo Bordeaux. **Investment portfolio formation via multicriteria decision aid: A Brazilian stock market study**. Journal of Modelling in Management, v. 13, n. 2, p. 394-417, 2018.
- BEHRMAN, Kennedy. **Fundamentos de Python para ciência de dados**. Porto Alegre: Bookman, 2023.
- CHANG, Wo.; GRADY, Nancy. **NIST Big Data Interoperability Framework: Volume 1**, NIST Pubs 2019.
- CHEN, Sheng-Qun; BAI, Jie. **Data-driven decision-making model for determining the number of volunteers required in typhoon disasters**. Journal of Safety Science and Resilience, v. 4, n. 3, p. 229-240, 2023.
- CHIMIDT, Rafael Fabri; KOZYREFF, Ernee. **Análise exploratória de uma década de competições no Kaggle**. Anais ... Encontro Nacional de Engenharia de Produção, 4 nov. 2020.
- DABBAS, Elias. **Interactive Dashboards and Data Apps with Plotly and Dash: Harness the power of a fully-fledged frontend web framework in Python—no JavaScript required**. Packt Publishing Ltd, 2021.
- DA SILVA, Leandro Augusto; PERES, Sarajane Marques; BOSCARIOLI, Clodis. **Introdução à mineração de dados: com aplicações em R**. Elsevier Brasil, 2017.
- DE OLIVEIRA, Sidney Lino; DE SOUZA, Cláudia Artane Lima; BENDER, Fábio Everton Alves Da Silva; RIBEIRO, Felipe Augusto Vtegas; GOMES, Guilherme March; FILHO, João Arcanjo Sampato. **Estratégias de investimento no mercado de capitais métricas para uma carteira de investimentos com foco em dividendos sob a ótica de quatro investidores renomados: Luiz Barsi, Décio Bazin, Joel Greenblatt e Benjamin Graham**. JAMES-Journal of Accounting Management Economics and Sustainability, v. 1, n. 1, 2023.
- DE SOUZA, Gabriela Maia. **Business Intelligence (BI) como uma ferramenta de gestão auxiliando na tomada de decisão**. Revista Ibero-Americana de Humanidades, Ciências e Educação, v. 8, n. 4, p. 1002-1019, 2022.

DOMINGUES, Augusto; FERREIRA, Márcio Antônio Couto; NEVES, Salomão Franco. **Aposentadoria privada: uma solução através do mercado de ações utilizando o método value investing.** OBSERVATÓRIO DE LA ECONOMÍA LATINOAMERICANA, v. 22, n. 3, p. e3606-e3606, 2024.

DOS SANTOS, Bruna; GONZAGA, Vinicius Drumond; STRAUSS, Edilberto; JUNIOR, Manoel Villas Bôas. **Tomada de decisão com business intelligence na gestão logística.** Projectus, v. 8, n. 1, p. 101-112, 2023.

FAES, Daniel. **Use of Python programming language in astronomy and science.** arXiv preprint arXiv:1807.04806, 2018.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Editora Atlas, 2002.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa.** 6. ed. São Paulo: Editora Atlas, 2017.

GELATTI, Elisangela; DO NASCIMENTO AFONSO, Paulo Fernando. **Desempenho de carteiras com diversificação de ativos locais e internacionais: uma análise para o período de pandemia e pós-pandemia do covid-19.** SINERGIA-Revista do Instituto de Ciências Econômicas, Administrativas e Contábeis, v. 28, n. 1, p. 123-141, 2024.

GRUS, Joel. **Data science do zero.** Rio de Janeiro: Alta books, 2016.

IBM. **What is ETL?** Disponível em: < <https://www.ibm.com/topics/etl> > Acessado em: 19 de novembro de 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Projeções da população por sexo e idade Brasil e Unidades da Federação 2010-2060.** Disponível em: < https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/78b962ace2419a967d57addd41377397.pdf >. Acessado em 21 de novembro de 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Projeto Atlas Escolar.** Disponível em: < <https://atlasescolar.ibge.gov.br/brasil/3051-caracteristicas-demograficas/idade/21898-piramide-etaria-1970-2022.html> >. Acessado em 26 de outubro de 2024.

KLEINERT, Lucas Schmidt. **Estudo comparativo entre duas estratégias de investimentos aplicadas no mercado de capitais.** Trabalho de Conclusão de Curso (Bacharelado em Engenharia de Produção) Universidade Federal de Santa Maria. Rio Grande do Sul. 2024.

LONG, Qiwen. **Investigation on big data evaluation and visualization of internet of things based on edge computing.** Measurement: Sensors, v. 33, p. 101177, 2024.

MANDAL, Arunima; SHAO, Yuanhang; LIU, Xiuwen. **Automatic Historical Stock Price Dataset Generation Using Python.** arXiv preprint arXiv:2308.13414, 2023.

MARTINS, Carlos Henrique Batista; JUNIOR, Fabio Lopes; BORTOLOTTI, Luis Marcelo. **Big data: proposta de um data warehouse para apoio à tomada de decisão em uma companhia de abastecimento de água e saneamento básico**. In: Congresso de Tecnologia-Fatec Mococa. 2023.

MARQUESONE, Rosangela. **Big Data: Técnicas e tecnologias para extração de valor dos dados**. Editora Casa do Código, 2016.

MCKINNEY, Wes. **Python for data analysis: Data wrangling with pandas, numpy, and jupyter**. " O'Reilly Media, Inc.", 2022.

MINISTÉRIO DO TRABALHO E PREVIDÊNCIA. **Módulo 07 RGPS e RPPS: O que é a Previdência Pública?** Disponível em: < <https://www.gov.br/previdencia/pt-br/assuntos/previdencia-complementar/mais-informacoes/arquivos/pbefrgps.pdf> >. Acessado em 26 de outubro de 2024.

MITTAL, Saksham; BHATTACHARYA, Sujoy; MANDAL, Satrajit. **Characteristics analysis of behavioural portfolio theory in the Markowitz portfolio theory framework**. Managerial finance, v. 48, n. 2, p. 277-288, 2022.

MIX VALE. **Tabela histórica dos salários de 1994 a 2024**. Disponível em: < https://www.mixvale.com.br/tabela-historica-dos-salarios-minimos-de-1994-a-2024/?gad_source=1&gclid=Cj0KCQiA3sq6BhD2ARIsAJ8MRwXtONZhQOMwk2TAOrHz5_fIMg_BC0WSo_U1ozMQYd20W21Ypjh_w3oaAgGsEALw_wcB >. Acessado em: 20 de dezembro de 2024.

MORAIS, Vitor Morita de. **Investimentos em renda variável no mercado brasileiro: da teoria na gestão de portfólios à prática no python**. Trabalho de conclusão de curso. (Bacharelado em Ciências Econômicas) – Instituto de Economia e Relações Internacionais. Universidade Federal de Uberlândia. Uberlândia.2024.

NAJAFABADI, Maryam M.; VILLSNUSTRE, Flavio; KHOSHGOFTAAR, Taghi M; SELIYA, Naeem; WALD, Randall; Muharemagic, Edin. **Deep learning applications and challenges in big data analytics**. Journal of big data, v. 2, p. 1-21, 2015.

OLIVEIRA, Cícero Elielson de Mendonça. **Métodos de seleção de carteira: uma análise dos métodos de Bazin, Graham e Markowitz**. Trabalho de Conclusão de Curso (Bacharelado em Administração). Universidade Federal do Rio Grande do Norte – Centro de Ciências Sociais Aplicadas Departamento de Ciências Administrativas. Rio Grande do Norte. 2019.

OLIVEIRA, Eliel Jocsã Marinho de. **Otimização de carteiras com inteligência artificial: uma aplicação prática da teoria de markowitz em python**. Trabalho de Conclusão de Curso (Bacharelado em Ciências Econômicas). Universidade Federal de Pernambuco – Centro de Acadêmico do Agreste. Pernambuco. 2024

OLIVEIRA, Moacir de Souza; MELO, Nicole de Farias; SANTOS, Leandro Colevati dos; OLIVEIRA, Wellington Pinto de. **Banco de dados no-sql x banco de dados sql: Estudo de Desempenho em Grandes Massas**. South American Development Society Journal, v. 4, n. 11, p. 298, 2018.

ORACLE. **O que é um banco de dados relacional?** Disponível em: < <https://www.oracle.com/br/database/what-is-a-relational-database/> >. Acessado em: 30 de novembro de 2024.

ORACLE. **O que é ETL?** Disponível em: < <https://www.oracle.com/br/integration/what-is-etl/> > acessado em: 19 de novembro de 2024.

PAULEEN, David J.; WANG, William YC. **Does big data mean big knowledge? KM perspectives on big data and analytics.** Journal of Knowledge Management, v. 21, n. 1, p. 1-6, 2017.

PYPI. **Investpy.** Disponível em :< <https://pypi.org/project/investpy/>>. Acessado em 24 de fevereiro de 2025.

PYTHON.LAND. **Python Tutorial.** Disponível em :< <https://python.land/python-tutorial> > . Acessado em 24 de fevereiro de 2025.

RAUTENBERG, Sandro; DO CARMO, Paulo Ricardo Viviurka. **Big data e ciência de dados: complementariedade conceitual no processo de tomada de decisão.** Brazilian Journal of Information Science, v. 13, n. 1, p. 56-67, 2019.

RECEITA FEDERAL. **Taxa de juros SELIC.** 2025. Disponível em: < <https://www.gov.br/receitafederal/pt-br/assuntos/orientacao-tributaria/pagamentos-e-parcelamentos/taxa-de-juros-selic> >. Acessado em: 05 de janeiro de 2025.

REIS, Tiago. **CAGR: o que é? Qual sua importância nos investimentos. 2019.** Disponível em: < <https://www.suno.com.br/artigos/cagr/> >. Acessado em: 20 de dezembro de 2024.

RICHARDS, Tyler. **Getting Started with Streamlit for Data Science: Create and deploy Streamlit web applications from scratch in Python.** Packt Publishing Ltd, 2021.

RODRIGUES, Marcio Jonavicius. **Data Science for Finance: automated investment recommendation with python.** Dissertação de mestrado (Mestrado em Estatística e Gestão de Informação), Universidade Nova de Lisboa. Lisboa, 2022

SANTIAGO NINA, Alex. **Análise de investimentos baseados em proventos de ações de bancos brasileiros e efeitos do coronavírus.** Revista Evidenciação Contábil & Finanças, v. 9, n. 2, 2021.

SANTOS, Rebeca Michaelis Sequeira. **Um estudo sobre a resiliência de ações e fundos imobiliários: uma observação no ano de 2020.** Trabalho de Conclusão de Curso (Bacharelado em Ciências Contábeis) -Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.

SANT'ANA, Ricardo César Gonçalves. **Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação.** Informação & Informação, v. 21, n. 2, p. 116–142-116–142, 2016.

SHARDA, Ramesh; DELEN, Dursun; TURBAN, Efraim. **Business Intelligence e Análise de Dados para Gestão do Negócio-4.** Bookman editora, 2019.

TAN, Kar Hooi. **Managerial perspectives of big data analytics capability towards product innovation.** Strategic Direction, v. 34, n. 8, p. 33-35, 2018.