



**UNIVERSIDADE FEDERAL DA BAHIA**  
**FACULDADE DE ECONOMIA**  
**CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS**

**CLEITON OTAVIO DA EXALTAÇÃO ROCHA**

**AVALIAÇÃO DE CARTEIRAS CRIADAS PELOS MODELOS DE  
FAMA-FRENCH E FRONTEIRA EFICIENTE DE MARKOWITZ NO  
MERCADO ACIONÁRIO BRASILEIRO ENTRE 2011 E 2021**

Salvador

2022

**CLEITON OTAVIO DA EXALTAÇÃO ROCHA**

**AVALIAÇÃO DE CARTEIRAS CRIADAS PELOS MODELOS DE  
FAMA-FRENCH E FRONTEIRA EFICIENTE DE MARKOWITZ NO  
MERCADO ACIONÁRIO BRASILEIRO ENTRE 2011 E 2021**

Trabalho de conclusão de curso apresentado ao curso de Ciências Econômicas da Universidade Federal da Bahia, como requisito parcial para a obtenção do grau de bacharel em Ciências Econômicas.

Área de Concentração: Finanças

Orientador: Prof. Dr. Reinan Ribeiro Souza Santos

Salvador

2022

R672

Rocha, Cleiton Otavio da Exaltação.

Avaliação de carteiras criadas pelos modelos de Fama-French e fronteira eficiente de Markowitz no mercado acionário brasileiro entre 2011 e 2021/ Cleiton Otavio da Exaltação Rocha. - - Salvador, 2022.

87 f.; il.; tab.; graf.

TCC (Graduação) – Universidade Federal da Bahia, Faculdade de Economia. Orientador: Prof. Dr. Reinan Ribeiro Souza Santos.

1. Mercado financeiro. 2. Mercado de ações. 3. Ações - portfólio. I. Universidade Federal da Bahia. II. Santos, Reinan Ribeiro Souza. III. Título.

CDD: 332.1

**CLEITON OTAVIO DA EXALTAÇÃO ROCHA**

**AVALIAÇÃO DE CARTEIRAS CRIADAS PELOS MODELOS DE  
FAMA-FRENCH E FRONTEIRA EFICIENTE DE MARKOWITZ NO MERCADO  
ACIONÁRIO BRASILEIRO ENTRE 2011 E 2021**

Trabalho de conclusão de curso apresentado ao curso de Ciências Econômicas da Universidade Federal da Bahia, como requisito parcial para a obtenção do grau de bacharel em Ciências Econômicas.

Aprovado em 04 de julho de 2022.

Banca Examinadora



---

Reinan Ribeiro Souza Santos

**Prof. Dr. Reinan Ribeiro Souza Santos**  
Universidade Federal da Bahia - UFBA



---

**Prof. Dr. Bruno dos Santos Gois**  
Universidade Federal do Rio Grande do Norte - UFRN



---

**Prof. Me. Rafael Sales Rios**  
Confederação Nacional da Indústria - CNI

## **AGRADECIMENTOS**

Agradeço primeiramente à minha mãe e minhas tias, cujo investimento, apoio e confiança foram fundamentais para que eu chegasse até aqui. Aos colegas da SEI por todo o conhecimento compartilhado durante minha estadia por lá. Aos amigos do ensino médio por todo o companheirismo. Ao Prof. Reinan Santos pela orientação desta monografia. Aos funcionários da FEUFBA por toda ajuda ao longo desses anos.

*“Não importa quão estreito o caminho, quão cheia de punições a minha sina, eu sou o mestre do meu destino, eu sou o comandante da minha alma.”*  
*(William Ernest Henley, 1875)*

## RESUMO

Neste trabalho comparamos os modelos de escolha de portfólio de ações trazendo informações sobre o melhor retorno dado o seu nível de risco considerando todas as ações listadas na Bolsa de Valores de São Paulo (BOVESPA). Os portfólios de ações foram construídos com base em dois modelos utilizados na literatura: 1) na explicação dos retornos dos ativos elaborada por Fama e French (1993), que busca explicar o retorno de mercado através de três grandes fatores e 2) na construção de um portfólio baseado no conceito chamado de Fronteira Eficiente, apresentado por Markowitz (1959) - nele é apresentado que o risco de uma carteira não é dado simplesmente pela média dos ativos individuais, mas sim pela diversificação da carteira de investimento como um todo. Posteriormente, derivado do trabalho de Markowitz, surge o modelo CAPM (elaborado por Sharpe (1964), Lintner (1965) e Mossin (1966)), que define a taxa de retorno teórica adequada para determinado ativo de acordo com a carteira de investimento previamente montada e o índice de Sharpe para seleção de carteiras. O estudo compreendeu as ações listadas na BOVESPA no período 2011-2021. Foi calculado os prêmios dos fatores de risco e dos retornos das ações e carteiras, e se testou o comportamento dos dois modelos através de medidas de risco e comparação com a performance do índice Ibovespa ao longo do mesmo período. O que se observou foi que, para o período 2011-2021, a seleção de um portfólio com base nos pressupostos de Markowitz trouxe melhores retornos em comparação ao modelo de Fama-French.

Palavras-chave: Portfólio. Risco. Retorno. Fama-French. Markowitz. CAPM.

## **ABSTRACT**

In this work we compare stock portfolio choice models providing information on the best return given their risk level considering all stocks listed on the São Paulo Stock Exchange (BOVESPA). The stock portfolios were built based on two models used in the literature: 1) in the explanation of asset returns elaborated by Fama and French (1993), which seeks to explain the market return through three major factors and 2) in the construction of a portfolio based on the concept called Efficient Frontier, presented by Markowitz (1959) - it is presented that the risk of a portfolio is not simply given by the average of individual assets, but by the diversification of the investment portfolio as a whole. Subsequently, derived from the work of Markowitz, the CAPM model appears (developed by Sharpe (1964), Lintner (1965) and Mossin (1966)), which defines the appropriate theoretical rate of return for a given asset according to the investment portfolio previously mounted and the Sharpe ratio for portfolio selection. The study comprised the shares listed on BOVESPA in the period 2011-2021. The premiums of the risk factors and the returns of the shares and portfolios were calculated, and the behavior of the two models was tested through risk measures and comparison with the performance of the Ibovespa index over the same period. What was observed was that, for the period 2011-2021, the selection of a portfolio based on the Markowitz assumptions brought better returns compared to the Fama-French model.

Keywords: Portfolio. Risk. Return. Fama-French. Markowitz. CAPM.



## LISTA DE ILUSTRAÇÕES

Figura 1.	Modelo teórico da fronteira eficiente e da taxa livre de risco . . . . .	19
Figura 2.	Fator SMB . . . . .	21
Figura 3.	Fator HML . . . . .	22
Figura 4.	Seis carteiras geradas através dos fatores SMB e HML . . . . .	22
Figura 5.	Histograma do preço de fechamento ajustado dos ativos . . . . .	27
Figura 6.	Risco-Retorno das Ações . . . . .	27
Figura 7.	Fator de Mercado - Fama e French . . . . .	31
Figura 8.	Fator SMB - Fama e French . . . . .	32
Figura 9.	Fator HML - Fama e French . . . . .	32
Figura 10.	Retorno diário - todos os ativos . . . . .	33
Figura 11.	Retorno diário - <i>SmallCaps</i> . . . . .	33
Figura 12.	Retorno e Volatilidade das Carteiras Simuladas . . . . .	34
Figura 13.	Retorno e Volatilidade das Carteiras Simuladas - <i>SmallCaps</i> . . . . .	34
Figura 14.	Fronteira Eficiente . . . . .	35
Figura 15.	Fronteira Eficiente - <i>SmallCaps</i> . . . . .	35
Figura 16.	Retorno das Carteiras Simuladas . . . . .	36

## LISTA DE TABELAS

Tabela 1.	Visualização parcial do banco de dados . . . . .	24
Tabela 2.	Bibliotecas R e Python Utilizadas . . . . .	25
Tabela 3.	Estatísticas descritivas das colunas do banco de dados . . . . .	26
Tabela 4.	Média diária do retorno e volatilidade, por setor do ativo . . . . .	28

## LISTA DE ABREVIATURAS E SIGLAS

Ação ON	Ação ordinária
Ação PN	Ação preferencial
B/M	<i>Book-to-market equity</i>
CAPM	<i>Capital Asset Pricing Model</i>
HML	<i>High minus Low</i>
IBOVESPA	Índice da Bolsa de Valores de São Paulo
ME	<i>Market equity</i>
NASDAQ	<i>NASDAQ Stock Market</i>
SMB	<i>Small minus Big</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>16</b>
2.1	MODELO DE FRONTEIRA EFICIENTE	16
<b>2.1.1</b>	<b>Risco</b>	<b>16</b>
<b>2.1.2</b>	<b>Retorno Esperado</b>	<b>17</b>
<b>2.1.3</b>	<b>A Fronteira Eficiente</b>	<b>17</b>
2.2	MODELO CAPM	18
2.3	ÍNDICE DE SHARPE	19
2.4	MODELO DE TRÊS FATORES DE FAMA E FRENCH	20
<b>2.4.1</b>	<b>Explicação dos fatores SMB e HML</b>	<b>21</b>
<b>3</b>	<b>BASE DE DADOS</b>	<b>24</b>
3.1	EXTRAÇÃO DOS DADOS	24
3.2	TRATAMENTO DOS DADOS	24
3.3	CONSTRUÇÃO DO MODELO DE MARKOWITZ	27
3.4	CONSTRUÇÃO DOS FATORES DE FAMA E FRENCH	29
<b>4</b>	<b>RESULTADOS</b>	<b>31</b>
4.1	MODELO FAMA-FRENCH	31
4.2	MODELO CAPM	33
4.3	RETORNO DAS CARTEIRAS SIMULADAS	35
4.4	LIMITAÇÕES	37
4.5	CONTINUIDADE DE PESQUISA	37
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>38</b>
	<b>REFERÊNCIAS</b>	<b>39</b>
	<b>APÊNDICE A – ESTATÍSTICAS DAS CARTEIRAS SIMULADAS</b>	<b>41</b>
	<b>APÊNDICE B – CORRELAÇÃO DE CADA CARTEIRA SIMULADA COM A TAXA LIVRE DE RISCO</b>	<b>42</b>
	<b>APÊNDICE C – CÓDIGO R: GERANDO CARTEIRA DE MARKOWITZ</b>	<b>43</b>
	<b>APÊNDICE D – CÓDIGO R: GERANDO CARTEIRA DE MARKOWITZ - <i>SMALL- CAPS</i></b>	<b>57</b>

<b>APÊNDICE E – CÓDIGO R: GERANDO FATOR SMB - FAMA-FRENCH . .</b>	<b>70</b>
<b>APÊNDICE F – CÓDIGO R: GERANDO FATOR HML - FAMA-FRENCH . .</b>	<b>74</b>
<b>APÊNDICE G – CÓDIGO R: GERANDO FATOR MERCADO - FAMA-FRENCH</b>	<b>79</b>
<b>APÊNDICE H – CÓDIGO R: GERANDO CARTEIRA FAMA-FRENCH . . .</b>	<b>82</b>
<b>APÊNDICE I – CÓDIGO R: UNINDO TODAS AS CARTEIRAS SIMULADAS</b>	<b>84</b>

## 1 INTRODUÇÃO

A determinação do risco-retorno aceitável ao realizar um investimento tem sido um tema debatido e pesquisado no campo das Finanças de forma recorrente. O marco temporal para o início desse debate foi com Markowitz (1952) e o que viria a ser conhecido posteriormente como a Teoria Moderna do Portfólio. Markowitz apresenta a diversificação de ativos como principal instrumento para a redução do risco de um portfólio de investimentos. Para Markowitz, o investidor leva em consideração dois fatores principais ao investir: o risco e o retorno da aplicação. Ao construir um portfólio com um menor risco, o investidor aceita um menor retorno, ao que pese uma menor flutuação no montante que fora aplicado. A forma de mensurar a eficiência de uma carteira, portanto, é relacionada pelo binômio risco e retorno; o investidor vai aceitar um maior risco apenas se a aplicação oferecer uma chance de maior retorno. Markowitz postula que dado dois ativos quaisquer com o mesmo e retorno e riscos diferentes, o investidor vai optar pelo ativo de menor risco sempre.

Entende-se o risco, segundo Assaf (2006), como o grau de incerteza sobre a rentabilidade de um investimento. O conceito de risco para Markowitz (1952) é separado em dois tipos: o risco idiossincrático e o sistemático. O Risco Idiossincrático é o risco relacionado ao próprio ativo em si, por exemplo, ao comprar uma ação, o investidor sempre conta com a possibilidade, mesmo que remota, daquela companhia detentora da ação falir. Ao comprar um título de dívida pública de um país, existe o risco de um calote, etc. Observa-se, portanto, que ao pulverizar sua aplicação em diversos ativos, o investidor consegue reduzir o risco idiossincrático. Isso porque, ao investir em mais de um ativo, mesmo que o valor investido não seja recuperado em algum dos ativos, o restante da alocação, de maneira geral, está preservada pelos demais ativos. O segundo risco, o Risco Sistemático, está associado à eventos exógenos que podem atingir todos os ativos, como a variação das taxas de juros, câmbio ou qualquer outro evento extraordinário que afete o mercado como um todo, ou seja, esse risco é aplicado a todos os ativos de modo geral e não por características únicas de cada ativo.

O retorno de um ativo é a relação entre a quantidade de dinheiro ganho (ou perdido) como resultado de um investimento e a quantidade de dinheiro investido nesse ativo.

Markowitz (1959) apresenta os investidores como seres que tem aversão ao risco, optando sempre por um maior retorno e menos risco; caso os investidores aceitem uma maior variância (ou seja, um maior risco) no seu portfólio de ativos, eles também vão exigir um maiores retornos esperados.

Baseando-se nos trabalhos e teorias propostas por Markowitz (1959), Sharpe (1964), Lintner (1965) e Mossin (1966) desenvolveram o Modelo de Precificação de Ativos Financeiros (CAPM - Capital Asset Pricing Model). De acordo com esta literatura, o retorno de um ativo é determinado

pelo retorno do ativo livre de risco e pelo prêmio de mercado ponderado pelo fator beta, (medida da sensibilidade dos retornos deste ativo em relação aos retornos da carteira de mercado); esse fator beta, por sua vez, seria o fator que explicaria a diferença de retorno exigido entre os ativos. Desta forma, ativos com diferentes betas apresentariam diferentes retornos exigidos, sendo a relação entre retorno e beta linear. O ativo livre de risco seria um ativo que, em tese, apresenta risco zero, ou muito próximo de zero, e retorno virtualmente garantido. No mercado acionário brasileiro, ao se deparar com a necessidade de escolher qual será o referencial que vai assumir o papel da taxa livre de risco, os pesquisadores costumam adotar a taxa básica de juros da economia, a Selic. Todavia, uma outra alternativa frequentemente adotada é a taxa de juros do CDI diário, denominada taxa DI.

Desde a sua formulação, o CAPM tem sido amplamente utilizado por diversos autores, que buscaram testar a validade do modelo e procuraram mensurar o poder de explicação do fator de risco na variação dos retornos dos ativos e a relação entre retorno e o fator beta; Em Rochman e Eid Jr. (2006), por exemplo, os autores buscaram investigar se é melhor investir em fundos ativos ou passivos, já Rabelo e outros (2007) analisou se as empresas que adotam práticas de governança corporativa rigorosas apresentam melhor performance ao longo do tempo. Apesar de ser bastante utilizado no campo das Finanças, o modelo CAPM sofreu com algumas críticas ao longo do tempo, entre elas, Fama e French (1992) identificaram ineficiências no modelo de Sharpe para a explicação dos retornos dos ativos. Para esses autores, além do fator beta, existem outros fatores que influenciam na variação dos retornos. Com base nessa crítica, Fama e French (1993) formularam um modelo de três fatores que, segundo os autores, são mais eficientes em explicar o retorno dos ativos. Os fatores de risco do modelo são o mercado, conforme definido por Sharpe e incluído no CAPM, o tamanho da empresa, definido pelo valor de mercado do patrimônio líquido, e o índice *Book-to-Market* ou índice B/M, definido pela relação entre o valor contábil e o valor de mercado do patrimônio líquido.

Com base nos estudos de Fama e French (1993), Sharpe (1970) e Markowitz (1959), o principal objetivo desse trabalho é investigar se o modelo de três fatores consegue explicar as variações dos retornos das ações listadas na Bolsa de Valores de São Paulo melhor do que o modelo CAPM. Este objetivo será abordado seguindo três etapas: em primeiro lugar, será aplicada sobre os dados do mercado acionário brasileiro a mesma metodologia de teste empregada por Fama e French (1993) no estudo do modelo no mercado acionário americano. Em segundo lugar, será aplicada a metodologia proposta por Markowitz (1959) e aprimorada por Sharpe (1970) também no mercado acionário brasileiro. Por fim, com os modelos de carteiras gerados, serão analisados os resultados obtidos decorrentes da aplicação das duas metodologias de teste chegando a conclusão se o modelo de três fatores é superior ao modelo CAPM na explicação dos retornos.

O trabalho está organizado em seis capítulos, o presente capítulo é a introdução. No capítulo 2 será discutido o referencial teórico do trabalho, com descrição histórica, teórica e matemática das teorias utilizadas no trabalho. No capítulo 3 é onde será abordado o banco de dados utilizado no

trabalho, suas características, quais *softwares* e fontes de dados foram utilizadas. No capítulo 4 é onde será detalhada a metodologia aplicada em cada teoria. No capítulo 5, os resultados obtidos são mostrados e analisados. Por fim, no capítulo 6 temos a conclusão do trabalho, demonstrando que, no mercado acionário brasileiro, apesar do modelo CAPM ser mais simples, uma alocação seguindo os preceitos de fronteira eficiente estabelecidas por Markowitz ainda traz melhores retornos do que o modelo Fama-French de três fatores.



## 2 REFERENCIAL TEÓRICO

### 2.1 MODELO DE FRONTEIRA EFICIENTE

A Fronteira Eficiente de Markowitz, também conhecida como Teoria do Portfólio Moderno, é uma representação gráfica da relação entre o Risco e o Retorno de um portfólio. A fronteira de Markowitz busca otimizar a composição de uma carteira para um determinado nível de risco de forma a obter o maior retorno possível.

Os conceitos desta teoria foram inicialmente formulados por Markowitz, com a publicação do artigo *Portfolio Selection*, em 1952. Markowitz (1952) define que as decisões de investimento dos agentes para seleção de uma carteira de ativos consideram dois fatores: o retorno esperado e a variância do retorno dos ativos financeiros. Para a tomada de decisão, Markowitz, considera que os investidores buscam maximizar o retorno esperado, considerando que são avessos ao risco (alta variância) e assim, caso tenham que se sujeitar a ativos com alta variância, os investidores exigirão altos retornos esperados desses ativos como forma de recompensa. As ideias propostas por Markowitz eram contrárias a ideia vigente na época, de que a melhor opção para a composição da carteira consistia na concentração de investimentos em ativos que ofereciam os maiores retornos, o que limitava as opções de investimentos e nem sempre fornecia o retorno desejado de fato.

A metodologia do modelo consiste em montar um portfólio que avalia os ativos com base no risco futuro, retorno e premissas de correlação. Uma série de alocações diferentes é determinada com base nos dados; entre essas possíveis alocações, as carteiras que maximizam o retorno projetado e minimizam o risco são chamadas de fronteira eficiente. A teoria considera que, dada uma certa tolerância ao risco, o investidor busca sempre a carteira com maior retorno esperado.

O retorno esperado dos ativos, sua respectiva volatilidade e a covariância entre os ativos são necessários para determinar a fronteira eficiente.

Um método popular de estimativa é assumir que os retornos futuros tenderão a seguir a média dos retornos anteriores. O desvio padrão dos retornos é comumente usado para medir a volatilidade.

A covariância descreve a relação entre a direção do movimento de dois ativos diferentes; se os ativos tiverem tendência a se mover na mesma direção, o resultado do cálculo será positivo; se os ativos tiverem tendência a se mover em direções opostas, o resultado do cálculo será negativo.

#### 2.1.1 Risco

Sendo uma medida do grau de incerteza sobre a rentabilidade de um investimento (ASSAF, 2006), o risco pode ser mensurado matematicamente como a volatilidade, ou desvio padrão, dos retornos de um título ou índice de mercado. Esse método é utilizado porque como o desvio padrão

representa uma medida de dispersão, quanto mais o preço de uma ação varia num período curto de tempo, maior o risco de se ganhar ou perder dinheiro negociando esta ação. Para Markowitz (1952), o risco é um dos fatores principais a ser considerado pelos investidores, ao lado da rentabilidade e do prazo de retorno. A equação do risco (desvio-padrão) adotada por Markowitz de uma carteira de dois ativos (A e B) é denotada da seguinte forma:

$$\sigma_p = \sqrt{(w_A^2 \cdot \sigma_A^2) + (w_B^2 \cdot \sigma_B^2) + 2 \cdot w_A \cdot w_B \cdot COV_{A,B}} \quad (2.1)$$

Onde,

$w_A$  e  $w_B$  representam as porcentagens do valor da carteira destinada ao ativo A e B.

$\sigma_A^2$  e  $\sigma_B^2$  representam a variância dos ativos da carteira.

$COV_{A,B}$  representa a covariância entre os ativos A e B.

### 2.1.2 Retorno Esperado

No campo das finanças, a expectativa em relação a um investimento é conhecida como retorno esperado do mesmo. Em síntese, é quanto se espera ganhar ao investir em determinado ativo. Segundo Samanez (2007), o retorno esperado de um ativo em uma carteira de ativos é a média central da distribuição probabilística dos retornos desse ativo. O retorno de uma carteira de ativos pode ser estimado calculando-se a soma dos produtos dos retornos dos ativos pelos respectivos pesos de cada ativo na carteira.

$$R(E) = X_A \bar{R}_A + X_B \bar{R}_B \quad (2.2)$$

Onde,

$X_A$  e  $X_B$  são os percentuais da participação dos ativos A e B na carteira (em termos de valor).

$\bar{R}_A$  e  $\bar{R}_B$  são os retornos individuais dos ativos A e B.

### 2.1.3 A Fronteira Eficiente

Fronteira Eficiente é um conceito que trata da combinação de um conjunto de ativos para construir infinitos portfólios diversificados; esse conjunto de ativos é chamado normalmente de 'carteiras'. As carteiras que apresentam o melhor retorno dado o nível de risco vão compor a fronteira eficiente e são denominadas, segundo Markowitz (1952), como 'carteiras ótimas'.

Uma carteira é eficiente, isto é, faz parte da fronteira eficiente, se três afirmações são respondidas: (i) não existe outra carteira com maior retorno e menor risco, (ii) não existe outra carteira com maior retorno e mesmo nível de risco, (iii) não existe outra carteira com mesmo nível de retorno e menor risco.

A fronteira é essencial para mostrar como a diversificação melhora o perfil de risco sobre retorno para o investidor. Mostrando que a relação entre risco e retorno é não linear. Há um retorno marginal decrescente ao risco; adicionar mais risco não gera um retorno igual. Em vez disso, cada unidade adicional de risco adiciona uma quantidade cada vez menor de retorno ao portfólio.

## 2.2 MODELO CAPM

Formulado por Sharpe (1964), Lintner (1965) e Mossin (1966), o modelo de precificação de ativos (do inglês, CAPM - Capital Asset Pricing Model), segundo Kopittke e Freitas (2001), é um modelo de precificação de ativos que estabelece uma relação entre o retorno esperado do ativo e a parcela de seu risco não diversificável.

Segundo Securato (1993), Sharpe, Lintner e Mossin perceberam as dificuldades de se estabelecer as covariâncias entre os retornos dos ativos que iriam compor as várias carteiras. A partir de tal raciocínio, iniciou-se a substituição das covariâncias pelo coeficiente de correlação linear (betas), o que originou o atual modelo CAPM, introduzindo a noção de risco sistemático (não-diversificável).

O modelo CAPM está fundamentado sobre a hipótese de mercado eficiente, ou seja, um mercado sem imperfeições, que possui as seguintes premissas:

- Todos os investidores são racionais,
- Os investidores são aversos ao risco,
- Nenhum investidor (tomador ou emprestador) consegue provocar oscilações nas taxas de mercado,
- Os retornos futuros são conhecidos e/ou previsíveis
- Não existem custos ou impostos sobre as transações
- As informações são livres, conhecidas e acessíveis a todos, sem custos
- Não há restrições aos investimentos;

Segundo Gitman (1997), usando o coeficiente beta para medir o risco não-diversificável, o modelo de formação de preços de ativos de capital (CAPM) é dado na equação:

$$E(r) = R_f + \beta_i(R_m - R_f) \quad (2.3)$$

Onde,

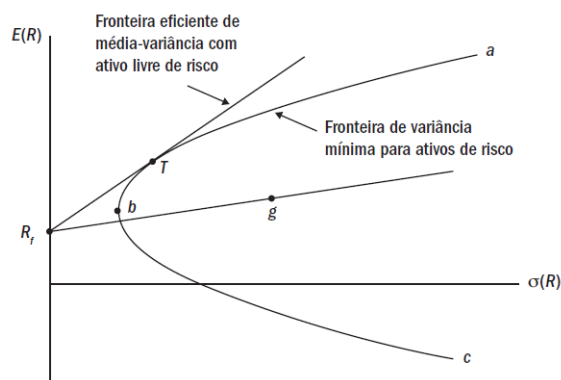
- $E(r)$  = Retorno exigido sobre o ativo j,

- $R_f$  = taxa de retorno livre de risco,
- $\beta_i$  = Coeficiente beta,
- $R_m$  = Retorno de mercado,

Observando o modelo, pode-se inferir que o retorno exigido para um determinado ativo em qualquer período de tempo é o valor da taxa livre de risco mais o resultado da operação envolvendo o coeficiente beta estimado multiplicado pelo resultado da subtração entre o retorno de mercado e o retorno do ativo livre de risco.

O modelo CAPM é representado graficamente na figura 1, onde o eixo X é percentual de risco e o eixo Y é o retorno esperado. A curva representa a fronteira eficiente e o ponto  $R_f$ , no eixo Y, é o valor da taxa livre de risco, no ponto 'T' é onde a curva da fronteira eficiente tangência a linha do ativo livre de risco, esse ponto é a carteira de melhor risco-retorno.

Figura 1. Modelo teórico da fronteira eficiente e da taxa livre de risco



Fonte: Fama e French (1993), traduzido para o português

### 2.3 ÍNDICE DE SHARPE

O Índice de Sharpe (1966) é uma métrica utilizada em finanças como uma medida da relação risco e retorno. Segundo Silva, Nogueira e Ribeiro (2015), o índice fornece o quanto de retorno adicional se tem em uma carteira para cada risco adicional. Dessa forma, o índice é a diferença entre o retorno total de uma carteira X e o retorno total do ativo livre de risco, dividido pelo desvio padrão da diferença entre o retorno total da carteira X e do retorno total do ativo livre de risco, aplicado a um período de tempo T. Sendo o ativo livre de risco normalmente representado por títulos atrelados à taxa básica de juros de um país como o Tesouro SELIC no Brasil ou *T-bonds* nos EUA, por exemplo.

A fórmula do índice de Sharpe é expressa como:

$$IS = \frac{R_a - R_b}{\sqrt{Var(R_a - R_b)}} \quad (2.4)$$

Onde:

- $IS$  = Índice de Sharpe,
- $R_a$  = Retorno da carteira,
- $R_b$  = Retorno do ativo livre de risco

O resultado da fórmula expressa é tratada como uma pontuação que cada carteira vai obter, dessa forma, torna-se possível comparar diversas carteiras e escolher aquela que possui o maior valor. A carteira com o maior índice de Sharpe é considerada a mais eficiente dentre as comparadas.

De acordo com Varga (2001), o índice de Sharpe se encaixa na teoria de seleção de carteira, mais especificamente no modelo CAPM, apontando as carteiras ótimas do modelo. "De acordo com o CAPM, nenhuma carteira pode ter um IS maior que o definido pela carteira de mercado. Carteiras com IS menor devem ser desprezadas." (Varga, 2001, p.229). Securato (1996, p. 206), por sua vez, define o índice como "a razão recompensa-variabilidade".

Para Milani e Ceretta (2012), Sharpe (1966) diferenciou-se dos demais estudos por levantar a discussão acerca da performance dos portfólio, enquanto estudos anteriores focavam apenas em métodos para a montagem de um portfólio eficiente.

## 2.4 MODELO DE TRÊS FATORES DE FAMA E FRENCH

O modelo de três fatores de Fama-French (1993) é um modelo estatístico projetado para descrever os retornos das ações.

O modelo mais tradicional de precificação de ativos, CAPM, elaborado por Sharpe (1964) com base no trabalho de Markowitz (1952) usa apenas uma variável para determinar o retorno de uma ação ou de um portfólio. Essa variável é o prêmio de mercado. Fama e French (1992), então, observaram que dois tipos de ações têm uma tendência a apresentar retornos superiores ao mercado: *smallcaps* e ações com um alto B/M (relação *book-to-market*, que compara o valor contábil da empresa com o preço atual de sua ação no mercado).

Segundo Fama e French (1993), os retornos dos ativos é explicado não somente pelos retornos de mercado, mas sim por outros dois fatores, conhecidos como os *Small minus Big* (SMB) e *High minus Low* (HML), que referem-se respectivamente aos fatores de tamanho da firma e "valor" precificado, no qual a relação valor contábil sobre valor de mercado é a *proxy*. Uma ação com alta relação contábil sobre mercado é considerada uma empresa de "valor", enquanto que no caso contrário se enquadram as chamadas empresas de "crescimento".

O modelo CAPM já trazia um fator essencial, a diferença entre o retorno médio da carteira e a taxa de juros livre de risco, também chamada de prêmio de mercado. Fama e French então chegam a conclusão que a união desses três fatores (Retorno de Mercado, SMB e HML) juntos deveriam ser significativos na explicação dos retornos dos ativos. Os autores utilizaram o método da carteira, que analisa o retorno de carteiras compostas por diferentes ações, ao invés dos retornos de cada ação individualmente.

Baseados nas anomalias identificadas em trabalhos anteriores (notadamente no trabalho de 1992 dos mesmos autores), Fama e French(1993) formularam um modelo de três fatores representado matematicamente pela expressão seguinte:

$$R_{ci,t} - R_{lrt} = a + b[R_{mt} - R_{lrt}] + s[SMB_t] + h[HML_t] + e_{i,t} \quad (2.5)$$

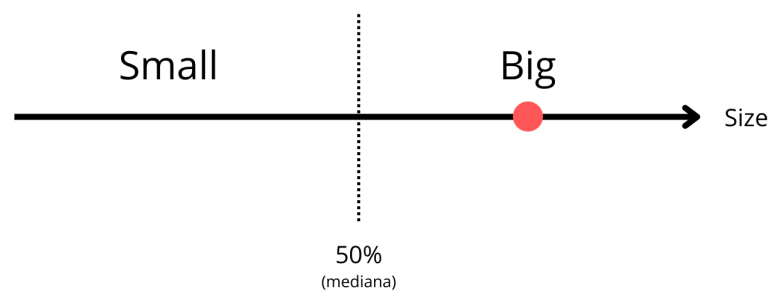
Onde,

- $R_{ci,t}$  = retorno da carteira i no mês t,
- $R_{mt} - R_{lrt}$  = prêmio da carteira de mercado no mês t,
- $SMB_t$  = prêmio pelo fator tamanho no mês t,
- $HML_t$  = prêmio pelo fator B/M no mês t,
- $e_{i,t}$  = resíduo do modelo referente a carteira i no mês t

#### 2.4.1 Explicação dos fatores SMB e HML

Fama e French (1993) explicam que para calcular o fator SMB o primeiro passo é dividir as ações pela mediana em dois grupos, denominados *Small* e *Big*. A mediana é utilizada no modelo porque ela é o valor que separa a metade maior e a metade menor de uma amostra, uma população ou uma distribuição de probabilidade.

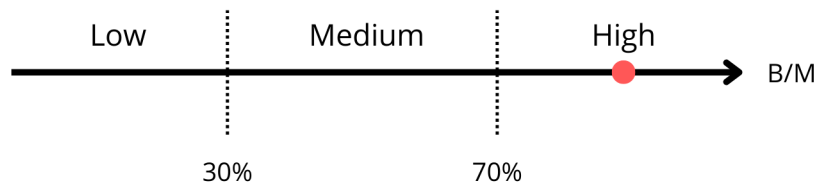
Figura 2. Fator SMB



Fonte: Elaboração própria.

Para obter o fator HMB, Fama e French (1993) dividem os ativos em três grupos, *Low*, *Medium* and *High* (L, M e H), onde o grupo *Low* é composto pelos ativos que possuem uma baixa razão B/M, esse grupo vai ser composto por cerca de 30% da amostra. O segundo grupo é o *Medium*, esse grupo é onde estão os ativos enquadrados na faixa intermediária da amostra, aqueles cuja razão B/M não é suficientemente grande para ser enquadrada no grupo *High*, nem tão pequena para ser enquadrado no grupo *Low*, o grupo *Medium* é composto por 40% da amostra. Por fim, o último grupo é o *High*, esse grupo é composto pelos ativos de maior razão B/M e é composto por 30% da amostra.

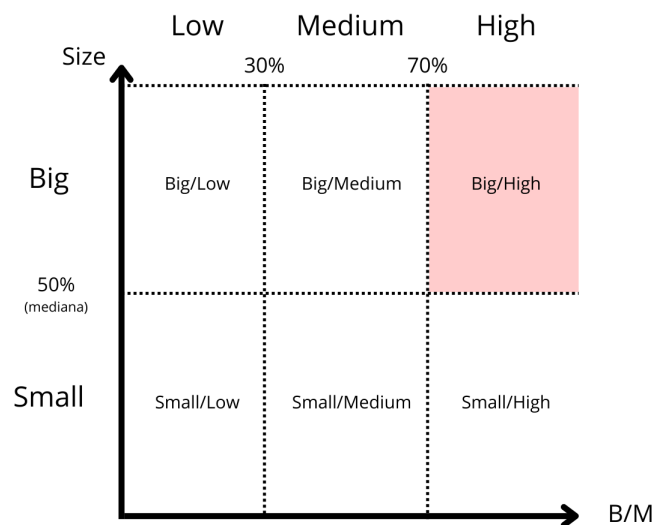
Figura 3. Fator HML



Fonte: Elaboração própria.

Através da união entre os fatores SMB e HML as carteiras são construídas: S/L, S/M, S/H, B/L, B/M e B/H, ou seja, Pequena/Baixa, Pequena/Média, Pequena/Alta, Grande/Baixa, Grande/Média e Grande/Alto, como é visto na figura 4. Com essas seis carteiras, derivam-se as variáveis explicativas, *small minus big* (SMB) e *high minus low* (HML).

Figura 4. Seis carteiras geradas através dos fatores SMB e HML



Fonte: Elaboração própria.

O retorno de mercado é representado pelo excesso de retorno, que é o retorno menos a taxa livre de risco (método CAPM), das seis carteiras de tamanho B/M mencionadas acima. O risco de tamanho é representado pela carteira *small minus big* (SMB) que é a diferença mensal entre o retorno médio das três carteiras pequenas (S/L, S/M, S/H) e o retorno médio das três grandes carteiras (B/L, B/M, B/H).

Dessa forma, o fator SMB pode ser representado algebricamente como:

$$SMB = \frac{(R_{Small/Low} + R_{Small/Medium} + R_{Small/High})}{3} - \frac{(R_{Big/Low} + R_{Big/Medium} + R_{Big/High})}{3} \quad (2.6)$$

O risco de B/M é representado pela carteira *high minus low* (HML) que é a diferença mensal entre o retorno médio das duas carteiras de B/M alto (S/H, B/H) e o retorno médio das duas carteiras B/M baixas (S/L, B/L). Observa-se que os dois portfólios médios S/M e B/M não estão incluídos no portfólio HML, a razão para isso é que Fama e French não estavam interessados nos ativos cuja razão B/M poderiam ser enquadradas como estáveis. O fator HML é feito, portanto, através da análise entre ativos com razão B/M discrepantes.

Algebricamente, o fator HML é representado como:

$$HML = \frac{(R_{Small/High} + R_{Big/High})}{2} - \frac{(R_{Small/Low} + R_{Big/Low})}{2} \quad (2.7)$$



### 3 BASE DE DADOS

#### 3.1 EXTRAÇÃO DOS DADOS

Os dados utilizados foram extraídos do Yahoo! Finance em janeiro de 2022. A extração foi feita através da linguagem de programação R e suas bibliotecas. O principal motivo para utilização do Yahoo! Finance como fonte de dados é pela sua facilidade de comunicação com as linguagens de programação R e Python, através de bibliotecas dedicadas, como a *getsymbols*, do R e a *yfinance*, do Python, outro fator que contribui para sua utilização é pela extensa gama de dados, sendo possível obter a cotação histórica de quase todos os ativos negociados em bolsa no mercado brasileiro desde 2002. O resultado da coleta dos dados foi salvo em arquivos *csv*'s (arquivos separados por vírgula) para posterior análise.

Através da tabela 1 é possível observar as características iniciais do banco de dados. Existem 8 colunas, sendo a coluna de nome *symbol* referente ao ativo negociado, a coluna *date* refere-se a data da negociação na Bolsa de Valores de São Paulo (BOVESPA), a coluna *open* refere-se ao preço de abertura do ativo na data de referência, a coluna *high* trata do preço mais alto atingido pelo ativo durante a data de negociação, a coluna *low* trata do preço mais baixo atingido pelo ativo durante a data de negociação, *close* refere-se ao preço de fechamento na data de referência, *adjusted* é o preço de fechamento (coluna *close*) ajustado para refletir um eventual desdobramento de ações (quando a companhia acrescenta ou retira de circulação determinada quantidade de ações). Por fim, a coluna *volume* mede a quantidade de ações de um ativo que são negociadas na data de referência.

Tabela 1. Visualização parcial do banco de dados

symbol	date	open	high	low	close	adjusted	volume
ATOM3.SA	2011-01-03	3,19	3,19	2,99	3,09	2,99	9663
CPL6.SA	2011-01-03	4,20	4,20	4,13	4,17	2,03	4047000
BRGE11.SA	2011-01-03	6,14	6,14	6,14	6,14	3,15	0
EZTC3.SA	2011-01-03	10,26	10,27	10,14	10,22	6,29	311060
HETA3.SA	2011-01-03	120,00	120,00	120,00	120,00	120,00	0
FRAS3.SA	2011-01-03	3,65	3,65	3,65	3,65	2,51	0
TKNO3.SA	2011-01-03	71,59	71,59	71,59	71,59	47,89	0
PAT3.SA	2011-01-03	11,78	11,78	11,78	11,78	9,31	3230
BGIP4.SA	2011-01-03	19,48	20,32	19,45	20,32	9,30	580
INEP3.SA	2011-01-03	5,61	5,87	5,61	5,83	5,71	30900

Fonte: Yahoo! Finance. Elaboração própria.

#### 3.2 TRATAMENTO DOS DADOS

Os dados foram analisados por meio das linguagens de programação R e Python, por meio de suas bibliotecas (códigos disponíveis em: <https://github.com/CleitonOERocha/tcc-finance-codes>).

Tabela 2. Bibliotecas R e Python Utilizadas

R	Python
Tidyverse	Pandas
Lubridate	Numpy
PerformanceAnalytics	yfinance
quantmod	datetime
fPortfolio	
tseries	
timetk	
xts	

Fonte: Elaboração própria.

Na tabela 2 são listadas as bibliotecas utilizadas e que serviram de base para fazer todo o tratamento dos dados, como a remoção de diversos valores ausentes nas colunas de preços, geração das carteiras em ambos os modelos, geração de gráficos, remoção de valores distorcidos aplicando o método de Tukey (1977) para detecção e remoção de *outliers* (Ex.: o ativo MMAQ3 possuía um valor ajustado variando de R\$ -47.490.514.248.794.112 até R\$ 300.201.790.035.460.096 em boa parte do ano de 2011), gerar as estatísticas descritivas do banco de dados, gerar as tabelas no formato Latex, dentre outras atividades menores.

O banco de dados possui informações de preços de abertura, fechamento, preço de fechamento ajustado para a data presente e volume de negociações, totalizando assim inicialmente 1.107.244 observações. Foram coletadas informações de 504 ativos que abrangem um período começando em 03/01/2011 e que se estende até 30/12/2021. Os dados são diários, ignorando-se finais de semana e feriados nacionais.

Para obter as carteiras em ambos os modelos (Fama e Markowitz) não será utilizado o banco de dados em sua completude, isso porque, para que se sirva de insumo para testar a qualidade das carteiras e poder compará-las é necessário um período de tempo não utilizado na aplicação dos modelos. Dessa forma, o banco de dados foi dividido em dois: baseline e teste. Onde, o baseline será utilizado na construção das carteiras e o banco de dados de teste será utilizado para observar o comportamento de ambos os portfólios uma vez criados. A divisão em baseline e teste foi aplicada na proporção 70% baseline e 30% para teste. Com isso, os dados do baseline indo até 10/09/2018. Totalizando 735.667 observações, o que representa 1.986 dias.

O banco de dados possui um grande número de microcaps (ações de empresas com valor de mercado menor ou igual a R\$1,5 bilhões), os valores de negociações de algumas dessas empresas provocavam distorções na modelagem dos dados, isso ocorre porque, alguns desses ativos apresentavam valores fora do comum (anomalias), com preços de fechamento na ordem de bilhões de reais. Por conta dessas distorções, fez-se necessário inserir um parâmetro para eliminar as ações dessas empresas muito pequenas e pouco negociadas. A abordagem utilizada foi filtrar os ativos pelo seu volume de negociação diário, o valor adotado como base foi uma

média de R\$ 1.000,00. Ou seja, permaneceram na base de dados apenas os ativos que durante todo o período analisado tiveram uma média de negociação diária de R\$ 1.000,00 ou mais. Essa média utilizada serviu para eliminar os ativos mais inconsistentes, mas ainda sim deixar ativos que apesar de serem pouco negociados, apresentam certa regularidade nos preços.

Após esse filtro, o número de ativos elegíveis para a análise foi de 345. Uma redução de 31,5% em relação aos dados originais, sem tratamento.

Para preservar a diversificação das carteiras - e por conta de muitos ativos do mercado acionário brasileiro terem suas ações divididas em ordinárias (ON), que são as ações que dão direito a voto nas decisões da empresa, e preferenciais (PN), ações sem poder de voto, mas com preferência no recebimento de dividendos - nos casos em que as carteiras geradas tinham mais de um ativo da mesma empresa (ex. alocação em PETR4 (PN) e PETR3 (ON)), foi removido da carteira o ativo de menor retorno da empresa (ex. PETR4 selecionada e PETR3 descartada) e substituído por outro ativo escolhido pelo modelo. Com isso, o número final de ativos elegíveis foi de 192.

Por fim, a variável utilizada para representar o preço dos ativos foi o preço de fechamento ajustado para a data presente (a coluna no banco de dados de nome *adjusted*), a escolha reflete a necessidade de levar em conta o desdobramento de algumas ações ao longo do período em que a análise é feita. Visto que o desdobramento não sinaliza uma queda real no preço da ação, mas apenas que mais ações estão em circulação.

A tabela 3 contém algumas estatísticas descritivas de acordo com as colunas de preço (*open*, *high*, *low*, *close*) e volume dos ativos, como a contagem de observações, média, desvio padrão e dispersão dos dados de acordo com o quartil depois da aplicação do tratamento dos dados.

Após o efeito de correção a média do preço de fechamento ajustado fica em R\$48,55 e o valor mínimo em R\$ 0,03. O número de observações no banco de dados fica em 638.416.

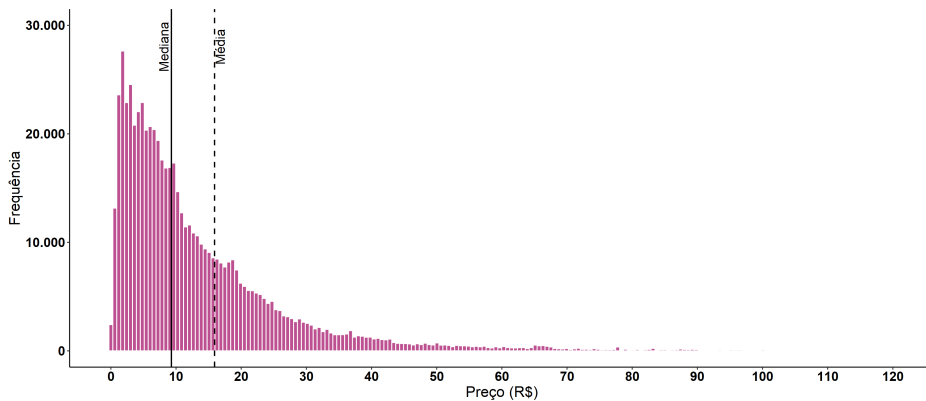
Tabela 3. Estatísticas descritivas das colunas do banco de dados

	open	high	low	close	adjusted	volume
count	638416,00	638416,00	638416,00	638416,00	638416,00	638416,00
mean	55,77	56,82	54,66	55,73	48,55	2744144,11
std	609,21	616,01	601,80	609,16	508,96	9917774,07
min	0,03	0,03	0,03	0,03	0,03	0,00
25%	5,30	5,40	5,19	5,30	4,60	4446,75
50%	11,60	11,80	11,37	11,59	9,60	198300,00
75%	22,29	22,67	21,90	22,27	18,96	1792600,00
max	49200,49	49500,50	45600,46	49200,49	44594,76	939592800,00

Fonte: Yahoo! Finance. Elaboração própria.

Observando o histograma na figura 5, observa-se que 50% (mediana) dos ativos tem um preço abaixo de R\$ 9,60. Com uma tendência decrescente na quantidade de ativos na medida em que o preço aumenta.

Figura 5. Histograma do preço de fechamento ajustado dos ativos



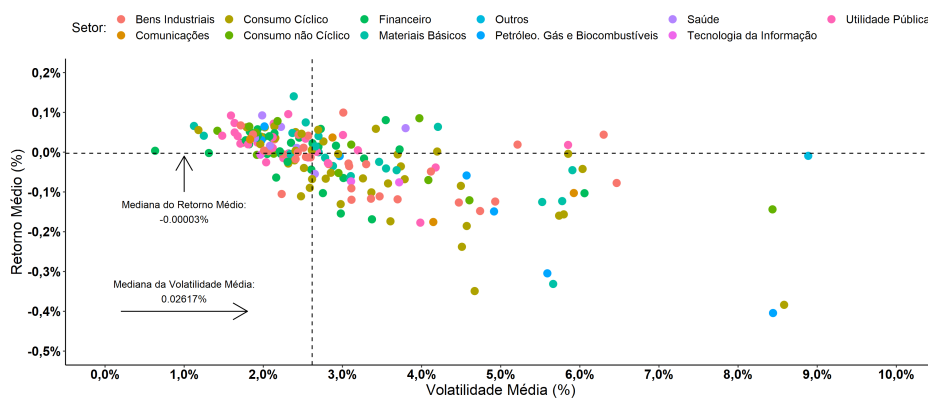
Fonte: Yahoo! Finance. Elaboração própria.

### 3.3 CONSTRUÇÃO DO MODELO DE MARKOWITZ

Para construir a carteira mais eficiente seguindo os pressupostos estabelecidos por Markowitz, é necessário ter a cotação diária de diversos ativos ao longo de um determinado período, com a cotação é possível estimar o retorno e a volatilidade desse ativo ao longo de determinado período. O período estabelecido nesse presente artigo começa na primeira semana de janeiro de 2011 e se encerra na última semana de dezembro de 2021.

A figura 6 mostra a relação entre volatilidade média diária e o retorno médio diário para todos os ativos no período analisado.

Figura 6. Risco-Retorno das Ações



Fonte: Yahoo! Finance. Elaboração própria.

Na tabela 4, os ativos foram agrupados por setor, a partir disso foi traçado o retorno médio e a volatilidade de cada grupo. O setor de Consumo Cíclico é o que possui o maior número de ativos, representando aproximadamente 23% do total. Esse é um setor que abrange empresas que

Tabela 4. Média diária do retorno e volatilidade, por setor do ativo

Setor	Total de Empresas	Retorno Médio Diário	Volatilidade Média Diária
Consumo Cíclico	44	-0,0472%	3,28%
Financeiro	35	-0,0021%	2,46%
Bens Industriais	32	-0,0392%	3,45%
Utilidade Pública	24	0,0157%	2,48%
Materiais Básicos	21	-0,0204%	3,32%
Consumo não Cíclico	13	-0,0019%	3,1%
Saúde	8	0,0335%	2,37%
Petróleo. Gás e Biocombustíveis	7	-0,1188%	4,35%
Comunicações	4	-0,0524%	3,69%
Tecnologia da Informação	3	-0,0519%	2,93%
Outros	1	-0,0093%	8,89%

Fonte: Fonte: Yahoo! Finance. Elaboração própria.

trabalham em ramos de atividades que apresentam desempenho conforme o ciclo econômico. Isso ocorre porque as empresas de consumo cíclicos são aquelas que mais se beneficiam com o crescimento econômico, visto que importam insumos baratos e tendem a ter aumento nas vendas, o que permite ter uma margem de lucro maior. São exemplos de empresas desse segmento a Cyrela (atua na construção cívil), Lojas Renner (vestuário), Magazine Luiza (venda de eletrodomésticos), dentre outras.

Para além da necessidade dos preços dos ativos, a construção do modelo CAPM envolve mais um pressuposto: a taxa de livre de risco. Não existe uma taxa livre de risco oficial, ao invés disso, o que existe são várias taxas já existentes que são utilizadas como *proxy* para a taxa estabelecida na literatura. No Brasil, por exemplo, é comum utilizar a taxa básica de juros da economia (denominada Selic) ou a Taxa DI como referencial. Contudo, é normal encontrar outros estudos em que os autores trazem uma abordagem diferente para a taxa livre. Barros, Famá e Silveira (2003), por exemplo, trazem o retorno mensal da caderneta de poupança como uma *proxy* para a taxa de retorno livre de risco. Para o presente artigo, será utilizada a taxa DI como referencial para o ativo livre de risco.

Para obter a taxa DI acumulada durante o período analisado, foi utilizado a calculadora elaborada pela B3<sup>1</sup>. Por meio dessa calculadora, é possível informar a data de início (03/01/2011) e fim (10/09/2018) desejada e obter a taxa acumulada durante esse período. O que representou uma taxa de 14,65% durante esse período. Dividindo esse valor pela quantidade de dias (1.986), chega-se ao número final: 0,058%. Esse número representa o retorno diário durante todo o período analisado, as carteiras elaboradas pelo modelo de Markowitz precisam superar esse *benchmark* para se tornarem elegíveis<sup>2</sup>.

Por fim, as carteiras são geradas através de milhares de simulações que envolvem a escolha de N

<sup>1</sup> <https://calculadorarendafixa.com.br/navbar/calculadora>

<sup>2</sup> Carteira elegível: carteira cujo retorno em determinado período de tempo é maior que o retorno da taxa livre de risco no mesmo período.

ativos com percentuais diferentes dentro do portfólio. As carteiras são então testadas com os dados de treino.

### 3.4 CONSTRUÇÃO DOS FATORES DE FAMA E FRENCH

Para se testar o modelo Fama e French, primeiro é necessário estimar o prêmio pelos riscos decorrentes dos fatores Mercado, Tamanho e índice B/M. A metodologia utilizada foi a mesma empregada por Fama e French (1993) na construção do modelo de três fatores. A saber:

- Em junho de cada ano  $t$ , começando em 2011 e terminando em 2021, todas as ações da amostra são ordenadas de acordo com o valor de mercado de junho das empresas que representam. O valor mediano é utilizado para dividir a amostra em dois grupos, classificados como B (*Big*) e S (*Small*), que contém as empresas de maior e menor valor de mercado respectivamente. Também neste mesmo mês, todas as ações amostrais são ordenadas de acordo com o índice B/M das empresas que representam. Este índice é calculado com valores contábeis e de mercado do patrimônio líquido referentes a dezembro do ano anterior ( $t-1$ ) ao de formação das carteiras. Após a ordenação, a amostra é segregada em três grupos: 30% inferior (*Low*), 40% médio (*Medium*) e 30% superior (*High*) de acordo com o valor do índice B/M. Em junho de cada ano  $t$ , após as duas ordenações anteriores, constroem-se seis carteiras, decorrentes da intersecção dos dois grupos ordenados de acordo com a variável Valor de Mercado e dos três grupos ordenados de acordo com a variável B/M. Estes seis grupos são:
  - *Small and low* (S/L): ações com baixo valor de mercado e baixo índice B/M,
  - *Small and medium* (S/M): ações com baixo valor de mercado e médio índice B/M,
  - *Small and high* (S/H): ações com baixo valor de mercado e alto índice B/M,
  - *Big and low* (B/L): ações com alto valor de mercado e baixo índice B/M,
  - *Big and medium* (B/M): ações com alto valor de mercado e médio índice B/M,
  - *Big and high* (B/H): ações com alto valor de mercado e alto índice B/M.
- De julho do ano  $t$  a junho do ano  $t+1$ , calcula-se o retorno mensal real de cada ação.
- De julho do ano  $t$  a junho do ano  $t+1$ , calcula-se o retorno real mensal de cada uma das seis carteiras através da ponderação, pelo valor de mercado da ação em relação ao valor de mercado da carteira, dos retornos das ações que as compõem.
- Mensalmente, calcula-se o prêmio do fator de risco tamanho (SMB) pela diferença entre a média dos retornos mensais das três carteiras S descritas anteriormente e a média dos retornos mensais das três carteiras B.

- Mensalmente, também se determina o prêmio pelo fator de risco B/M (HML), através da diferença entre a média dos retornos mensais das duas carteiras H e a média dos retornos mensais das duas carteiras L.
- Finalmente, todo mês calcula-se o prêmio pelo fator de risco mercado, subtraindo-se do retorno da carteira de mercado a taxa livre de risco. O retorno da carteira de mercado é encontrado através da ponderação, pelo valor, do retorno das ações amostrais

Estes três prêmios mensais serão utilizados como variáveis explicativas nas regressões temporais para validação do modelo e teste de hipóteses. Finalmente, todas as carteiras são reformuladas em junho de cada ano, repetindo-se os procedimentos descritos.

## 4 RESULTADOS

### 4.1 MODELO FAMA-FRENCH

Com o objetivo demonstrar o comportamento dos fatores ao longo do período analisado, foram adicionados gráficos do retorno acumulado de cada fator partindo de junho/2011 até dezembro de 2021, compreendendo todo o período da amostra. Com isso, pode-se observar quais fatores impactaram positivamente e negativamente no modelo final de Fama-French.

O primeiro fator observado é o fator de mercado do CAPM, que se configura como a diferença entre o retorno médio da carteira de mercado e a taxa de juros livre de risco. Neste trabalho o retorno do índice Ibovespa foi adotado como o retorno de mercado e a taxa CDI como taxa de juros livre de risco.

Figura 7. Fator de Mercado - Fama e French



Fonte: Yahoo! Finance. Elaboração própria.

O fato de no início da amostra a taxa DI ser elevada quando comparada aos últimos anos da análise, bem a crise interna que atingiu o país em 2014 e que se seguiu por pouco mais de dois anos e posteriormente a pandemia da Covid-19, que assolou os mercados internacionais e o Brasil impactaram consideravelmente o desempenho deste fator. Encerrando o período com um retorno negativo de aproximadamente 57%, como é visto na figura 7.

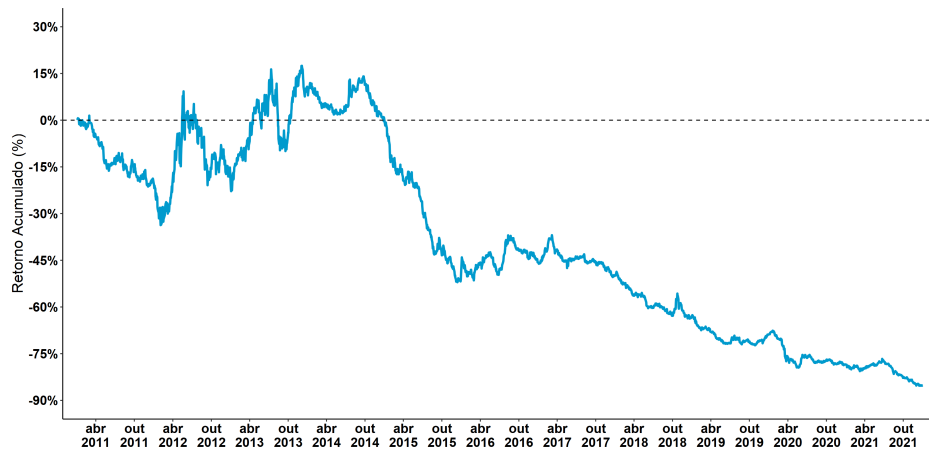
Seja pelo baixo desempenho do índice Ibovespa ou pela elevada taxa DI durante o período analisado, ambos bastante influenciados por fatores externos, não foi observado prêmio de risco pelo investimento em ações.

Na figura 8, partindo de junho/2011, observa-se que o fator tamanho teve desempenho muito inferior ao fator mercado, bem como os demais fatores que serão apresentados a seguir. Isto



demonstra que as ações de empresas classificadas como *Small* apresentaram desempenho muito inferior as ações das *Big*.

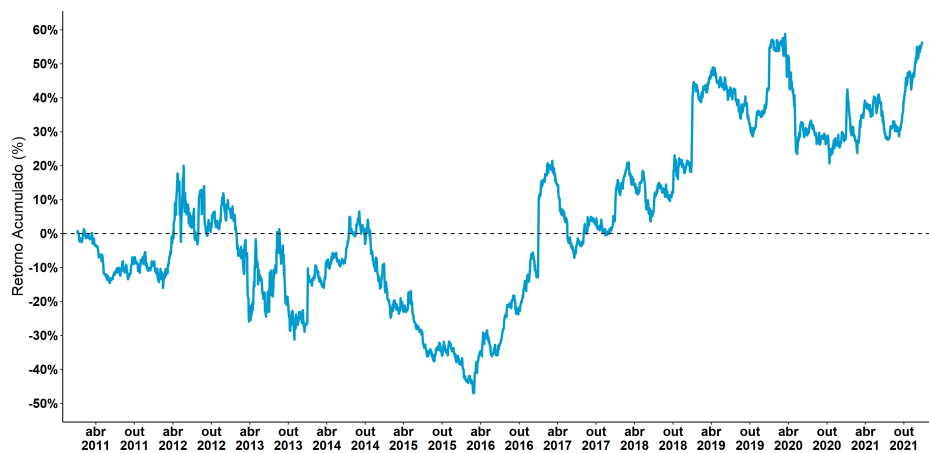
Figura 8. Fator SMB - Fama e French



Fonte: Yahoo! Finance. Elaboração própria.

Observa-se na figura 9, partindo de junho/2011, o diferencial de desempenho entre as empresas com alto índice B/M menos o desempenho das empresas com baixo índice B/M. No final do período, o indicador encerrou com um retorno positivo de aproximadamente 55%, demonstrando um rendimento muito superior para as empresa com elevado índice B/M.

Figura 9. Fator HML - Fama e French



Fonte: Elaboração própria do autor com dados do Yahoo! Finance

## 4.2 MODELO CAPM

Para entender como chegar a um portfólio final elaborado segundo o modelo CAPM, é preciso entender como a alocação é feita segundo os preceitos de retorno e volatilidade capítulo 2 - Referencial Teórico; Onde o melhor ativo é aquele que se sobressai em relação ao benchmark (taxa livre de risco, aqui sendo a Taxa DI) e que possui um volatilidade adequada ao retorno no período analisado.

Para fins de simulação de carteiras segundo o CAPM, este trabalho estabelece dois tipos de alocação: 1) com todos os ativos listados após o tratamento dos dados, como explicado no tópico 3.2 - Tratamento dos dados e 2) apenas com os ativos caracterizados como "*smallcaps*".

Na figura 10 é possível observar o comportamento do retorno diário de todos os ativos, salvo exceções, boa parte dos ativos possui retorno diário próximo a zero.

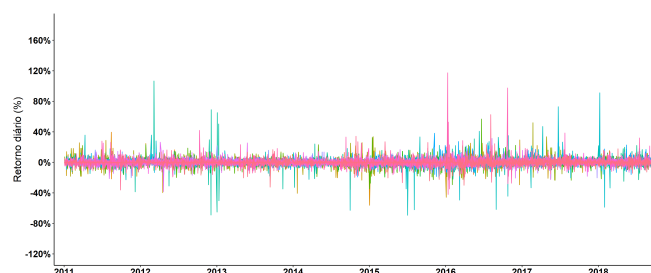
Figura 10. Retorno diário - todos os ativos



Fonte: Yahoo! Finance. Elaboração própria.

A figura 11 contém o retorno diário apenas das *smallcaps*, observa-se um comportamento de retorno similar ao da figura 10, com o comportamento do retorno diário de todos os ativos, salvo exceções, boa parte dos ativos possui retorno diário próximo a zero.

Figura 11. Retorno diário - *SmallCaps*

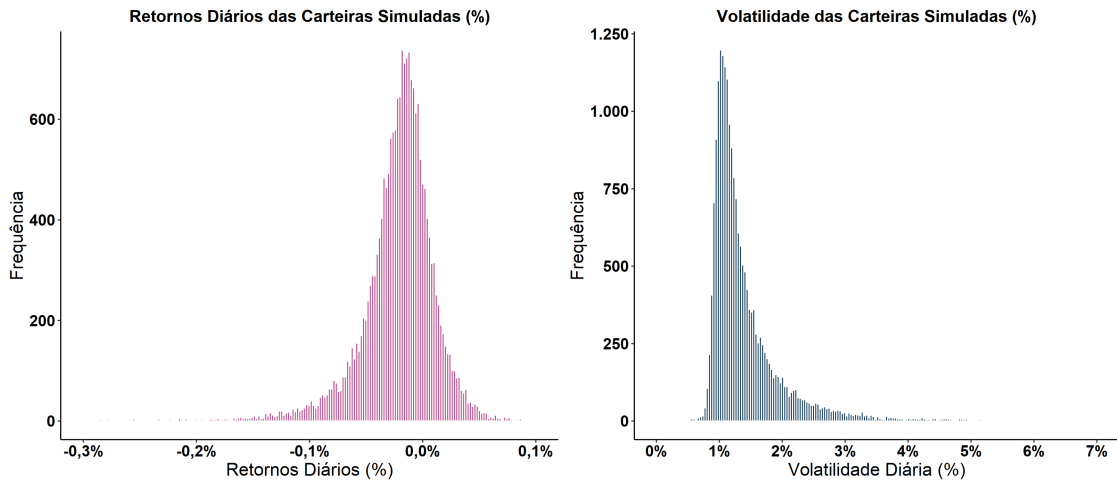


Fonte: Elaboração própria do autor com dados do Yahoo! Finance

Milhares de composições de carteiras com ativos e percentuais alocativos diferentes são simuladas. Na figura 12 é possível observar a distribuição dos retornos e a volatilidade diária dessas carteiras.

90% das carteiras simuladas possuem um retorno diário na faixa de  $-0,01\%$  até  $0,01\%$  e volatilidade diária de até  $4\%$ .

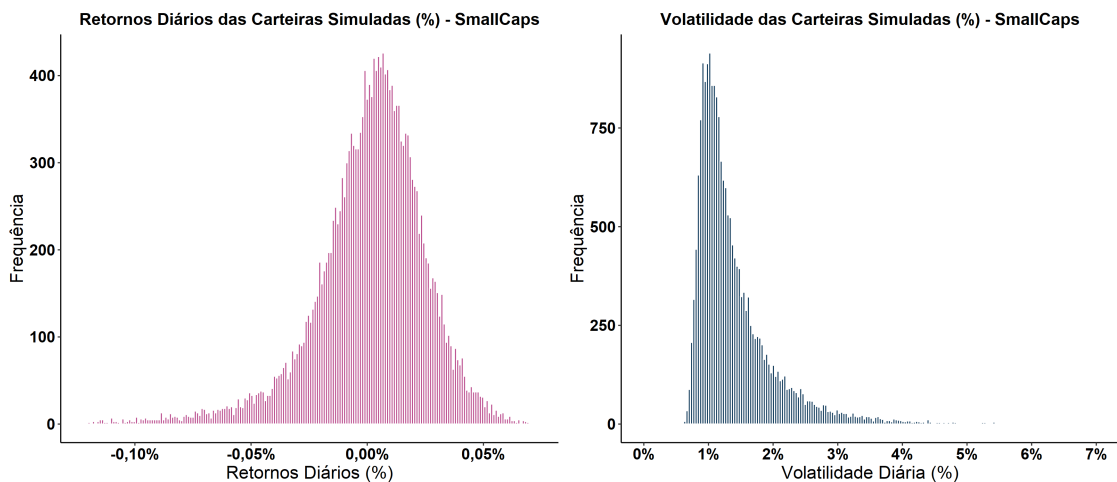
Figura 12. Retorno e Volatilidade das Carteiras Simuladas



Fonte: Yahoo! Finance. Elaboração própria.

Na figura 13 é possível observar as simulações de carteiras para os ativos *smallcaps*. 90% da distribuição dos retornos das carteiras desses ativos estão na faixa de  $0,05\%$  até  $0,05\%$ . Se tratando da volatilidade diária, 90% se encontra num intervalo de até  $4\%$ .

Figura 13. Retorno e Volatilidade das Carteiras Simuladas - *SmallCaps*

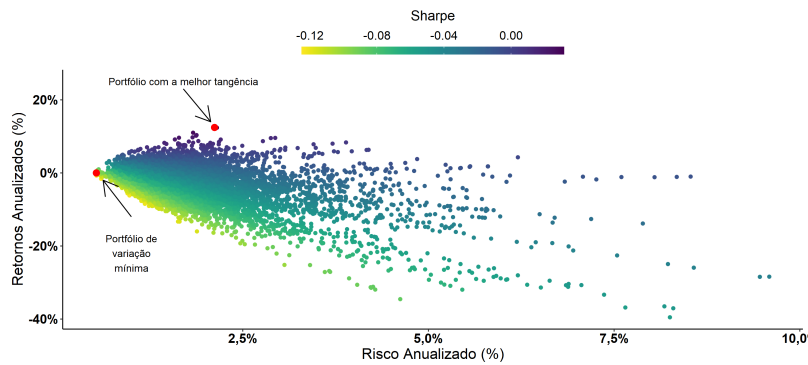


Fonte: Yahoo! Finance. Elaboração própria.

A partir dos dados de retorno e risco das carteira simuladas, foi possível otimizar os resultados, obtendo assim, a fronteira eficiente proposta por Markowitz (1952). Como é visto na figura 14. A carteira selecionada então é aquele que dentre todas presentes na fronteira eficiente, possui o

melhor índice Sharpe (1964). A carteira de melhor Sharpe forneceu um retorno anualizado de aproximadamente 10% e um risco (volatilidade) anualizada de 2,4%.

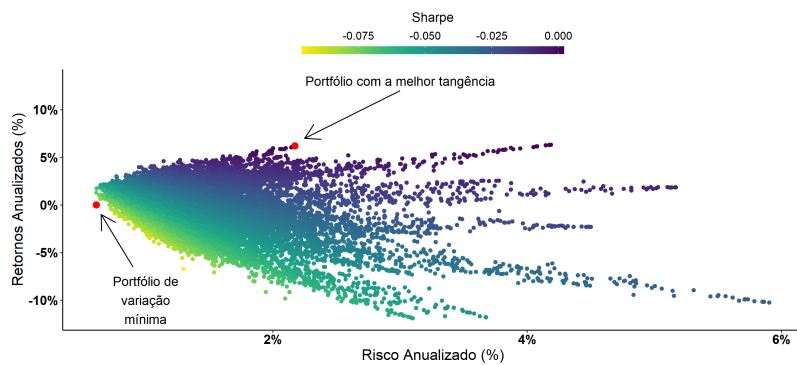
Figura 14. Fronteira Eficiente



Fonte: Yahoo! Finance. Elaboração própria.

Na figura 15, temos a carteira selecionada dentre os ativos *smallcaps*. A carteira de melhor Sharpe forneceu um retorno anualizado de aproximadamente 5% e um risco (volatilidade) anualizada de 2,1%.

Figura 15. Fronteira Eficiente - *SmallCaps*

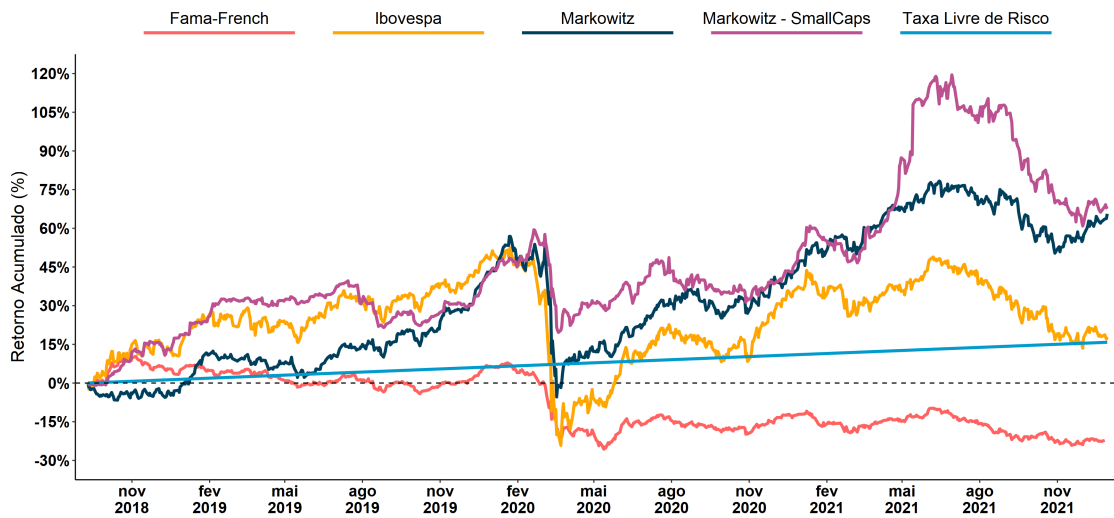


Fonte: Yahoo! Finance. Elaboração própria.

#### 4.3 RETORNO DAS CARTEIRAS SIMULADAS

O resultado das simulações das carteiras geradas através do modelo Fama-French e CAPM com dados do baseline indo até 10/09/2018 foram aplicadas nos dados de teste, que abrange o período de 11/09/2018 até 30/12/2021. O resultado final é observado na figura 16. Onde é possível ver o retorno acumulado de cada carteiras, assim como da Taxa Livre de Risco (Taxa DI) e do Ibovespa no período.

Figura 16. Retorno das Carteiras Simuladas



Fonte: Yahoo! Finance. Elaboração própria.

Nota-se num primeiro momento uma performance melhor do índice Ibovespa e da carteira de *smallcaps*. A carteira de Fama-French fica com um desempenho próximo ao da taxa livre de risco. Com a pandemia da Covid-19, os investidores temendo o impacto no mercado financeiro, acabam por retirar seus investimentos dos ativos de renda variável, o que provoca uma queda abrupta de todas as carteiras simuladas, exceto pela taxa livre de risco. O efeito é maior no índice Ibovespa e na carteira de Markowitz simulada com todos os ativos. Nesse mesmo período, a carteira composta apenas por *smallcaps* teve uma menor queda em comparação com as outras carteiras.

O período entre maio e agosto de 2020 é marcado pela recuperação das carteiras simuladas, exceto pela carteira Fama-French, que permanece a área de retorno negativo.

Entre novembro de 2020 até maio de 2021, o que se observa é um desempenho muito similar entre as carteiras criadas através dos preceitos do modelo CAPM. O retorno acumulado dessas carteiras ao final de maio de 2021 era de 45% em relação ao momento inicial.

Entre junho até agosto de 2021, a carteira de *smallcaps* se desloca positivamente, com um desempenho muito superior as demais. Todavia, ao final de agosto a carteira sofre uma queda no desempenho, chegando em dezembro com um desempenho muito similar a carteira de Markowitz com todos os ativos.

Ao final de dezembro de 2021, o retrato final é que o desempenho das carteiras geradas pelo modelo CAPM (Markowitz e Markowitz - *SmallCaps*) se sobressaem em relação as demais carteiras. O índice Ibovespa, que representa o mercado, e a Taxa Livre de Risco, que representa a alocação mais segura possível, chegam ao final do período com um desempenho muito similar,

todavia, o Ibovespa apresentou muito mais volatilidade na comparação. Por fim, o modelo Fama-French teve o pior desempenho dentre todos os ativos analisados, apresentando um retorno acumulado negativo no período.

O resultado encontrado aqui evidencia a hipótese de que, para o mercado brasileiro, o modelo Fama-French produz resultados inferiores ao modelo CAPM. Em outros trabalhos, por exemplo, em Chague (2007), o autor encontrou evidências que o sucesso do modelo multifatorial (Fama-French) sobre o CAPM nos dados dos EUA não coincide com os dados do Brasil. O autor usa uma abordagem econométrica para mensurar a capacidade dos modelos em explicar o excesso de retorno dos portfólios de mercado no período entre 1999 até 2006.

#### 4.4 LIMITAÇÕES

Assim como em outros trabalhos do mesmo tema, fatores oriundos da obtenção de dados, que vão desde a ausência de muitas fontes gratuitas e confiáveis, baixo número de ativos líquidos e a falta de um histórico consistente para a análise dificultam a replicação do modelo Fama-French no mercado brasileiro.

O tratamento nos dados tenta mitigar os efeitos da ausência e da dificuldade em se obter dados mais abrangentes e precisos, em especial em relação aos dados de tamanho das empresas, todavia, esse é um fator que mesmo mitigado ainda impacta no resultado final da projeção do cenário estabelecido.

#### 4.5 CONTINUIDADE DE PESQUISA

A partir dos resultados apresentados aqui, hipóteses podem surgir sobre os retornos abaixo do esperado do modelo Fama-French para o mercado brasileiro, visto que, observando o mercado americano, o modelo traz resultados satisfatórios. Alguns pontos que podem ser levados adiante em busca de respostas podem partir de uma análise mais robusta do fator SMB (*Small minus Big*). A quantidade de empresas do setor de tecnologia - que nos últimos anos tiveram uma forte expansão e um rápido crescimento - no mercado americano faz com que o fator SMB traga mais retorno do que na comparação com o mercado brasileiro, que possui menos empresas da área tecnológica e tem suas atenções voltadas para empresas bem estabelecidas como Vale, Petrobras, Itaú? Algumas empresas brasileiras, como Nubank e XP, recentemente optaram por abrir capital no mercado americano, em especial, na NASDAQ. Isso de algum modo impacta nos retornos do modelo Fama-French no Brasil? São hipóteses que podem ser testadas em outras linhas de pesquisas.

## 5 CONSIDERAÇÕES FINAIS

A teoria moderna do portfólio postula que sob as devidas condições de equilíbrio, os retornos de um ativo podem ser descritos por uma relação linear ao seu risco. A partir dessa visão, o modelo CAPM surge como uma forma de se calcular esses retornos. Todavia, o mercado financeiro traz consigo anomalias ou "falhas de mercado", isso porque, em última instância, o preço de um ativo é formado pela crença/temor dos investidores no futuro desse ativo.

Novos modelos alternativos ao CAPM foram desenvolvidos para tentar capturar e explicar essas anomalias, entre eles, um dos modelos mais destacados é o modelo de três fatores de Fama e French (1993).

Este trabalho procurou avaliar o desempenho de ambos os modelos com dados do mercado financeiro brasileiro para observar se existia vantagem em usar o modelo de Fama e French ao invés do modelo CAPM para obter maiores retornos no mercado brasileiro.

As conclusões finais foram: 1) as carteiras geradas pelo modelo CAPM mostrou um desempenho superior as demais no período analisado; 2) de maneira geral, a carteira gerada seguindo os três fatores propostas por Fama e French (1993) não obteve resultados significativos ao longo do período analisado.

O modelo CAPM, apesar das suas limitações, foi superior ao modelo de três fatores na obtenção de melhores retornos dado o nível de risco.

## REFERÊNCIAS

- ASSAF NETO, Alexandre. **Mercado Financeiro**. 7. ed. São Paulo: Atlas, 2006.
- CHAGUE, Fernando Daniel. **The CAPM and Fama-French models: a comparative study**. 2007. 44 f. Dissertação (Mestrado em Economia) - Escola de Economia de São Paulo (EESP), Fundação Getúlio Vargas (FGV), São Paulo, 2007.
- FAMA, Eugene; FRENCH, Kenneth. The cross-section of expected stock returns. **The Journal of Finance**, v. 47, n. 2, p. 427-465, 1992.
- FAMA, Eugene; FRENCH, Kenneth. Common risk factors in the returns on stocks and bonds. **Journal of Financial Economics**, v. 33, n. 1, p. 3-56, 1993.
- KOPITTKÉ, B. H.; FREITAS, S. C. Considerações acerca do Capital Asset Pricing Model (CAPM) e sua utilização nos dias atuais. *In*: ENCONTRO NACIONAL DE ENGENHARIA DA PRODUÇÃO - ENEGEP, 2001, Salvador, Bahia. **Anais[...]**. Bahia: ENEGEP, 2001.
- LINTNER, John. Security prices, risk, and maximal gains from diversification. **Journal of Finance**, v. 20, n. 4, p. 587-615, 1965.
- MARKOWITZ, H. Portfolio selection. **Journal of Finance**, v. 7, n.1, p. 77-91, 1952.
- MARKOWITZ, H. **Portfolio selection: efficient diversification of investment**. New York: Wiley, 1959.
- MILANI, Bruno; CERETTA, Paulo Sérgio. Avaliação de performance de fundos de investimento: a história do pensamento atual. **Revista Estudo & Debate**, v. 19, n. 1, 2012.
- MOSSIN, J. Equilibrium in a capital asset market. **Econometrica**, v. 34, n. 4, p. 768-783, 1966.
- RABELO, S. S. T.; ROGERS, P.; RIBEIRO, K. C. S.; MENDES-DA-SILVA, W. Performance das melhores práticas de governança corporativa no Brasil: um estudo de carteiras. *In*: ENCONTRO DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO - ANPAD, 2007, Rio de Janeiro, RJ. **Anais[...]**. Rio de Janeiro: ANPAD, 2007.
- ROCHMAN, R. R.; EID JÚNIOR, W. Fundos de investimento ativos e passivos no Brasil: comparando e determinando seus desempenhos. *In*: ENCONTRO DA ASSOCIAÇÃO NACIONAL DOS PROGRAMAS DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO - ANPAD, 2006, Salvador, Bahia. **Anais[...]**. Salvador: ANPAD, 2006.
- SECURATO, José Roberto. **Decisões financeiras em condições de risco**. 2. ed. São Paulo: Atlas, 1993.



SHARPE, William. Capital asset prices: a theory of market equilibrium under conditions of risk. **The Journal of Finance**, v. 19, n. 3, p. 425-442, 1964.

SHARPE, William. Mutual fund performance. **The Journal of Business**, v. 39, n. 1, p. 119-138, 1966.

SHARPE, William. **Portfolio theory and capital markets**. New York: McGraw-Hill, 1970.

SILVA, Breno; NOGUEIRA, Sérgio; RIBEIRO, Kárem. Aplicação prática do índice de Sharpe na determinação de um portfólio ótimo de ativos. **Revista Eletrônica de Administração - REA**, v. 14, n. 1, p. 85-99, 2015.

SILVEIRA, A. M.; BARROS, L. A. B. C.; FAMÁ, R. Estrutura de governança e valor nas companhias abertas brasileiras. **Caderno de Pesquisas em Administração**, v. 10, n. 1, p. 57-71, 2003.

TUKEY, John Wilder. **Exploratory data analysis**. Massachusetts: Addison-Wesley, 1977.

VARGA, Gyorgy. Índice de Sharpe e outros indicadores de performance aplicados a fundos de ações brasileiros. **Revista de Administração Contemporânea**, v. 5, n. 3, p. 215-245, 2001.

## APÊNDICE A . ESTATÍSTICAS DAS CARTEIRAS SIMULADAS

Estatística	Fama-French	Ibovespa	Mark.	Mark. - SmallCaps	Tx. Livre
Valor Mínimo	-25,66	-24,17	-6,63	-0,93	0,01
Quartil 1	-16,76	16,17	10,30	30,50	3,99
Mediana	-12,40	25,98	29,03	37,42	7,94
Média Aritmética	-8,08	24,98	32,08	46,06	7,90
Quartil 3	1,32	35,50	54,57	57,46	11,88
Valor Máximo	10,73	52,39	78,31	119,48	15,77
Erro Padrão da Média	0,36	0,50	0,87	0,97	0,16
Nvl. de conf. inf. da média (95%)	-8,79	23,99	30,37	44,15	7,59
Nvl. de conf. sup. da média (95%)	-7,37	25,97	33,80	47,97	8,22
Variância	1,06	2,06	6,18	7,65	0,21
Desvio Padrão	10,29	14,35	24,87	27,67	4,56

Fonte: Elaboração própria do autor baseado nos dados do Yahoo! Finance

## APÊNDICE B . CORRELAÇÃO DE CADA CARTEIRA SIMULADA COM A TAXA LIVRE DE RISCO

Portfólio	Correlação
Fama-French e Taxa Livre de Risco	-86,06%
Ibovespa e Taxa Livre de Risco	24,96%
Markowitz e Taxa Livre de Risco	89,47%
Markowitz - SmallCaps e Taxa Livre de Risco	81,85%
Taxa Livre de Risco e Taxa Livre de Risco	100,00%

Fonte: Elaboração própria do autor baseado nos dados do Yahoo! Finance

## APÊNDICE C . CÓDIGO R: GERANDO CARTEIRA DE MARKOWITZ

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

options(scipen = 999)

# pacotes R -----

library(tidyverse)
library(gridExtra)
library(PerformanceAnalytics)
library(quantmod)
library(fPortfolio)
library(tseries)
library(ggrepel)
library(xts)
library(lubridate)
library(viridis)
library(timetk)
library(tidyquant)

#install.packages("IntroCompFinR", repos="http://R-Forge.R-project.org")
library(IntroCompFinR)

setwd('C:\\Users\\cleit\\Desktop\\TCC\\Dados')

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

#####
##### Carregando a ajustando dados -----
#####

# Carregando lista de ativos -----
indicadores_setores_v2 <- read.csv2("indicadores_setores_v3.csv")

# Carregando dados de preços dos ativos até o dia 11/11/2021 -----
price_data <- read.csv2("acoes_setores_2022_04_21.csv")

# Transformando em date -----
price_data$date <- ymd(price_data$date)

datas <- data.frame(datas = unique(price_data$date)) %>% arrange(datas) %>% as.vector()
```

```

datas[round(length(datas$datas)*0.7,0),]

# Criando dataframe dos ativos únicos em cada ano de acordo com o maior volume -----
ativos_unico <- price_data %>%
  dplyr::mutate(ano = lubridate::year(date)) %>%
  dplyr::group_by(ano, symbol) %>%
  dplyr::summarise(media_vol = sum(volume, na.rm = TRUE)) %>%
  as.data.frame() %>%
  dplyr::mutate(ativo_no_number = stringr::str_extract(symbol, ".{4}") %>%
  dplyr::group_by(ano, ativo_no_number) %>%
  top_n(1, media_vol) %>%
  as.data.frame() %>%
  dplyr::mutate(id_ativo = paste0(symbol,"_",ano))

price_data <- price_data %>%
  dplyr::mutate(ano = lubridate::year(date)) %>%
  dplyr::mutate(id_ativo = paste0(symbol,"_",ano))

price_data <- price_data %>% dplyr::filter(symbol != 'CEGR3.SA')

table_pos_ajuste <- price_data %>%
  dplyr::filter(id_ativo %in% ativos_unico$id_ativo) %>%
  dplyr::filter(adjusted >= 0) %>%
  dplyr::group_by(symbol) %>%
  dplyr::mutate(media_volume = mean(volume, na.rm = TRUE)) %>%
  as.data.frame() %>%
  dplyr::filter(media_volume >= 1000) %>%
  dplyr::select(!c(id_ativo, media_volume, ano))

write.csv2(table_pos_ajuste, 'table_pos_ajuste.csv', row.names = F)

price_data <- price_data %>%
  dplyr::filter(id_ativo %in% ativos_unico$id_ativo) %>%
  dplyr::filter(adjusted >= 0) %>%
  dplyr::filter(date <= datas[round(length(datas$datas)*0.7,0),]) %>%
  dplyr::group_by(symbol) %>%
  dplyr::mutate(media_volume = mean(volume, na.rm = TRUE)) %>%
  as.data.frame() %>%
  dplyr::filter(media_volume >= 1000) %>%
  dplyr::select(symbol, date, adjusted) %>%
  spread(symbol,adjusted)

price_data <- price_data[,!is.na(price_data[1,])]

price_data <- price_data %>% as.data.frame()

price_data <- price_data %>% tk_xts()

# Se a ação n tiver preço, retorne o último valor dela -----
price_data <- na.locf(price_data, fromLast = FALSE)

```

```
#####
# Calculando os retornos do ativos -----
#####

# Retorno pelo método 'Log' -----
retornos = Return.calculate(price_data, method = "log")

# Retornos começam na primeira data; nenhuma posição anterior, retorno 0
retornos[1,] = 0

#####
### Calculando o Benchmark: Taxa DI diária no período -----
#####

# https://calculadorarendafixa.com.br/#
# https://www.contadordedias.com.br/contador-de-dias/

# percentual em decimal/quantidade de dias

datas_split <- datas %>% dplyr::filter(datas >= as.Date("2011-01-03") &
                                     datas <= as.Date("2018-09-10"))

nrow(datas_split)

benchmark = 1.1465/nrow(datas_split)

#####
### Calculando a média do retorno diário e a vol dos ativos -----
#####

# Matriz de covariância de todas as ações -----
Cov = cov(retornos)

# df to xts -----
retornos_df <- data.frame(date=index(retornos), coredata(retornos))

# Média do retorno e vol -----
retornos_df <- retornos_df %>%
  gather(acao, retorno_acao, -date) %>%
  dplyr::group_by(acao) %>%
  dplyr::summarise(media = mean(retorno_acao, na.rm = TRUE),
                  var = stats::sd(retorno_acao, na.rm = TRUE))

# nome das colunas -----
colnames(retornos_df) = c("Acao", "Retorno Medio", "Volatilidade")

# Calculando o Sharpe -----
retornos_df <- retornos_df %>% dplyr::mutate(Sharpe = ('Retorno Medio' - benchmark)/Volatilidade)

# Trazendo o setor para o df -----
retornos_df <- left_join(retornos_df,
                        indicadores_setores_v2 %>% select(acoes_sa, Setor),
                        by = c("Acao" = "acoes_sa"))
```

```
#####
### Gráfico da volatilidade/retorno de cada ação -----
#####

retorno_plot <- ggplot(retornos_df, aes(x = Volatilidade,
                                         y = 'Retorno Medio',
                                         label = Acao,
                                         colour = Setor)) +

  geom_point(size = 3.5) +
  geom_hline(yintercept = median(retornos_df$'Retorno Medio', na.rm = TRUE), linetype = "dashed") +
  geom_vline(xintercept = median(retornos_df$Volatilidade, na.rm = TRUE), linetype = "dashed") +
  annotate('text', x = 0.01, y = -0.0015,
            label = paste0("Mediana do Retorno Médio Diário:\n",
                           round(median(retornos_df$'Retorno Medio',
                                         na.rm = TRUE),5), "%")) +
  annotate(geom = 'segment',
          x = 0.01,
          xend = 0.01,
          y = -0.001,
          yend = -0.0001, color = 'black', arrow = arrow(type = "open")) +
  annotate('text', x = 0.01, y = -0.0035,
            label = paste0("Mediana da Volatilidade Média Diário:\n",
                           round(median(retornos_df$Volatilidade,
                                         na.rm = TRUE),5), "%")) +
  annotate(geom = 'segment',
          y = -0.004,
          yend = -0.004,
          x = 0.002,
          xend = 0.018, color = 'black', arrow = arrow(type = "open")) +
  labs(x = "Volatilidade Média (%)",
       y = "Retorno Médio (%)",
       colour = "Setor:") +
  theme_classic() +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(0,0.1, 0.01),
                    limits = c(0,0.1)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(-0.005, 0.002, 0.001),
                    limits = c(-0.005, 0.002)) +
  theme(legend.position = "top",
        legend.title = element_text(size = 14),
        plot.caption = element_text(size = 13),
        legend.text = element_text(size = 12),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_colour_discrete(
    drop = FALSE,
    guide = guide_legend(
      direction = "horizontal",
      keyheight = unit(2, units = "mm"),
      keywidth = unit(30/length(labels), units = "mm"),
      label.hjust = 0.5,
```

```

    nrow = 1,
    byrow = TRUE,
    reverse = FALSE,
    label.position = "top")) +
  guides(col = guide_legend(nrow = 2))

### Salvando em PNG -----
ggsave(plot = retorno_plot,
        paste0(pasta_graficos, "retorno_risco_plot2.png"),
        width = 14, height = 6, dpi = 250, units = "in", type = "cairo-png")

ret_setor <- retornos_df %>%
  dplyr::group_by(Setor) %>%
  dplyr::summarise('Total de Empresas' = n(),
                  'Retorno Médio Diário' = paste0(format(round(mean('Retorno Medio',
                                                                    na.rm = TRUE)*100,
                                                                    4),
                                                                    big.mark = ".",
                                                                    decimal.mark = ",", "%"),
                  'Volatilidade Média Diária' = paste0(format(round(mean(Volatilidade,
                                                                    na.rm = TRUE)*100,2),
                                                                    big.mark = ".",
                                                                    decimal.mark = ",", "%")) %>%

  as.data.frame() %>%
  arrange(desc('Total de Empresas'))

sum(ret_setor$'Total de Empresas')

write.csv2(ret_setor, 'ret_setor.csv', row.names = F)

# Observando subidas e descidas de cada uma das ações -----
x_date <- data.frame(date=index(retornos), coredata(retornos))
x_date[,1]

retornos_plot_df <- data.frame(date=index(retornos), coredata(retornos))
retornos_plot_df <- retornos_plot_df %>% gather(ativo, retorno, -date)

retornos_plot_df <- retornos_plot_df[-209362,]
retornos_plot_df <- retornos_plot_df[-15377,]

ret_diario_plot <- ggplot(retornos_plot_df, aes(x = date, y = retorno, colour = ativo)) +
  geom_line() +
  theme_classic() +
  labs(x = "",
       y = "Retorno diário (%)",
       colour = "Setor:") +
  scale_x_date(date_labels = "%Y", breaks = "1 year", expand = c(0,50)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",",
                                                    breaks = seq(-2, 2.2, 0.4),
                                                    limits = c(-2, 2.2)) +
  theme(legend.position = "none",
        legend.title = element_text(size = 14),
        plot.caption = element_text(size = 13),
        legend.text = element_text(size = 12),
        strip.text = element_text(face = "bold", size = "11"),

```



```

plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
axis.text.y = element_text(face="bold", color="#000000",size = 14),
axis.text.x = element_text(face="bold", color="#000000",size = 14),
axis.title.y = element_text(size = 16),
axis.title.x = element_text(size = 16)

### Salvando em PNG -----
ggsave(plot = ret_diario_plot,
        paste0(pasta_graficos, "retorno_diario_22042022.png"),
        width = 14, height = 6,
        dpi = 250, units = "in",type = "cairo-png")

#####
### Montando a fronteira eficiente de Markowitz -----
#####

# n observações para realizar Markowitz
n = 20000

# ativos para realizar a simulação
set.seed(1544)
x = nrow(retornos_df)

# Sorteando as ações -----
acoes_sorteadas <- sample(retornos_df$Acao,
                          size = x,
                          replace = FALSE)

print(acoes_sorteadas)

# Retorno/Vol das ações escolhidas -----

acoes_escolhidas_aleatorio <- retornos_df %>%
  dplyr::filter(Acao %in% acoes_sorteadas) %>%
  arrange(factor(Acao, levels = acoes_sorteadas))

retornos_sorteados <- retornos[,acoes_sorteadas]

# Matriz de covariância das ações -----
Cov_sorteados = cov(retornos_sorteados) # Matriz de covariância das ações

#View(cor(Cov_sorteados))

# Matriz de pesos -----
w <- matrix(data = actuar::rpareto(n*x,
                                   shape = 1,
                                   scale = 1),
            nrow = n,
            ncol = x,
            byrow = TRUE)

for(i in c(1:n)) {

```

```

w[i,] = w[i,] / sum(w[i,])
}

# Montando a volatilidade para cada peso e cada carteira -----
# Usaremos a forma algébrica

vol_w = rep(NA, n)

for(i in c(1:n)) {

  w_momento = matrix(data = w[i,],
                    ncol = 1)

  # Transposta do vetor de peso da carteira i
  # multiplicação matricial matriz de covariância
  # multiplicação matricial do vetor de peso da carteira i

  vol_momento = sqrt( t(w_momento) %% Cov_sorteados %% w_momento )

  # Armazenando a volatilidade de cada carteira

  vol_w[i] = vol_momento
}

# Montando o retorno para cada peso -----

ret_w = rep(NA, n)

# Peso de cada ação no retorno;
# será apagado a cada iteração

soma = rep(NA, x)

for(i in c(1:n)) {

  for(j in c(1:x)) {

    # Peso aleatório vezes retorno médio de cada ação

    soma[j] = w[i,j] * acoes_escolhidas_aleatorio$'Retorno Medio'[j]

  }

  # Retorno da carteira é a soma dos retornos de cada ação

  ret_w[i] = sum(soma)
}

#####
### Gráfico da fronteira -----
#####

# Unindo ambas informações em um dataframe apenas -----

markowitz <- tibble(ret_w, vol_w)

```

```

markowitz <- markowitz %>%
  mutate(Sharpe = (ret_w-benchmark)/vol_w)

names(markowitz) = c("Retorno Medio", "Volatilidade", "Sharpe")

min_var <- markowitz[which.min(markowitz$Volatilidade),]
max_sr <- markowitz[which.max(markowitz$Sharpe),]

p <- markowitz %>%
  ggplot(aes(x = Volatilidade, y = 'Retorno Medio'*100, color = Sharpe)) +
  geom_point() +
  theme_classic() +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +
  theme(legend.position = "top",
        legend.title = element_text(size = 15),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  labs(x = 'Risco Anualizado (%)',
        y = 'Retornos Anualizados (%)') +
  geom_point(aes(x = Volatilidade,
                 y = 'Retorno Medio'), data = min_var, color = 'red', size = 3) +
  geom_point(aes(x = Volatilidade,
                 y = 'Retorno Medio'*100), data = max_sr, color = 'red', size = 3) +
  annotate('text', x = 0.017, y = 0.25, label = "Portfólio com a melhor tangência") +
  annotate('text', x = 0.01, y = -0.22, label = "Portfólio de\nvariação\nmínima") +
  annotate(geom = 'segment', x = 0.017, xend = 0.020685, y = 0.23,
          yend = 0.14, color = 'black', arrow = arrow(type = "open")) +
  annotate(geom = 'segment', x = 0.01, xend = 0.0062, y = -0.12,
          yend = -0.02, color = 'black', arrow = arrow(type = "open")) +
  scale_colour_viridis(
    option = "viridis",
    direction = -1,
    guide = guide_colorbar(
      direction = "horizontal",
      barheight = unit(3, units = "mm"),
      barwidth = unit(100, units = "mm"),
      draw.ulim = F,
      title.position = 'top',
      label.position = "top",
      title.hjust = 0.5,
      label.hjust = 0.5
    )
  )

### Salvando em PNG -----
ggsave(plot = p, paste0(pasta_graficos, "fronteira_todas_acoes.png"),
        width = 12, height = 5.5, dpi = 250, units = "in", type = "cairo-png")

#####
# Montando a fronteira eficiente -----

```

```
#####

# Pacote {fPortfolio}
# Monta a simulação apenas com esse código -----

fronteira = portfolioFrontier(as.timeSeries(retornos_sorteados))

# Gráfico da fronteira -----

frontierPlot(fronteira,
             pch = 20,
             cex = 1.5,
             type = "o",
             lwd = 2)

singleAssetPoints(fronteira,
                  col = c(1:x),
                  pch = 20,
                  cex = 1.5) # Retorno/Vol de cada ação

# Quantidade de simulações -----
monteCarloPoints(fronteira,
                 mcSteps = n,
                 pch = 20,
                 cex = 0.1,
                 col = "steelblue") # Simulações com pesos aleatórios

# Distribuição/comportamento de retorno e vol das carteiras simuladas -----
vol_w_df <- data.frame(valor = vol_w)

# Distribuição/comportamento de retorno e vol das carteiras simuladas -----
vol_w_df <- data.frame(valor = vol_w)

vol_diaria_carteiras <- ggplot(data = vol_w_df, aes(x = valor)) +
  geom_histogram(bins = 200, fill = "#173F5F", colour = "white") +
  labs(x="Volatilidade Diária (%)", y="Frequência",
       colour = "",
       title = "Volatilidade das Carteiras Simuladas (%)") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 14),
        plot.caption = element_text(size = 13),
        legend.text = element_text(size = 12),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 15, hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_y_continuous(labels = scales::number_format(big.mark = ".", decimal.mark = ",")) +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(0, 0.07, .01),
                    limits = c(0,0.07))

ret_w_df <- data.frame(valor = ret_w)
```

```

ret_diaria_carteiras <- ggplot(data = ret_w_df, aes(x = valor)) +
  geom_histogram(bins = 200, fill = "#bc5090", colour = "white") +
  labs(x="Retornos Diários (%)", y="Frequência",
       colour = "",
       title = "Retornos Diários das Carteiras Simuladas (%)") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 14),
        plot.caption = element_text(size = 13),
        legend.text = element_text(size = 12),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 15, hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(-0.003, 0.001, .001),
                    limits = c(-0.003,0.001))

plot_vol_ret_diaria <- arrangeGrob(ret_diaria_carteiras, vol_diaria_carteiras, ncol = 2)

### Salvando em PNG ###
ggsave(plot = plot_vol_ret_diaria,
       paste0(pasta_graficos, "retorno_vol_carteiras_simulacao_22042022.png"),
       width = 14, height = 6,
       dpi = 250, units = "in",type = "cairo-png")

#####
### Obtendo carteiras de min var e carteira ótima -----
#####

# Descobrimo a carteira ótima -----
ótimo <- IntroCompFinR::tangency.portfolio(er = acoes_escolhidas_aleatorio$'Retorno Medio',
                                           cov.mat = Cov_sorteados,
                                           risk.free = 0.0001,
                                           shorts = FALSE)

ret_esp_ótimo = ótimo[[2]] # Retorno esperado da carteira ótima
vol_esp_ótimo = ótimo[[3]] # Volatilidade da carteira ótima

carteira_ótima = tibble(ret_esp_ótimo, vol_esp_ótimo)

names(carteira_ótima) = c("Retorno Medio", "Volatilidade")

#####
### Tabela com resultados -----
#####

# Para os pesos ótimos para cada ação numa tabela -----

pesos_ótimos = ótimo[[4]]

matriz_ótima = matrix(data = NA,

```

```

        nrow = 3,
        ncol = length(acoas_sorteadas))

matriz_ótima[1,] = paste0(round(pesos_ótimos*100, digits = 2))
matriz_ótima[2,] = paste0(round(acoas_escolhidas_aleatorio$'Retorno Medio'*100, digits = 2))
matriz_ótima[3,] = paste0(round(acoas_escolhidas_aleatorio$Volatilidade*100, digits = 2))

matriz_ótima = as.data.frame(matriz_ótima,
                             row.names = c("Pesos", "Retorno", "Volatilidade"))

names(matriz_ótima) = acoes_escolhidas_aleatorio$Acao

matriz_otima_long <- matriz_ótima %>% gather(ativo, valor)

matriz_otima_long <- matriz_otima_long %>% mutate(info = rep(c("Pesos",
                                                             "Retorno",
                                                             "Volatilidade"),
                                                             nrow(retornos_df)))

matriz_otima_long <- matriz_otima_long %>% spread(info, valor)

matriz_otima_long <- matriz_otima_long %>% dplyr::filter(Pesos != 0)

matriz_otima_long$Pesos <- as.numeric(matriz_otima_long$Pesos)
matriz_otima_long$Retorno <- as.numeric(matriz_otima_long$Retorno)
matriz_otima_long$Volatilidade <- as.numeric(matriz_otima_long$Volatilidade)

write.csv2(matriz_otima_long,
           "carteira_eficiente2.csv",
           row.names = F)

#####
##### Gerando dados -----
#####

carteira_eficiente <- read.csv2("carteira_eficiente2.csv")

carteira_eficiente$Pesos <- as.numeric(carteira_eficiente$Pesos)

cotacao_2 <- tq_get(carteira_eficiente$ativo,
                   from = "2018-09-10",
                   to = "2021-12-31",
                   get = "stock.prices")

cotacao_2 <- cotacao_2 %>%
  dplyr::select(symbol, date, adjusted) %>%
  spread(symbol,adjusted) %>%
  tk_xts()

# Se a ação n tiver preço, retorne o último valor dela -----
cotacao_2 <- na.locf(cotacao_2, fromLast = TRUE)

```

```

# calculando o retorno discreto
retornos = Return.calculate(cotacao_2, method = "log")

# Retornos começam na primeira data; nenhuma posição anterior, retorno 0
retornos[1,] = 0

# pesos gerados do outro script
portReturns <- Return.portfolio(retornos,
                                (carteira_eficiente$Pesos/100),
                                rebalance_on = "years")
portReturns_est <- Return.portfolio(retornos, (carteira_eficiente$Pesos/100),
                                    rebalance_on = "years", verbose = TRUE, wealth.index = TRUE)

# transformando xts em dataframe
carteira <- data.frame(date=index(portReturns), coredata(portReturns))

# criando coluna com nome da carteira e mudando nome do retorno
carteira <- carteira %>%
  dplyr::mutate(tipo = "Carteira Eficiente pelo método de Markowitz") %>%
  dplyr::rename(retorno_diario = "portfolio.returns")

# retorno acumulado
port_cumulative_ret <- carteira %>% mutate(cr = cumprod(1 + retorno_diario))

port_cumulative_ret_xts <- port_cumulative_ret %>%
  dplyr::select(date, cr) %>%
  dplyr::rename('Carteira Eficiente pelo método de Markowitz' = "cr") %>%
  tk_xts()

#####
##### Ibovespa -----
#####

# gerando dados
bovespa <- getSymbols("^BVSP",
                      src = "yahoo",
                      from = "2018-09-10",
                      to = "2021-12-31",
                      periodicity="daily",
                      auto.assign = T,
                      warnings = F) %>%
  na.omit()

# calculando retorno simples
retornos_ibovespa = Return.calculate(BVSP[,4],
                                     method = "log")

# transformando NA em zero
retornos_ibovespa[1,] = 0

# removendo NA
retornos_ibovespa <- na.omit(retornos_ibovespa)

```

```

# transformando em dataframe
indice <- data.frame(date=index(retornos_ibovespa), coredata(retornos_ibovespa))

# criando coluna com nome do indice e renomenado
indice <- indice %>% dplyr::mutate(tipo = "Ibovespa") %>%
  dplyr::rename(retorno_diario = "BVSP.Close")

# calculando retorno acumulado
port_cum_ret_indice <- indice %>% mutate(cr = cumprod(1 + retorno_diario))

port_cumulative_ibov_xts <- port_cum_ret_indice %>%
  dplyr::select(date, cr) %>%
  dplyr::rename("Ibovespa" = "cr") %>% tk_xts()

merge_port_ibov <- merge(port_cumulative_ret_xts,
  port_cumulative_ibov_xts)[time(port_cumulative_ret_xts), ]

#####
##### Unindo bancos e gerando gráfico -----
#####

merge_port_ibov_df <- data.frame(date=index(merge_port_ibov), coredata(merge_port_ibov))
merge_port_ibov_df <- merge_port_ibov_df %>% gather(tipo, media_ret, -date)

merge_port_ibov_df <- merge_port_ibov_df %>% mutate(media_ret = media_ret - 1)

merge_port_ibov_df$tipo <- ifelse(
  merge_port_ibov_df$tipo == "Carteira.Eficiente.pelo.metódo.de.Markowitz",
  "Carteira Eficiente pelo método de Markowitz",
  merge_port_ibov_df$tipo)

merge_port_ibov_df$tipo <- ifelse(
  merge_port_ibov_df$tipo == "Carteira Eficiente pelo método de Markowitz",
  "Portfólio - Markowitz",
  merge_port_ibov_df$tipo)

ret_acum_graph <- ggplot(merge_port_ibov_df, aes(x = date, y = media_ret, colour = tipo)) +
  geom_line(size = 1.3) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(x="", y="Retorno Acumulado",
  colour = "") +
  theme_classic() +
  theme(legend.position = "top",
  legend.title = element_text(size = 10),
  plot.caption = element_text(size = 7),
  legend.text = element_text(size = 15),
  strip.text = element_text(face = "bold", size = "11"),
  plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
  plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
  axis.text.y = element_text(face="bold", color="#000000",size = 14),
  axis.text.x = element_text(face="bold", color="#000000",size = 14),
  axis.title.y = element_text(size = 16),
  axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", breaks = "2 months", expand = c(0,50)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +

```



```
scale_colour_manual(values = c("#173F5F", "#bc5090"),
  drop = FALSE,
  guide = guide_legend(
    direction = "horizontal",
    keyheight = unit(2, units = "mm"),
    keywidth = unit(30/length(labels), units = "mm"),
    label.hjust = 0.5,
    nrow = 1,
    byrow = TRUE,
    reverse = FALSE,
    label.position = "top"))

### Salvando csv -----
write.csv2(merge_port_ibov_df,
  "retorno_acum_ibov_markowitz_22042022.csv",
  row.names = F)

### Salvando em PNG -----
ggsave(plot = ret_acum_graph,
  paste0(pasta_graficos, "ret_acum_graph_2.png"),
  width = 14, height = 7,
  dpi = 250, units = "in", type = "cairo-png")
```

## APÊNDICE D . CÓDIGO R: GERANDO CARTEIRA DE MARKOWITZ - *SMALLCAPS*

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

options(scipen = 999)

# pacotes R -----

library(plotly)
library(tidyverse)
library(gridExtra)
library(ggplot2)
library(PerformanceAnalytics)
library(quantmod)
library(fPortfolio)
library(tseries)
library(ggrepel)
library(actuar)
library(tseries)
library(reticulate)
library(xts)
library(lubridate)
library(viridis)
library(timetk)
library(tidyquant)

#install.packages("IntroCompFinR", repos="http://R-Forge.R-project.org")
library(IntroCompFinR)

setwd("C:\\Users\\pc\\Desktop\\TCC\\Dados")

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

#####
##### Carregando a ajustando dados -----
#####

# Carregando lista de ativos -----
indicadores_setores_v2 <- read.csv2("indicadores_setores_v3.csv")
small_caps <- indicadores_setores_v2 %>% dplyr::filter(Classificação == "Small Cap")

indicadores_setores_v2$Classificação

# Carregando dados de preços dos ativos até o dia 11/11/2021 -----
```

```

price_data <- read.csv2("acoes_setores_2022_04_21.csv")

# Transformando em date -----
price_data$date <- ymd(price_data$date)

price_data <- price_data %>% dplyr::filter(symbol %in% small_caps$acoes_sa)

datas <- data.frame(datas = unique(price_data$date)) %>% arrange(datas) %>% as.vector()

datas[round(length(datas$datas)*0.7,0),]

# Criando dataframe dos ativos únicos em cada ano de acordo com o maior volume -----
ativos_unico <- price_data %>%
  dplyr::mutate(ano = lubridate::year(date)) %>%
  dplyr::group_by(ano, symbol) %>%
  dplyr::summarise(media_vol = sum(volume, na.rm = TRUE)) %>%
  as.data.frame() %>%
  dplyr::mutate(ativo_no_number = stringr::str_extract(symbol, "^.{4}") %>%
  dplyr::group_by(ano, ativo_no_number) %>%
  top_n(1, media_vol) %>%
  as.data.frame() %>%
  dplyr::mutate(id_ativo = paste0(symbol,"_",ano))

price_data <- price_data %>%
  dplyr::mutate(ano = lubridate::year(date)) %>%
  dplyr::mutate(id_ativo = paste0(symbol,"_",ano))

price_data <- price_data %>% dplyr::filter(symbol != 'CEGR3.SA')

price_data <- price_data %>%
  dplyr::filter(id_ativo %in% ativos_unico$id_ativo) %>%
  dplyr::filter(adjusted >= 0) %>%
  dplyr::filter(date <= datas[round(length(datas$datas)*0.7,0),]) %>%
  as.data.frame() %>%
  dplyr::select(symbol, date, adjusted) %>%
  spread(symbol,adjusted)

price_data <- price_data[,!is.na(price_data[1,])]

price_data <- price_data %>% as.data.frame()

price_data <- price_data%>% tk_xts()

# Se a ação n tiver preço, retorne o último valor dela -----
price_data <- na.locf(price_data, fromLast = FALSE)

#####
# Calculando os retornos do ativos -----
#####

# Retorno pelo método 'Log' -----

```



```
#####
### Gráfico da volatilidade/retorno de cada ação -----
#####

retorno_plot <- ggplot(retornos_df, aes(x = Volatilidade,
                                         y = 'Retorno Medio',
                                         label = Acao,
                                         colour = Setor)) +

  geom_point(size = 3.5) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  labs(x = "Volatilidade",
       y = "Retorno Diário",
       colour = "Setor:") +
  theme_classic() +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(0,0.1, 0.02),
                    limits = c(0,0.1)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +
  theme(legend.position = "top",
       legend.title = element_text(size = 14),
       plot.caption = element_text(size = 13),
       legend.text = element_text(size = 12),
       strip.text = element_text(face = "bold", size = "11"),
       plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
       plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
       axis.text.y = element_text(face="bold", color="#000000",size = 14),
       axis.text.x = element_text(face="bold", color="#000000",size = 14),
       axis.title.y = element_text(size = 16),
       axis.title.x = element_text(size = 16)) +
  scale_colour_discrete(
    drop = FALSE,
    guide = guide_legend(
      direction = "horizontal",
      keyheight = unit(2, units = "mm"),
      keywidth = unit(30/length(labels), units = "mm"),
      label.hjust = 0.5,
      nrow = 1,
      byrow = TRUE,
      reverse = FALSE,
      label.position = "top")) +
  guides(col = guide_legend(nrow = 2))

### Salvando em PNG ###
ggsave(plot = retorno_plot,
       paste0(pasta_graficos, "retorno_risco_smallcaps.png"),
       width = 14, height = 6, dpi = 250, units = "in",type = "cairo-png")

x_date <- data.frame(date=index(retornos), coredata(retornos))
x_date[,1]

retornos_plot_df <- data.frame(date=index(retornos), coredata(retornos))
retornos_plot_df <- retornos_plot_df %>% gather(ativo, retorno, -date)

ret_diario_plot <- ggplot(retornos_plot_df, aes(x = date, y = retorno, colour = ativo)) +
```

```

geom_line() +
theme_classic() +
labs(x = "",
     y = "Retorno diário (%)",
     colour = "Setor:") +
scale_x_date(date_labels = "%Y", breaks = "1 year", expand = c(0,50)) +
scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                   breaks = seq(-1.2, 1.8, 0.4),
                   limits = c(-1.2, 1.8)) +
theme(legend.position = "none",
      legend.title = element_text(size = 14),
      plot.caption = element_text(size = 13),
      legend.text = element_text(size = 12),
      strip.text = element_text(face = "bold", size = "11"),
      plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
      plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
      axis.text.y = element_text(face="bold", color="#000000",size = 14),
      axis.text.x = element_text(face="bold", color="#000000",size = 14),
      axis.title.y = element_text(size = 16),
      axis.title.x = element_text(size = 16))

### Salvando em PNG ###
ggsave(plot = ret_diario_plot,
       paste0(pasta_graficos, "retorno_diario_smallcaps.png"),
       width = 14, height = 6, dpi = 250, units = "in",type = "cairo-png")

#####
### Montando a fronteira eficiente de Markowitz -----
#####

# n observações para realizar Markowitz
n = 20000

# ativos para realizar a simulação
set.seed(1544)
x = nrow(retornos_df)

# Sorteando as ações -----
acoes_sorteadas <- sample(retornos_df$Acao,
                          size = x,
                          replace = FALSE)

print(acoes_sorteadas)

# Retorno/Vol das ações escolhidas -----

acoes_escolhidas_aleatorio <- retornos_df %>%
  dplyr::filter(Acao %in% acoes_sorteadas) %>%
  arrange(factor(Acao, levels = acoes_sorteadas))

retornos_sorteados <- retornos[,acoes_sorteadas]

```

```

# Matriz de covariância das ações -----
Cov_sorteados = cov(retornos_sorteados) # Matriz de covariância das ações

#View(cor(Cov_sorteados))

# Matriz de pesos -----
w <- matrix(data = actuar::rpareto(n*x,
                                shape = 1,
                                scale = 1),
            nrow = n,
            ncol = x,
            byrow = TRUE)

for(i in c(1:n)) {

  w[i,] = w[i,] / sum(w[i,])

}

# Montando a volatilidade para cada peso e cada carteira -----
# Usaremos a forma algébrica

vol_w = rep(NA, n)

for(i in c(1:n)) {

  w_momento = matrix(data = w[i,],
                    ncol = 1)

  # Transposta do vetor de peso da carteira i
  # multiplicação matricial matriz de covariância
  # multiplicação matricial do vetor de peso da carteira i

  vol_momento = sqrt( t(w_momento) %%% Cov_sorteados %%% w_momento )

  # Armazenando a volatilidade de cada carteira

  vol_w[i] = vol_momento

}

# Montando o retorno para cada peso -----

ret_w = rep(NA, n)

# Peso de cada ação no retorno;
# será apagado a cada iteração

soma = rep(NA, x)

for(i in c(1:n)) {

  for(j in c(1:x)) {

    # Peso aleatório vezes retorno médio de cada ação

    soma[j] = w[i,j] * acoes_escolhidas_aleatorio$'Retorno Medio'[j]
  }
}

```

```

}

# Retorno da carteira é a soma dos retornos de cada ação

ret_w[i] = sum(soma)

}

#####
### Gráfico da fronteira -----
#####

# Unindo ambas informações em um dataframe apenas:

markowitz <- tibble(ret_w, vol_w)

markowitz <- markowitz %>%
  dplyr::mutate(Sharpe = (ret_w-benchmark)/vol_w)

names(markowitz) = c("Retorno Medio", "Volatilidade", "Sharpe")

min_var <- markowitz[which.min(markowitz$Volatilidade),]
max_sr <- markowitz[which.max(markowitz$Sharpe),]

p <- markowitz %>%
  ggplot(aes(x = Volatilidade, y = 'Retorno Medio'*100, color = Sharpe)) +
  geom_point() +
  theme_classic() +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +
  scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ",")) +
  theme(legend.position = "top",
        legend.title = element_text(size = 14),
        plot.caption = element_text(size = 13),
        legend.text = element_text(size = 12),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  labs(x = 'Risco Anualizado (%)',
       y = 'Retornos Anualizados (%)'
  ) +
  geom_point(aes(x = Volatilidade,
                y = 'Retorno Medio'), data = min_var, color = 'red', size = 3) +
  geom_point(aes(x = Volatilidade,
                y = 'Retorno Medio'*100), data = max_sr, color = 'red', size = 3) +
  annotate('text', x = 0.030, y = 0.13, label = "Portfólio com a melhor tangência", size = 5) +
  annotate('text', x = 0.01, y = -0.10, label = "Portfólio de\nvariação\nmínima", size = 5) +
  annotate(geom = 'segment', x = 0.030, xend = 0.0225, y = 0.12,
        yend = 0.07, color = 'black', arrow = arrow(type = "open")) +
  annotate(geom = 'segment', x = 0.009, xend = 0.0065, y = -0.065,
        yend = -0.01, color = 'black', arrow = arrow(type = "open")) +
  scale_colour_viridis(
    option = "viridis",
    direction = -1,
    guide = guide_colorbar(

```



```

direction = "horizontal",
barheight = unit(3, units = "mm"),
barwidth = unit(100, units = "mm"),
draw.ulim = F,
title.position = 'top',
label.position = "top",
title.hjust = 0.5,
label.hjust = 0.5
))

### Salvando em PNG ###
ggsave(plot = p, paste0(pasta_graficos, "fronteira_smallcaps.png"),
       width = 12, height = 5.5, dpi = 250, units = "in", type = "cairo-png")

#####
# Montando a fronteira eficiente -----
#####

# Pacote {fPortfolio}
# Monta a simulação apenas com esse código -----

fronteira = portfolioFrontier(as.timeSeries(retornos_sorteados))

# Gráfico da fronteira -----

frontierPlot(fronteira,
             pch = 20,
             cex = 1.5,
             type = "o",
             lwd = 2)

singleAssetPoints(fronteira,
                  col = c(1:x),
                  pch = 20,
                  cex = 1.5) # Retorno/Vol de cada ação

# Quantidade de simulações -----
monteCarloPoints(fronteira,
                 mcSteps = n,
                 pch = 20,
                 cex = 0.1,
                 col = "steelblue") # Simulações com pesos aleatórios

# Distribuição/comportamento de retorno e vol das carteiras simuladas -----

vol_w_df <- data.frame(valor = vol_w)

vol_diaria_carteiras <- ggplot(data = vol_w_df, aes(x = valor)) +
  geom_histogram(bins = 200, fill = "#173F5F", colour = "white") +
  labs(x="Volatilidade Diária (%)", y="Frequência",
       colour = "",
       title = "Volatilidade das Carteiras Simuladas (%) - SmallCaps") + #,
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 14),

```

```

plot.caption = element_text(size = 13),
legend.text = element_text(size = 12),
strip.text = element_text(face = "bold", size = "11"),
plot.title = element_text(size = 15, hjust = 0.5, face="bold"),
plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
axis.text.y = element_text(face="bold", color="#000000",size = 14),
axis.text.x = element_text(face="bold", color="#000000",size = 14),
axis.title.y = element_text(size = 16),
axis.title.x = element_text(size = 16)) +
scale_y_continuous(labels = scales::number_format(big.mark = ".", decimal.mark = ",")) +
scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
breaks = seq(0, 0.07, .01),
limits = c(0,0.07))

ret_w_df <- data.frame(valor = ret_w)

ret_diaria_carteras <- ggplot(data = ret_w_df, aes(x = valor)) +
  geom_histogram(bins = 200, fill = "#bc5090", colour = "white") +
  labs(x="Retornos Diários (%)", y="Frequência",
  colour = "",
  title = "Retornos Diários das Carteiras Simuladas (%) - SmallCaps") +
  theme_classic() +
  theme(legend.position = "none",
  legend.title = element_text(size = 14),
  plot.caption = element_text(size = 13),
  legend.text = element_text(size = 12),
  strip.text = element_text(face = "bold", size = "11"),
  plot.title = element_text(size = 15, hjust = 0.5, face="bold"),
  plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
  axis.text.y = element_text(face="bold", color="#000000",size = 14),
  axis.text.x = element_text(face="bold", color="#000000",size = 14),
  axis.title.y = element_text(size = 16),
  axis.title.x = element_text(size = 16)) +
scale_x_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","))

plot_vol_ret_diaria <- arrangeGrob(ret_diaria_carteras,
  vol_diaria_carteras, ncol = 2)

### Salvando em PNG ###
ggsave(plot = plot_vol_ret_diaria,
  paste0(pasta_graficos, "retorno_vol_carteras_simulacao_smallcaps.png"),
  width = 14, height = 6,
  dpi = 250, units = "in", type = "cairo-png")

#####
#### Obtendo carteiras de min var e carteira ótima -----
#####

# Carteira de mínima variância -----

# mínima_variância = fronteira_by_hand %>%
#   arrange(Volatilidade) %>%
#   slice(1)

```

```

# Descobrimos a carteira ótima -----

ótimo <- IntroCompFinR::tangency.portfolio(er = acoes_escolhidas_aleatorio$'Retorno Medio',
                                           cov.mat = Cov_sorteados,
                                           risk.free = 0.0001,
                                           shorts = FALSE)

ret_esp_ótimo = ótimo[[2]] # Retorno esperado da carteira ótima
vol_esp_ótimo = ótimo[[3]] # Volatilidade da carteira ótima

carteira_ótima = tibble(ret_esp_ótimo, vol_esp_ótimo)

names(carteira_ótima) = c("Retorno Medio", "Volatilidade")

#####
### Tabela com resultados -----
#####

# Para os pesos ótimos para cada ação numa tabela -----

pesos_ótimos = ótimo[[4]]

matriz_ótima = matrix(data = NA,
                      nrow = 3,
                      ncol = length(acoes_sorteadas))

matriz_ótima[1,] = paste0(round(pesos_ótimos*100, digits = 2))
matriz_ótima[2,] = paste0(round(acoes_escolhidas_aleatorio$'Retorno Medio'*100, digits = 2))
matriz_ótima[3,] = paste0(round(acoes_escolhidas_aleatorio$Volatilidade*100, digits = 2))

matriz_ótima = as.data.frame(matriz_ótima,
                              row.names = c("Pesos",
                                             "Retorno",
                                             "Volatilidade"))

names(matriz_ótima) = acoes_escolhidas_aleatorio$Acao

matriz_otima_long <- matriz_ótima %>% gather(ativo, valor)

matriz_otima_long <- matriz_otima_long %>%
  mutate(info = rep(c("Pesos",
                     "Retorno",
                     "Volatilidade"), nrow(retornos_df)))

matriz_otima_long <- matriz_otima_long %>% spread(info, valor)

matriz_otima_long <- matriz_otima_long %>% dplyr::filter(Pesos != 0)

matriz_otima_long$Pesos <- as.numeric(matriz_otima_long$Pesos)
matriz_otima_long$Retorno <- as.numeric(matriz_otima_long$Retorno)
matriz_otima_long$Volatilidade <- as.numeric(matriz_otima_long$Volatilidade)

write.csv2(matriz_otima_long, "carteira_eficiente_smallcaps.csv", row.names = F)

```

```
#####
##### Gerando dados -----
#####

carteira_eficiente <- read.csv2("carteira_eficiente_smallcaps.csv")

carteira_eficiente$Pesos <- as.numeric(carteira_eficiente$Pesos)

cotacao_2 <- tq_get(carteira_eficiente$ativo,
                   from = "2018-09-10",
                   to = "2021-12-31",
                   get = "stock.prices")

# Removendo ações cujo volume no período seja menor que R$ 100.000,00 e transformando df em xts -----
cotacao_2 <- cotacao_2 %>%
  dplyr::select(symbol, date, adjusted) %>%
  spread(symbol,adjusted) %>%
  tk_xts()

# Se a ação n tiver preço, retorne o último valor dela -----
cotacao_2 <- na.locf(cotacao_2, fromLast = TRUE)

# calculando o retorno discreto
retornos = Return.calculate(cotacao_2, method = "log")

# Retornos começam na primeira data; nenhuma posição anterior, retorno 0
retornos[1,] = 0

# pesos gerados do outro script
portReturns <- Return.portfolio(retornos,
                                (carteira_eficiente$Pesos/100),
                                rebalance_on = "years")

portReturns_est <- Return.portfolio(retornos,
                                    (carteira_eficiente$Pesos/100),
                                    rebalance_on = "years",
                                    verbose = TRUE,
                                    wealth.index = TRUE)

# transformando xts em dataframe
carteira <- data.frame(date=index(portReturns), coredata(portReturns))

# criando coluna com nome da carteira e mudando nome do retorno
carteira <- carteira %>%
  dplyr::mutate(tipo = "Carteira Eficiente Markowitz - SmallCaps") %>%
  dplyr::rename(retorno_diario = "portfolio.returns")

# retorno acumulado
port_cumulative_ret <- carteira %>% mutate(cr = cumprod(1 + retorno_diario))
```

```

port_cumulative_ret_xts <- port_cumulative_ret %>%
  dplyr::select(date, cr) %>%
  dplyr::mutate(cr = cr - 1) %>%
  dplyr::rename('Carteira Eficiente Markowitz - SmallCaps' = "cr") %>%
  tk_xts()

#####
#### Lendo dados da carteira 1 -----
#####

carteira_1 <- read.csv2("retorno_acum_ibov_markowitz_22042022.csv")

carteira_1$date <- lubridate::ymd(carteira_1$date)

carteira_1 <- carteira_1 %>% spread(tipo, media_ret) %>% tk_xts()

merge_port_full <- merge(port_cumulative_ret_xts,
                        carteira_1)[time(port_cumulative_ret_xts), ]

#####
##### Unindo bancos e gerando gráfico -----
#####

merge_port_full_df <- data.frame(date=index(merge_port_full),
                                coredata(merge_port_full))
merge_port_full_df <- merge_port_full_df %>% gather(tipo, media_ret, -date)

merge_port_full_df$tipo <- ifelse(
  merge_port_full_df$tipo == "Portfólio...Markowitz",
  "Markowitz",
  merge_port_full_df$tipo)

merge_port_full_df$tipo <- ifelse(
  merge_port_full_df$tipo == "Carteira.Eficiente.Markowitz...SmallCaps",
  "Markowitz - SmallCaps",
  merge_port_full_df$tipo)

merge_port_full_df <- merge_port_full_df %>% drop_na(media_ret)

ret_acum_graph <- ggplot(merge_port_full_df, aes(x = date, y = media_ret, colour = tipo)) +
  geom_line(size = 1.3) +
  labs(x="", y="Retorno Acumulado",
       colour = "",
       title = "Carteira de Markowitz x Ibovespa") +
  theme_minimal() +
  geom_hline(yintercept = 0, linetype = "dashed") +
  theme(legend.position = "top",
        legend.title = element_text(size = 10),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size=17, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size=10),
        axis.text.x = element_text(face="bold", color="#000000",size=10),
        axis.title.y = element_text(size = 10)) +
  scale_x_date(date_labels = "%b\n%Y", breaks = "2 months", expand = c(0,50)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),

```

```
breaks = seq(-.3,1.3,.3),
limits = c(-.3,1.3) +
scale_colour_manual(values = c("#173F5F", "#bc5090", "#ffa600"),
drop = FALSE,
guide = guide_legend(
direction = "horizontal",
keyheight = unit(2, units = "mm"),
keywidth = unit(50/length(labels), units = "mm"),
label.hjust = 0.5,
nrow = 1,
byrow = TRUE,
reverse = FALSE,
label.position = "top"))

### Salvando csv -----
write.csv2(merge_port_full_df, "retorno_acum_ibov_markowitz_all.csv", row.names = F)

### Salvando em PNG -----
ggsave(plot = ret_acum_graph,
"ret_acum_graph_all.png",
width = 14, height = 7, dpi = 250, units = "in",type = "cairo-png")
```

## APÊNDICE E . CÓDIGO R: GERANDO FATOR SMB - FAMA-FRENCH

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

library(tidyverse)
library(lubridate)
library(readxl)

options(scipen = 999)

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

pasta_dados <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\'

#####
##### Carregando a ajustando dados -----
#####

setwd("C:\\Users\\cleit\\Desktop\\TCC\\Dados\\MarketCap")

# Carregando lista de ativos -----
indicadores_setores_v2 <- read.csv2(paste0(pasta_dados, "indicadores_setores_v3.csv"))

indicadores_setores_v2$Ações <- ifelse(indicadores_setores_v2$Ações == "",
                                     NA,
                                     indicadores_setores_v2$Ações)

indicadores_setores_v2 <- indicadores_setores_v2 %>% drop_na(Ações)

num_acoes_df <- read.csv2("num_acoes_df.csv")

num_acoes_df <- num_acoes_df %>%
  dplyr::mutate(acoes_sa = paste0(symbol, ".SA")) %>% dplyr::select(!symbol)

# Carregando dados de preços dos ativos até o dia 11/11/2021 -----
price_data <- read.csv2(paste0(pasta_dados, "acoes_setores_2022_04_21.csv"))

price_data_num_acoes <- inner_join(num_acoes_df,
                                  price_data, by = c("acoes_sa" = "symbol"))

price_data_num_acoes <- price_data_num_acoes %>%
  dplyr::filter(num_acoes > 0) %>%
  dplyr::filter(!acoes_sa %in% c("MMAQ3.SA", "MMAQ4.SA")) %>%
  dplyr::mutate(ano = lubridate::year(date)) %>%
```

```

dplyr::group_by(ano, acoes_sa) %>%
dplyr::mutate(media_volume = mean(volume, na.rm = TRUE)) %>%
as.data.frame() %>%
dplyr::filter(media_volume >= 50000) %>%
dplyr::mutate(marketcap = close*num_acoes) %>%
dplyr::relocate(num_acoes, .after = adjusted)

Market_cap <- price_data_num_acoes

#####
##### Calculando Fator SMB -----
#####

# Obtendo o market cap do último dia do ano -----
Market_cap <- Market_cap %>%
  dplyr::group_by(ano, acoes_sa) %>%
  slice(tail(row_number(), 1)) %>%
  as.data.frame()

# Calculando os decis de cada ativo -----
Market_cap <- Market_cap %>%
  dplyr::group_by(ano) %>%
  dplyr::mutate(Decil = ntile(-adjusted, 10)) %>%
  as.data.frame() %>%
  dplyr::select(acoes_sa, Decil, ano)

Market_cap_price <- dplyr::inner_join(price_data_num_acoes,
                                     Market_cap,
                                     by = c("acoes_sa", "ano"))

# Obtendo o preço médio de cada decil por data -----
Decil_Price <- Market_cap_price %>%
  dplyr::group_by(date, Decil) %>%
  dplyr::summarise(Mean_Price = mean(adjusted))

# Criando as variáveis long e short -----
Decil_Price <- Decil_Price %>%
  dplyr::mutate(Buy_or_Short = case_when(Decil <= 3 ~ "Short", Decil >= 8 ~ "Long" ))

# Agrupando por data e variável (long/short) e obtendo a média de preço -----
Decil_Price <- Decil_Price %>%
  dplyr::group_by(date, Buy_or_Short) %>%
  dplyr::summarise(Mean_Price_Class = mean(Mean_Price))

# Removendo NA's (ativos que não são usados na estratégia long/short) -----
Decil_Price <- Decil_Price %>% drop_na(Buy_or_Short)

# Agrupando por variável e obtendo o retorno de cada estratégia -----
Decil_Price <- Decil_Price %>%
  dplyr::group_by(Buy_or_Short) %>%
  dplyr::mutate(Return = Mean_Price_Class/lag(Mean_Price_Class)-1) %>%
  dplyr::select(date, Buy_or_Short, Return)

```



```

# Pivotando o banco -----
Decil_Price <- Decil_Price %>% pivot_wider(names_from = Buy_or_Short , values_from = Return)

# Criando variável SMB -----
Decil_Price <- Decil_Price %>% dplyr::mutate(SMB = Long - Short)

#Decil_Price <- Decil_Price[-1,]

Decil_Price <- Decil_Price %>% drop_na(Long, SMB)

Decil_Price <- Decil_Price %>% dplyr::filter(SMB >= -1)
Decil_Price <- Decil_Price %>% dplyr::filter(SMB <= 1.5)

Decil_Price <- Decil_Price %>%
  dplyr::filter(!date %in% c('2014-01-02', '2015-01-02', '2016-01-02',
                             '2017-01-02', '2018-01-02', '2019-01-02',
                             '2020-01-02', '2018-05-02', '2018-05-11',
                             '2020-04-16', '2021-01-04'))

Decil_Price <- Decil_Price %>% mutate(return_acum = cumprod(1 + SMB)-1)

Decil_Price$date <- lubridate::ymd(Decil_Price$date)

# Gráfico -----
smb_plot <- ggplot(data = Decil_Price, aes(x = date, y = return_acum, group = "all")) +
  geom_line(size = 1.2, color = "deepskyblue3") +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 10),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", date_breaks = "6 months", expand = c(0,70)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                     breaks = seq(-0.9, 0.3, 0.15),
                     limits = c(-0.9,0.3))

### Salvando em PNG -----
ggsave(plot = smb_plot,
        paste0(pasta_graficos, "smb_graph_ff.png"),
        width = 14, height = 7, dpi = 250, units = "in",type = "cairo-png")

fator_smb <- Decil_Price %>%
  dplyr::select(date, SMB, return_acum) %>%
  'colnames<-'(c('date', 'SMB', 'SMB_acum'))

```

```
write.csv2(fator_smb,
           'fator_smb.csv', row.names = F)

### Salvando em PNG -----
ggsave(plot = smb_plot, "smb_graph_ff.png",
        width = 14, height = 7, dpi = 250, units = "in", type = "cairo-png")
```

## APÊNDICE F . CÓDIGO R: GERANDO FATOR HML - FAMA-FRENCH

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

library(tidyverse)
library(lubridate)
library(readxl)

options(scipen = 999)

#####
##### Carregando dados do MarketCap -----
#####

setwd("C:\\Users\\cleit\\Desktop\\TCC\\Dados\\MarketCap")

pasta_dados <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\'

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

# Carregando lista de ativos -----
indicadores_setores_v2 <- read.csv2(paste0(pasta_dados, "indicadores_setores_v3.csv"))

indicadores_setores_v2$Ações <- ifelse(
  indicadores_setores_v2$Ações == "",
  NA,
  indicadores_setores_v2$Ações)

indicadores_setores_v2 <- indicadores_setores_v2 %>% drop_na(Ações)

num_acoes_df <- read.csv2("num_acoes_df.csv")

num_acoes_df <- num_acoes_df %>%
  dplyr::mutate(acoes_sa = paste0(symbol, ".SA")) %>%
  dplyr::select(!symbol)

# Carregando dados de preços dos ativos até o dia 11/11/2021 -----
price_data <- read.csv2(paste0(pasta_dados, "acoes_setores_2022_04_21.csv"))

price_data_num_acoes <- inner_join(num_acoes_df,
  price_data, by = c("acoes_sa" = "symbol"))

price_data_num_acoes <- price_data_num_acoes %>%
```

```

dplyr::filter(num_acoes > 0) %>%
dplyr::filter(!acoes_sa %in% c("MMAQ3.SA", "MMAQ4.SA")) %>%
dplyr::mutate(ano = lubridate::year(date)) %>%
dplyr::group_by(ano, acoes_sa) %>%
dplyr::mutate(media_volume = mean(volume, na.rm = TRUE)) %>%
as.data.frame() %>%
dplyr::filter(media_volume >= 50000) %>%
dplyr::mutate(marketcap = close*num_acoes) %>%
dplyr::relocate(num_acoes, .after = adjusted)

rm(price_data, num_acoes_df, indicadores_setores_v2)

#####
##### Carregando dados do Patrimonio Liq -----
#####

# csv com CNPJ e ticket da ação -----
cnpj_acoes <- read.csv2('C:\\Users\\cleit\\Desktop\\TCC\\Dados\\cnpj_acoes.csv')

# Ajuste no CNPJ da Tim
cnpj_acoes$CNPJ <- ifelse(cnpj_acoes$CNPJ == '02.421.421/0001-11',
                        '02.558.115/0001-21', cnpj_acoes$CNPJ)
# Ajuste no CNPJ da Natura
cnpj_acoes$CNPJ <- ifelse(cnpj_acoes$CNPJ == '32.785.497/0001-97',
                        '71.673.990/0001-77', cnpj_acoes$CNPJ)
# Ajuste no CNPJ da Terra Santa
cnpj_acoes$CNPJ <- ifelse(cnpj_acoes$CNPJ == '40.337.136/0001-06',
                        '05.799.312/0001-20', cnpj_acoes$CNPJ)

# csv com CNPJ e cód CVM da ação -----
cnpj_cod_cvm_acoes <- read.csv2(paste0(pasta_dados, 'cnpj_cod_cvm_acoes.csv'))

# csv com PL da ação e cód CVM -----
patrimonio_liq_empresas <- read.csv2(paste0(pasta_dados, 'patrimonio_liq_empresas.csv'))

# Unindo PL e CNPJ das ações -----
patrimonio_liq_empresas_v2 <- patrimonio_liq_empresas %>%
  left_join(cnpj_cod_cvm_acoes, by = 'CD_CVM')

# Removendo CNPJ's ausentes -----
patrimonio_liq_empresas_v2 <- patrimonio_liq_empresas_v2 %>% drop_na(CNPJ)

# Unindo PL com ticket da ação -----
patrimonio_liq_empresas_v3 <- left_join(patrimonio_liq_empresas_v2,
                                       cnpj_acoes, by = 'CNPJ')

# Removendo ticket ausente -----
patrimonio_liq_empresas_v3 <- patrimonio_liq_empresas_v3 %>% drop_na(Ticker)

patrimonio_liq_empresas_v3 <- patrimonio_liq_empresas_v3 %>%
  dplyr::select(DT_REFER, Ticker, VL_CONTA) %>%
  dplyr::mutate(Ticker = paste0(Ticker, ".SA")) %>%
  dplyr::rename(acoes_sa = Ticker,
               ano = DT_REFER,
               patrim_liq = VL_CONTA)

```

```

# Obtendo o market cap no meio do ano -----
Market_cap <- price_data_num_acoes %>%
  dplyr::mutate(mes = lubridate::month(date)) %>%
  dplyr::mutate(mes_ano = paste0(mes,"_",ano)) %>%
  dplyr::filter(mes_ano %in% c('6_2011', '6_2012', '6_2013',
                              '6_2014', '6_2015', '6_2016',
                              '6_2017', '6_2018', '6_2019',
                              '6_2020', '6_2021')) %>%
  dplyr::group_by(ano, acoes_sa) %>%
  slice(tail(row_number(), 1)) %>%
  as.data.frame()

# Unindo dataframe do marketcap com df do PL -----
Market_cap_PL <- left_join(Market_cap,
                          patrimonio_liq_empresas_v3,
                          by = c('acoes_sa', 'ano'))

# Removendo ativos em PL -----
Market_cap_PL <- Market_cap_PL %>% drop_na(patrim_liq)

#####
##### Criando Indice HML -----
#####

# Criando indice BM -----
Market_cap_PL <- Market_cap_PL %>% dplyr::mutate(bm_index = patrim_liq/marketcap)

# Agrupando por ativo e ano para obter a média do BM em cada ano -----
Market_cap_PL_group <- Market_cap_PL %>%
  dplyr::group_by(ano, acoes_sa) %>%
  dplyr::summarise(bm_index = mean(bm_index, na.rm = TRUE)) %>%
  as.data.frame()

# Criando grupos do HML -----
Market_cap_PL_group_v2 <- Market_cap_PL_group %>%
  arrange(desc(bm_index)) %>%
  dplyr::group_by(ano) %>%
  dplyr::mutate(grupo = ntile(-bm_index, 10)) %>%
  as.data.frame() %>%
  dplyr::select(acoes_sa, bm_index, grupo, ano) %>%
  dplyr::mutate(grupo = case_when(grupo %in% 1:3 ~ 'High',
                                  grupo %in% 4:7 ~ 'Medium',
                                  grupo %in% 8:10 ~ 'Low'))

# Removendo ações intermediárias (Medium) -----
Market_cap_PL_group_v2 <- Market_cap_PL_group_v2 %>%
  dplyr::filter(grupo != 'Medium')

# Unindo fator HML com os ativos reais -----
Market_cap_PL_v3 <- dplyr::inner_join(price_data_num_acoes,
                                     Market_cap_PL_group_v2,
                                     by = c("acoes_sa", "ano"))

# Obtendo o preço médio de cada grupo por data -----

```

```

grupo_Price <- Market_cap_PL_v3 %>%
  dplyr::group_by(date, grupo) %>%
  dplyr::summarise(Mean_Price = mean(adjusted)) %>%
  as.data.frame()

# Removendo NA's (ativos que não são usados na estratégia long/short) -----
grupo_Price <- grupo_Price %>% drop_na(grupo)

# Agrupando por variável e obtendo o retorno de cada estratégia -----
grupo_Price <- grupo_Price %>%
  dplyr::group_by(grupo) %>%
  dplyr::mutate(Return = Mean_Price/lag(Mean_Price)-1) %>%
  as.data.frame() %>%
  dplyr::select(date, grupo, Return)

# Pivotando o banco -----
grupo_Price <- grupo_Price %>% pivot_wider(names_from = grupo , values_from = Return)

# Criando variável HML -----
grupo_Price <- grupo_Price %>% dplyr::mutate(HML = High - Low)

grupo_Price <- grupo_Price %>% drop_na(High, HML)

grupo_Price <- grupo_Price %>% dplyr::filter(HML >= -0.3)
grupo_Price <- grupo_Price %>% dplyr::filter(HML <= 0.3)

grupo_Price <- grupo_Price %>% dplyr::mutate(return_acum = cumprod(1 + HML)-1)

grupo_Price$date <- lubridate::ymd(grupo_Price$date)

hml_plot <- ggplot(data = grupo_Price, aes(x = date, y = return_acum, group = "all")) +
  geom_line(size = 1.2, color = "deepskyblue3") +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 10),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", date_breaks = "6 months", expand = c(0,70)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                     breaks = seq(-0.5,0.6,.1),
                     limits = c(-0.5,0.6))

### Salvando em PNG -----

```

```
ggsave(plot = hml_plot,  
        paste0(pasta_graficos, "hml_graph_ff.png"),  
        width = 14, height = 7, dpi = 250, units = "in", type = "cairo-png")  
  
fator_hml <- grupo_Price %>%  
  dplyr::select(date, HML, return_acum) %>%  
  'colnames<-'(c('date', 'HML', 'HML_acum'))  
  
write.csv2(fator_hml,  
           paste0(pasta_dados, 'fator_hml.csv'), row.names = F)
```

## APÊNDICE G . CÓDIGO R: GERANDO FATOR MERCADO - FAMA-FRENCH

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

library(tidyverse)
library(lubridate)
library(quantmod)

setwd('C:\\Users\\cleit\\Desktop\\TCC\\Dados')

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

pasta_dados <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\'

quantmod::getSymbols('^BVSP')

bovespa <- data.frame(date=index(BVSP), coredata(BVSP))

bovespa <- bovespa %>%
  dplyr::select(date, BVSP.Close, BVSP.Adjusted) %>%
  'colnames<-'(c('date', 'close', 'adjusted')) %>%
  dplyr::filter(date >= as.Date("2011-01-03") & date <= as.Date("2021-12-30"))

#####
##### Retorno Mensal -----
#####

bovespa_mensal <- bovespa %>%
  dplyr::mutate(mes = lubridate::month(date),
               ano = lubridate::year(date)) %>%
  dplyr::mutate(mes_ano = paste0(mes, "_", ano)) %>%
  dplyr::group_by(mes_ano) %>%
  dplyr::summarise(media_mes = mean(adjusted, na.rm = TRUE)) %>%
  as.data.frame()

benchmark = 1.1465/nrow(bovespa_mensal)

bovespa_mensal <- bovespa_mensal %>%
  dplyr::mutate(mes_ano = lubridate::dmy(paste0('01_', mes_ano))) %>%
  arrange(mes_ano) %>%
  dplyr::mutate(Return = media_mes/lag(media_mes)-1) %>%
  drop_na(Return) %>%
  dplyr::mutate(Return_menos_txa_livre = Return - benchmark) %>%
```



```

dplyr::mutate(return_acum = cumprod(1 + Return)-1,
              return_acum_menos_txa_livre = cumprod(1 + Return_menos_txa_livre)-1)

ggplot(data = bovespa_mensal, aes(x = mes_ano,
                                y = return_acum_menos_txa_livre, group = "all")) +
  geom_line(size = 1.2, color = "deepskyblue3") +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "",
       title = "Fator Mercado - Fama French",
       caption = "Fonte: Yahoo! Finance | Elaboração: Cleiton Otavio da Exaltação Rocha") +
  theme_minimal() +
  theme(legend.position = "top",
        legend.title = element_text(size = 10),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size=17, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size=10),
        axis.text.x = element_text(face="bold", color="#000000",size=10),
        axis.title.y = element_text(size = 10)) +
  scale_x_date(date_labels = "%b\n%Y", date_breaks = "1 years", expand = c(0,50)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(-0.8,0.1,.1),
                    limits = c(-0.8,0.1))

#####
##### Retorno Diário -----
#####

benchmark_diario = 1.1465/nrow(bovespa)

bovespa_diario <- bovespa %>%
  dplyr::mutate(Return = adjusted/lag(adjusted)-1) %>%
  drop_na(Return) %>%
  dplyr::mutate(Return_menos_txa_livre = Return - benchmark_diario) %>%
  dplyr::mutate(return_acum = cumprod(1 + Return)-1,
                return_acum_menos_txa_livre = cumprod(1 + Return_menos_txa_livre)-1)

fator_mercado <- ggplot(data = bovespa_diario, aes(x = date,
                                                  y = return_acum_menos_txa_livre, group = "all")) +
  geom_line(size = 1.2, color = "deepskyblue3") +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 10),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),

```

```

    axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", date_breaks = "6 months", expand = c(0,70)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
    breaks = seq(-0.8,0.1,.1),
    limits = c(-0.8,0.1))

### Salvando em PNG -----
ggsave(plot = fator_mercado,
  paste0(pasta_graficos, "fator_mercado_graph_ff.png"),
  width = 14, height = 7, dpi = 250, units = "in",type = "cairo-png")

fator_mercado <- bovespa_diario %>%
  dplyr::select(date,
    Return_menos_txa_livre,
    return_acum_menos_txa_livre) %>%
  'colnames<-'(c('date', 'Market', 'Market_acum'))

write.csv2(fator_mercado,
  paste0(pasta_dados, 'fator_mercado.csv'), row.names = F)

```

## APÊNDICE H . CÓDIGO R: GERANDO CARTEIRA FAMA-FRENCH

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

library(tidyverse)
library(lubridate)
library(timetk)
library(PerformanceAnalytics)

options(scipen = 999)

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

pasta_dados <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\'

setwd('C:\\Users\\cleit\\Desktop\\TCC\\Dados')

fator_hml <- read.csv2('fator_hml.csv')
fator_smb <- read.csv2('fator_smb.csv')
fator_mercado <- read.csv2('fator_mercado.csv')

fator_hml$date <- lubridate::ymd(fator_hml$date)
fator_smb$date <- lubridate::ymd(fator_smb$date)
fator_mercado$date <- lubridate::ymd(fator_mercado$date)

fatores_ff <- purrr::reduce(list(fator_hml, fator_smb, fator_mercado),
                           dplyr::inner_join, by = 'date')

fatores_ff$ano <- lubridate::year(fatores_ff$date)

fatores_ff <- fatores_ff %>% dplyr::rename('Rm_minus_Rf' = 'Market')

fatores_ff <- fatores_ff %>% dplyr::select(-c("HML_acum", "SMB_acum", "Market_acum"))

rm(fator_hml, fator_mercado, fator_smb)

factor_df_train <- fatores_ff %>%
  dplyr::filter(date >= as.Date("2011-01-03") & date <= as.Date("2018-09-10"))

factor_df_test <- fatores_ff %>%
  dplyr::filter(date > as.Date("2018-09-10"))

portfolio <- factor_df_test %>% dplyr::select(-c('ano'))
```

```

portfolio <- portfolio %>% timetk::tk_xts()

portfolio_retorno <- PerformanceAnalytics::Return.portfolio(portfolio,
                                                             weights = rep(.33,3))

portfolio_retorno <- data.frame(date=index(portfolio_retorno),
                               coredata(portfolio_retorno))

portfolio_retorno <- portfolio_retorno %>%
  dplyr::mutate(return_acum = cumprod(1 + portfolio.returns)-1)

ff_portfolio <- portfolio_retorno %>%
  ggplot(aes(x = date, y = return_acum, group = "all")) +
  geom_line(size = 1.2, color = "deepskyblue3") +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "") +
  theme_classic() +
  theme(legend.position = "none",
        legend.title = element_text(size = 10),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 13),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", date_breaks = "2 months", expand = c(0,50)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(-0.3,0.15,.05),
                    limits = c(-0.3,0.15))

### Salvando em PNG -----
ggsave(plot = ff_portfolio,
        paste0(pasta_graficos, "ff_portfolio_retorno.png"),
        width = 14, height = 7, dpi = 250, units = "in",type = "cairo-png")

portfolio_retorno <- portfolio_retorno %>%
  dplyr::select(date, return_acum) %>%
  'colnames<-'(c('date', 'fama_french'))

write.csv2(portfolio_retorno,
           paste0(pasta_dados, 'fama_french_portfolio.csv'), row.names = F)

```

## APÊNDICE I . CÓDIGO R: UNINDO TODAS AS CARTEIRAS SIMULADAS

```
#####
#####
## ESCRITO POR: Cleiton Rocha - www.linkedin.com/in/cleitonoerocha/
## EMAIL: cleitonotavio058@gmail.com
## LICENÇA: GPLv3
## DATA: 05/07/2022
## CÓDIGO DISPONÍVEL EM:
## https://github.com/CleitonOERocha/tcc-finance-codes
#####
#####

library(tidyverse)
library(lubridate)
library(Cairo)

setwd("C:\\Users\\cleit\\Desktop\\TCC\\Dados")

pasta_graficos <- 'C:\\Users\\cleit\\Desktop\\TCC\\Dados\\Graficos\\'

#####
##### Carteiras de Markowitz e Ibovespa -----
#####

markowitz <- read.csv2("retorno_acum_ibov_markowitz_all.csv")

markowitz <- markowitz %>%
  dplyr::mutate(date = lubridate::ymd(date)) %>%
  dplyr::filter(date <= "2021-12-31")

markowitz$date <- lubridate::ymd(markowitz$date)

#####
##### Carteiras de Fama French -----
#####

fama_french <- read.csv2('fama_french_portfolio.csv')

fama_french <- fama_french %>%
  dplyr::mutate(tipo = 'Fama-French') %>%
  dplyr::rename(media_ret = fama_french)

fama_french$date <- lubridate::ymd(fama_french$date)

# Unindo Markowitz e Fama-French -----

carteiras <- bind_rows(fama_french, markowitz)

#####
##### Taxa Livre de Risco -----
#####
```

```

# https://calculadorarendafixa.com.br#
# https://www.contadordedias.com.br/contador-de-dias/

benchmark = (1.15793155-1)/1207

txa_livre_risco <- data.frame(date = seq(as.Date('2018-09-10'),as.Date('2021-12-30'),1))

txa_livre_risco <- txa_livre_risco %>%
  dplyr::mutate(txa_livre_risco = benchmark) %>%
  dplyr::mutate(txa_livre_risco = cumsum(txa_livre_risco)) %>%
  'colnames<-'(c('date', 'media_ret')) %>%
  dplyr::mutate(tipo = 'Taxa Livre de Risco')

todos_ativos <- bind_rows(carteiras, txa_livre_risco)

todos_ativos$tipo <- ifelse(
  todos_ativos$tipo == 'Carteira Markowitz - SmallCaps', 'Markowitz - SmallCaps',
  todos_ativos$tipo)

todos_ativos$tipo <- ifelse(
  todos_ativos$tipo == 'Carteira Markowitz',
  'Markowitz',
  todos_ativos$tipo)

ret_acum_graph <- ggplot(todos_ativos, aes(x = date, y = media_ret, colour = tipo)) +
  geom_line(size = 1.3) +
  geom_hline(yintercept = 0, linetype = 'dashed') +
  labs(x="", y="Retorno Acumulado (%)",
       colour = "") +
  theme_classic() +
  theme(legend.position = "top",
        legend.title = element_text(size = 10),
        plot.caption = element_text(size = 7),
        legend.text = element_text(size = 15),
        strip.text = element_text(face = "bold", size = "11"),
        plot.title = element_text(size = 13, color = "dodgerblue4", hjust = 0.5, face="bold"),
        plot.subtitle = element_text(hjust = 0.5, face="bold", size = 12),
        axis.text.y = element_text(face="bold", color="#000000",size = 14),
        axis.text.x = element_text(face="bold", color="#000000",size = 14),
        axis.title.y = element_text(size = 16),
        axis.title.x = element_text(size = 16)) +
  scale_x_date(date_labels = "%b\n%Y", breaks = "3 month", expand = c(0,15)) +
  scale_y_continuous(labels = scales::percent_format(big.mark = ".", decimal.mark = ","),
                    breaks = seq(-0.3,1.20,0.15),
                    limits = c(-0.3,1.20)) +
  scale_colour_manual(values = c("#ff6361", "#ffa600", "#003f5c", "#bc5090", "deepskyblue3"),
                    drop = FALSE,
                    guide = guide_legend(
                      direction = "horizontal",
                      keyheight = unit(2, units = "mm"),
                      keywidth = unit(60/length(labels), units = "mm"),
                      title.position = 'top',
                      title.hjust = 0.5,
                      label.hjust = 0.5,
                      nrow = 1,
                      byrow = TRUE,

```

```
reverse = FALSE,  
label.position = "top"))
```

```
### Salvando em PNG ###
```

```
ggsave(plot = ret_acum_graph,  
        paste0(pasta_graficos, "all_metodos.png"),  
        width = 14, height = 7, dpi = 250, units = "in", type = "cairo-png")
```