

Trabalho final ciência de dados

Yuri Fonseca Lopes

Mudanças em relação aos anteriores

- Novo Banco de dados
- Novo somatório, agora ponderado
- Tentar corroborar com a hipótese do artigo
 - Empresas com alto DR tem retornos menores comparados aos demais
- Tentar encontrar valores de AC, CA, DR, OFC, ECI que indicam empresas com alto retorno anual (r.12.4)
- Entre outras pequenas mudanças (descritas no .ipynb)

Banco de dados

Index = Data

Firm = Número da firma (cada firma recebeu um número, temos diferentes amostras para a mesma firma)

BE = Patrimônio líquido das empresas

ME = Valor de mercado

r.12.4 = Retorno acumulado nos últimos 12 meses

AC = Acionistas

CA = Conselho e Administração

DR = Diretoria

OFC = Órgãos de fiscalização e Controle

ECI = Ética e conflitos de interesse

Industry = Tipo de indústria, podendo conter: Agro e Pesca, Alimentos e Bebidas, Comércio, Construção, Eletro-eletrônicos, Energia Elétrica, Máquinas Industriais, Mineração, Minerais não Metálicos, Outros, Papel e Celulose, Petróleo e Gás, Química, Siderurgia e Metalurgia, Software e Dados, Telecomunicações, Têxtil, Transporte Serviços, Veículos e peças, “Finanças e Seguros” (financeiras)

Banco de dados

Foi disponibilizado uma nova database que já realiza essa soma. Essa nova soma ponderada contém um mapeamento das respostas das empresas, seguindo:

- - 0 -> não se aplica
- - 1 -> não
- - 2 -> parcialmente
- - 3 -> sim

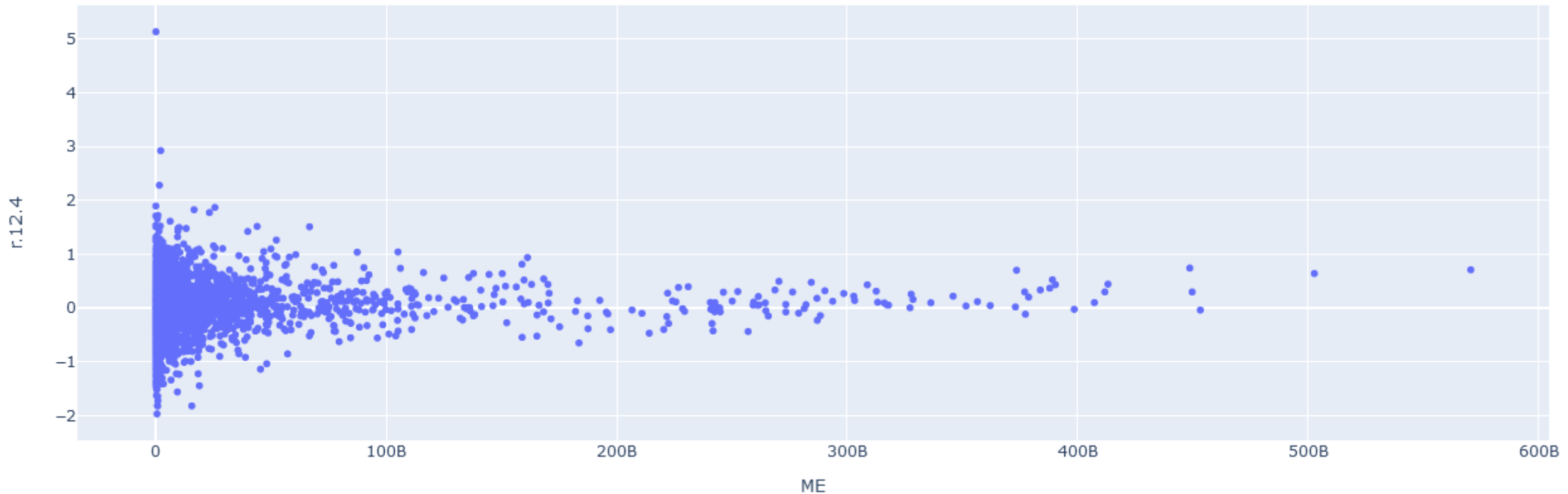
$$GC_i = \left(\frac{1}{N_i \cdot W_{max}} \right) \cdot \sum_{j=1}^{N_i} w_{ij}$$

"Em que GC i é indicador de governança corporativa padronizado, correspondente ao capítulo i; Ni representa o número total de perguntas no questionário do capítulo i; wij refere-se ao peso atribuído à resposta j no capítulo i, de acordo com a escala escolhida; Wmax representa o peso máximo atribuído a uma resposta, conforme definido pela escala escolhida, ou seja, o valor é igual a 3 para este estudo".

Exploração de dados

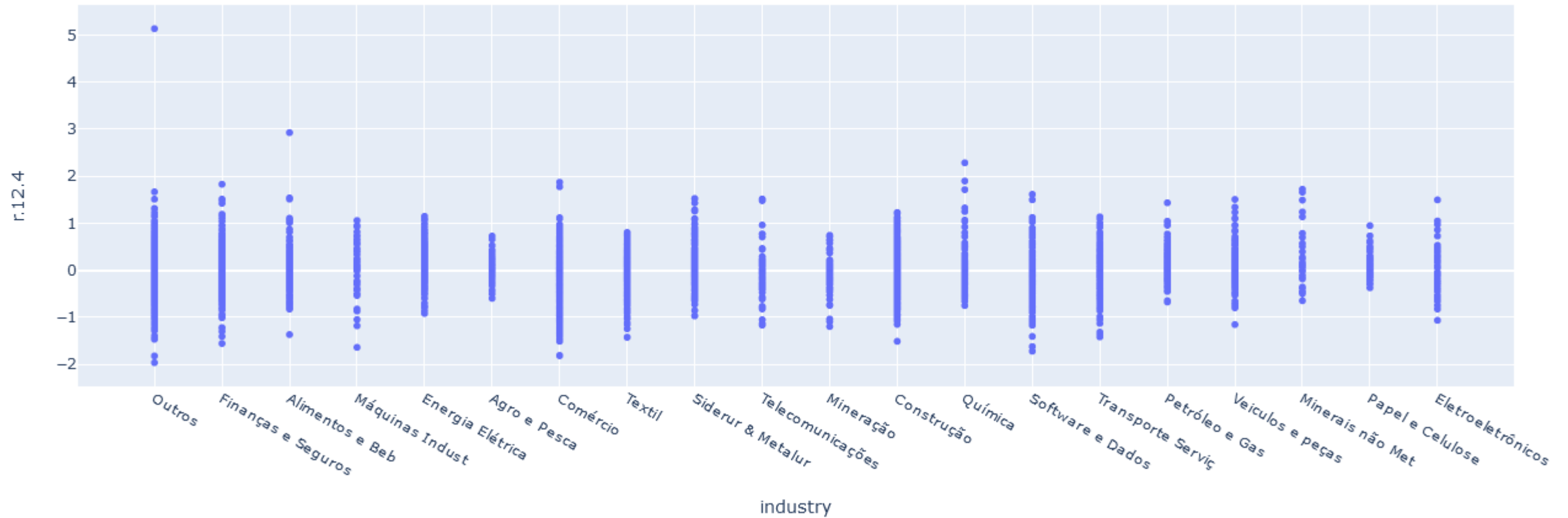
- Não temos valores faltantes

Market Value por Retorno Anual



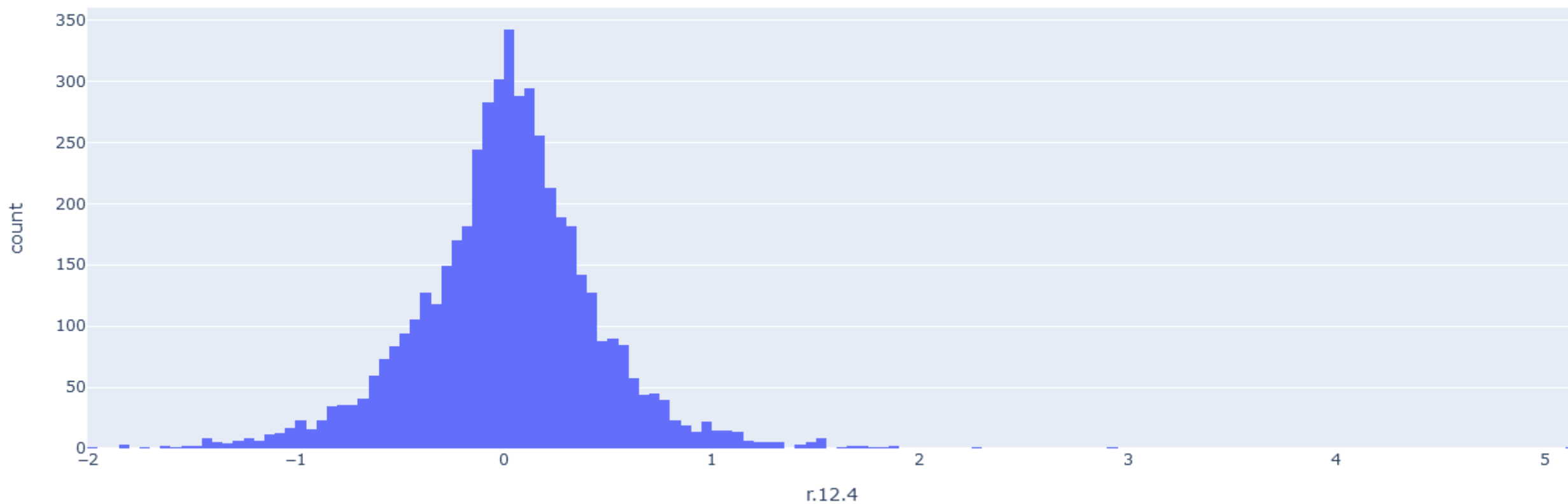
Exploração de dados

Tipo de industria por Retorno Anual



Exploração de dados

Histograma dos valores



Estatística descritiva

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	4.944000e+03	4.944000e+03	4944.000000	4944.000000	4944.000000	4944.000000	4944.000000	4944.000000
mean	9.244580e+09	1.647379e+10	0.017635	0.528727	0.781031	0.845292	0.710720	0.764979
std	2.990092e+10	4.501468e+10	0.441591	0.148566	0.130017	0.135652	0.165161	0.139815
min	5.600000e+04	1.597143e+06	-1.960695	0.194444	0.361111	0.333333	0.233333	0.305556
25%	7.253755e+08	8.281819e+08	-0.211830	0.416667	0.694444	0.791667	0.600000	0.694444
50%	1.794868e+09	3.431592e+09	0.021291	0.527778	0.805556	0.875000	0.733333	0.777778
75%	6.121884e+09	1.161053e+10	0.254705	0.638889	0.888889	0.958333	0.833333	0.861111
max	4.341940e+11	5.705210e+11	5.134580	0.833333	1.000000	1.000000	1.000000	1.000000

Clusterização

- O que é kmeans?

Dropei a coluna industry, ME_lag1, r_i, r.3.0

Separei por ano

Normalizei os dados, clusterizei usando AC, CA, DR, OFC, ECI

Reduzi a dimensionalidade

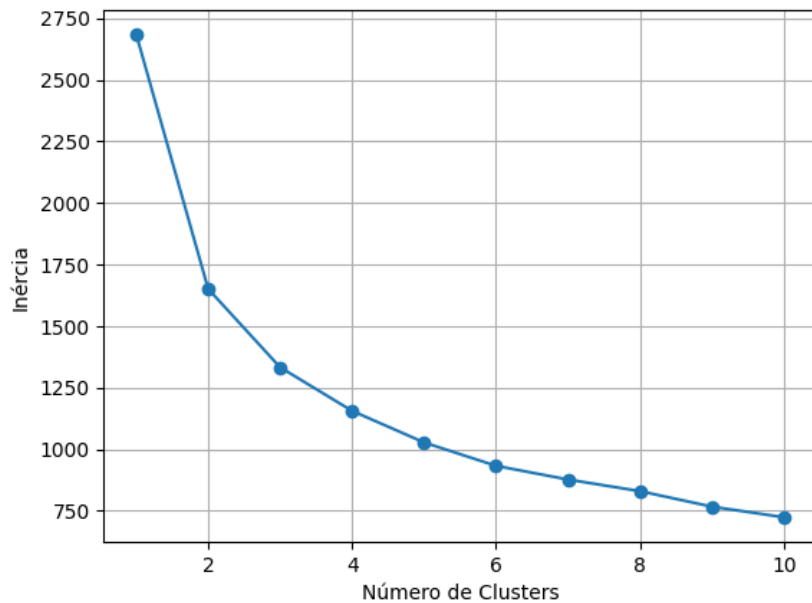
```
[84] # @title Normalizando os dados e clusterizando
```

```
def processar_dados(df_ano, n_clusters=3):  
    # Normalizando os dados  
    scaler = StandardScaler()  
    values = scaler.fit_transform(df_ano[["AC", "CA", "DR", "OFC", "ECI"]])  
  
    # Clusterizando com KMeans  
    kmeans = KMeans(n_clusters=n_clusters, random_state=42, n_init=10)  
    clusters = kmeans.fit_predict(values)  
  
    # Reduzindo dimensionalidade com PCA  
    pca = PCA(n_components=2)  
    X_reduced = pca.fit_transform(values)  
  
    # Convertendo X_reduced para DataFrame e adicionando as colunas originais  
    df_reduced = pd.DataFrame(X_reduced, columns=['PC1', 'PC2'])  
    df_reduced['Cluster'] = clusters.astype(str)  
    df_reduced['Index'] = df_ano['Index'].values  
    df_reduced['firm'] = df_ano['firm'].values  
    df_reduced['BE'] = df_ano['BE'].values  
    df_reduced['ME'] = df_ano['ME'].values  
    df_reduced['r.12.4'] = df_ano['r.12.4'].values  
    df_reduced['AC'] = df_ano['AC'].values  
    df_reduced['CA'] = df_ano['CA'].values  
    df_reduced['DR'] = df_ano['DR'].values  
    df_reduced['OFC'] = df_ano['OFC'].values  
    df_reduced['ECI'] = df_ano['ECI'].values  
  
    # Retorna o DataFrame reduzido com informações completas  
    return df_reduced  
  
df_2019_reduced = processar_dados(df_2019)  
df_2020_reduced = processar_dados(df_2020)  
df_2021_reduced = processar_dados(df_2021)  
df_2022_reduced = processar_dados(df_2022)  
df_2023_reduced = processar_dados(df_2023)
```

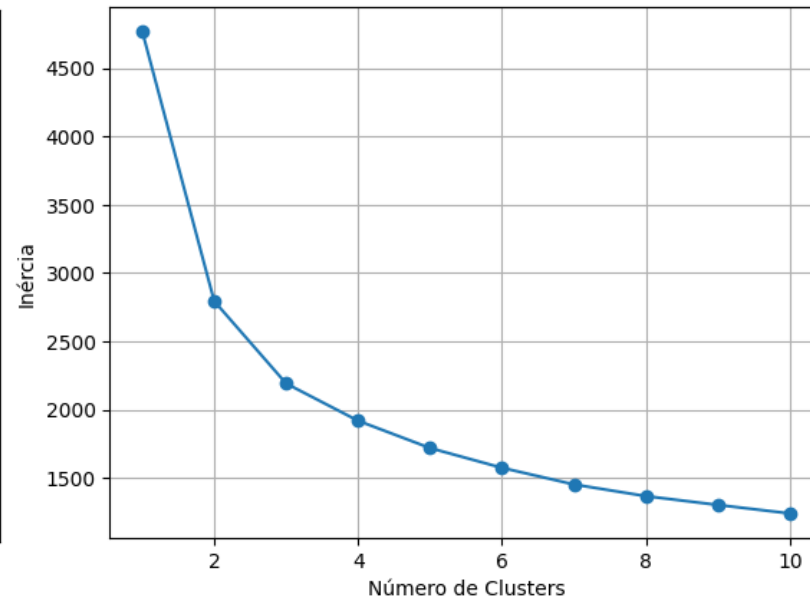
Hiperparâmetros

- Utilizei o teste do cotovelo, silhueta, e índice Calinski-hazabasz
Cotovelo sugeriu 3 clusters, os demais 2. Optei por 3
- Fiz os testes para todos os dataframes divididos por ano, apresentando o mesmo resultado.

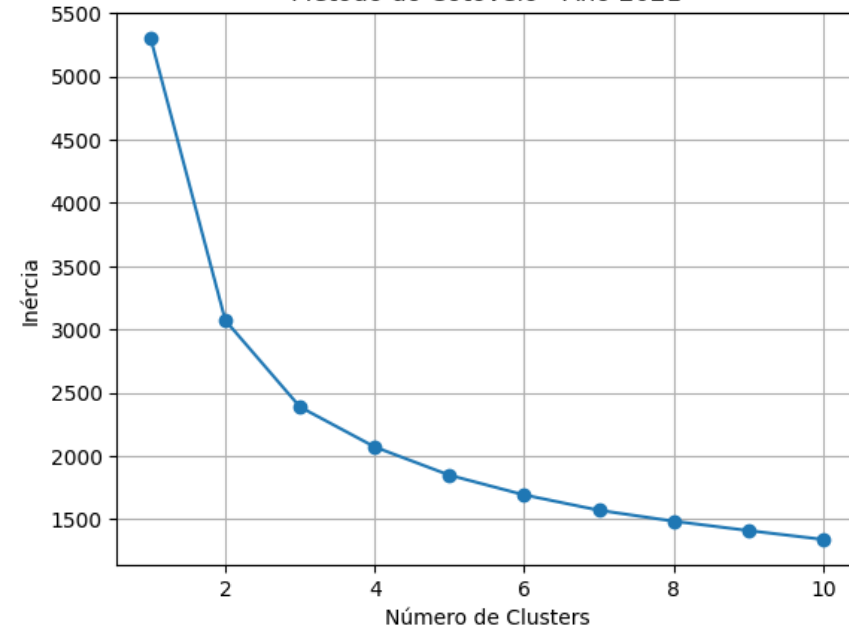
Método do Cotovelo - Ano 2019



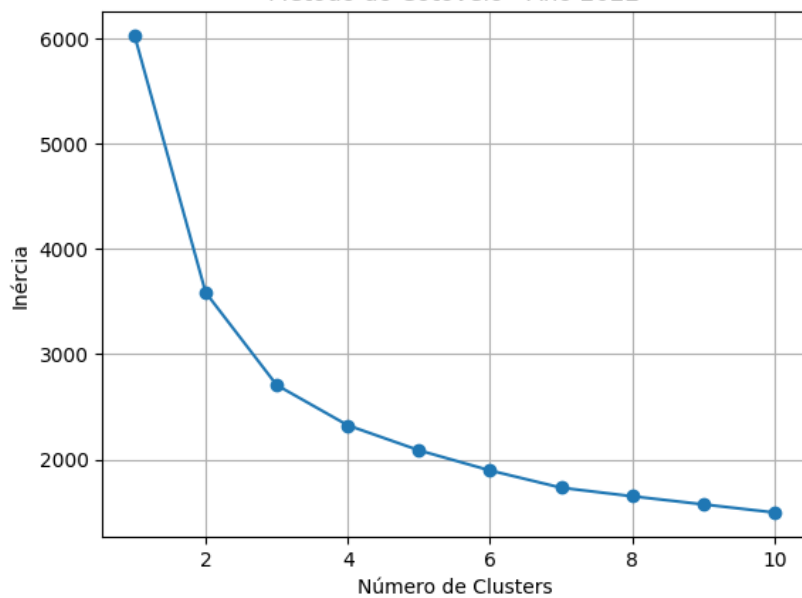
Método do Cotovelo - Ano 2020



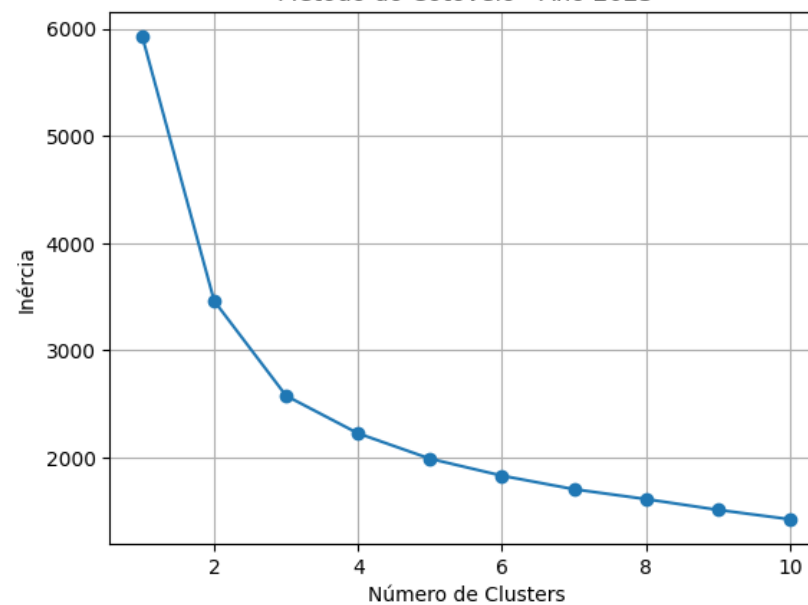
Método do Cotovelo - Ano 2021



Método do Cotovelo - Ano 2022



Método do Cotovelo - Ano 2023



```
Ano 2020: Para n_clusters = 2, a pontuação média da silhueta é: 0.3449
Ano 2020: Para n_clusters = 3, a pontuação média da silhueta é: 0.2419
Ano 2020: Para n_clusters = 4, a pontuação média da silhueta é: 0.2468
Ano 2020: Para n_clusters = 5, a pontuação média da silhueta é: 0.2294
Ano 2020: Para n_clusters = 6, a pontuação média da silhueta é: 0.2327
Ano 2020: Para n_clusters = 7, a pontuação média da silhueta é: 0.2304
Ano 2020: Para n_clusters = 8, a pontuação média da silhueta é: 0.2298
Ano 2020: Para n_clusters = 9, a pontuação média da silhueta é: 0.2351
Ano 2020: Para n_clusters = 10, a pontuação média da silhueta é: 0.2424
-----
Ano 2021: Para n_clusters = 2, a pontuação média da silhueta é: 0.3422
Ano 2021: Para n_clusters = 3, a pontuação média da silhueta é: 0.2651
Ano 2021: Para n_clusters = 4, a pontuação média da silhueta é: 0.2569
Ano 2021: Para n_clusters = 5, a pontuação média da silhueta é: 0.2315
Ano 2021: Para n_clusters = 6, a pontuação média da silhueta é: 0.2305
Ano 2021: Para n_clusters = 7, a pontuação média da silhueta é: 0.2224
Ano 2021: Para n_clusters = 8, a pontuação média da silhueta é: 0.2163
Ano 2021: Para n_clusters = 9, a pontuação média da silhueta é: 0.2256
Ano 2021: Para n_clusters = 10, a pontuação média da silhueta é: 0.2136
-----
Ano 2021: Para n_clusters = 2, a pontuação média da silhueta é: 0.3505
Ano 2021: Para n_clusters = 3, a pontuação média da silhueta é: 0.2704
Ano 2021: Para n_clusters = 4, a pontuação média da silhueta é: 0.2366
Ano 2021: Para n_clusters = 5, a pontuação média da silhueta é: 0.2420
Ano 2021: Para n_clusters = 6, a pontuação média da silhueta é: 0.2377
Ano 2021: Para n_clusters = 7, a pontuação média da silhueta é: 0.2292
Ano 2021: Para n_clusters = 8, a pontuação média da silhueta é: 0.2297
Ano 2021: Para n_clusters = 9, a pontuação média da silhueta é: 0.2078
Ano 2021: Para n_clusters = 10, a pontuação média da silhueta é: 0.2084
```

```
-----
Ano 2022: Para n_clusters = 2, a pontuação média da silhueta é: 0.3746
Ano 2022: Para n_clusters = 3, a pontuação média da silhueta é: 0.2815
Ano 2022: Para n_clusters = 4, a pontuação média da silhueta é: 0.2540
Ano 2022: Para n_clusters = 5, a pontuação média da silhueta é: 0.2493
Ano 2022: Para n_clusters = 6, a pontuação média da silhueta é: 0.2358
Ano 2022: Para n_clusters = 7, a pontuação média da silhueta é: 0.2413
Ano 2022: Para n_clusters = 8, a pontuação média da silhueta é: 0.2199
Ano 2022: Para n_clusters = 9, a pontuação média da silhueta é: 0.2122
Ano 2022: Para n_clusters = 10, a pontuação média da silhueta é: 0.2102
-----
Ano 2023: Para n_clusters = 2, a pontuação média da silhueta é: 0.4078
Ano 2023: Para n_clusters = 3, a pontuação média da silhueta é: 0.2945
Ano 2023: Para n_clusters = 4, a pontuação média da silhueta é: 0.2779
Ano 2023: Para n_clusters = 5, a pontuação média da silhueta é: 0.2382
Ano 2023: Para n_clusters = 6, a pontuação média da silhueta é: 0.2287
Ano 2023: Para n_clusters = 7, a pontuação média da silhueta é: 0.2242
Ano 2023: Para n_clusters = 8, a pontuação média da silhueta é: 0.2207
Ano 2023: Para n_clusters = 9, a pontuação média da silhueta é: 0.2170
Ano 2023: Para n_clusters = 10, a pontuação média da silhueta é: 0.2172
-----
```

```
Ano 2020: Para n_clusters = 2, o índice Calinski-Harabasz é: 334.9935
Ano 2020: Para n_clusters = 3, o índice Calinski-Harabasz é: 270.9471
Ano 2020: Para n_clusters = 4, o índice Calinski-Harabasz é: 235.0455
Ano 2020: Para n_clusters = 5, o índice Calinski-Harabasz é: 214.5065
Ano 2020: Para n_clusters = 6, o índice Calinski-Harabasz é: 199.5362
Ano 2020: Para n_clusters = 7, o índice Calinski-Harabasz é: 182.0813
Ano 2020: Para n_clusters = 8, o índice Calinski-Harabasz é: 168.9293
Ano 2020: Para n_clusters = 9, o índice Calinski-Harabasz é: 165.0652
Ano 2020: Para n_clusters = 10, o índice Calinski-Harabasz é: 158.6647
-----
Ano 2021: Para n_clusters = 2, o índice Calinski-Harabasz é: 672.6976
Ano 2021: Para n_clusters = 3, o índice Calinski-Harabasz é: 559.3189
Ano 2021: Para n_clusters = 4, o índice Calinski-Harabasz é: 470.3698
Ano 2021: Para n_clusters = 5, o índice Calinski-Harabasz é: 420.9120
Ano 2021: Para n_clusters = 6, o índice Calinski-Harabasz é: 384.6631
Ano 2021: Para n_clusters = 7, o índice Calinski-Harabasz é: 360.2619
Ano 2021: Para n_clusters = 8, o índice Calinski-Harabasz é: 336.2185
Ano 2021: Para n_clusters = 9, o índice Calinski-Harabasz é: 314.2177
Ano 2021: Para n_clusters = 10, o índice Calinski-Harabasz é: 298.2409
-----
Ano 2021: Para n_clusters = 2, o índice Calinski-Harabasz é: 769.9524
Ano 2021: Para n_clusters = 3, o índice Calinski-Harabasz é: 646.9303
Ano 2021: Para n_clusters = 4, o índice Calinski-Harabasz é: 550.4566
Ano 2021: Para n_clusters = 5, o índice Calinski-Harabasz é: 494.2221
Ano 2021: Para n_clusters = 6, o índice Calinski-Harabasz é: 450.8816
Ano 2021: Para n_clusters = 7, o índice Calinski-Harabasz é: 418.0029
Ano 2021: Para n_clusters = 8, o índice Calinski-Harabasz é: 387.6037
Ano 2021: Para n_clusters = 9, o índice Calinski-Harabasz é: 363.5530
Ano 2021: Para n_clusters = 10, o índice Calinski-Harabasz é: 345.9045
```

```
-----
Ano 2022: Para n_clusters = 2, o índice Calinski-Harabasz é: 820.2509
Ano 2022: Para n_clusters = 3, o índice Calinski-Harabasz é: 738.7654
Ano 2022: Para n_clusters = 4, o índice Calinski-Harabasz é: 638.8838
Ano 2022: Para n_clusters = 5, o índice Calinski-Harabasz é: 567.5228
Ano 2022: Para n_clusters = 6, o índice Calinski-Harabasz é: 523.7092
Ano 2022: Para n_clusters = 7, o índice Calinski-Harabasz é: 496.2798
Ano 2022: Para n_clusters = 8, o índice Calinski-Harabasz é: 454.4058
Ano 2022: Para n_clusters = 9, o índice Calinski-Harabasz é: 424.0663
Ano 2022: Para n_clusters = 10, o índice Calinski-Harabasz é: 403.0057
-----
Ano 2023: Para n_clusters = 2, o índice Calinski-Harabasz é: 844.1187
Ano 2023: Para n_clusters = 3, o índice Calinski-Harabasz é: 769.3752
Ano 2023: Para n_clusters = 4, o índice Calinski-Harabasz é: 655.1923
Ano 2023: Para n_clusters = 5, o índice Calinski-Harabasz é: 584.2954
Ano 2023: Para n_clusters = 6, o índice Calinski-Harabasz é: 527.8621
Ano 2023: Para n_clusters = 7, o índice Calinski-Harabasz é: 486.5714
Ano 2023: Para n_clusters = 8, o índice Calinski-Harabasz é: 450.0246
Ano 2023: Para n_clusters = 9, o índice Calinski-Harabasz é: 429.3226
Ano 2023: Para n_clusters = 10, o índice Calinski-Harabasz é: 412.7578
-----
```

Descrição estatística dos clusters em 2019

Resumo estatístico do Cluster 0:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	2.500000e+02	2.500000e+02	250.000000	250.000000	250.000000	250.000000	250.000000	250.000000
mean	7.483088e+09	1.340651e+10	0.141880	0.492556	0.756333	0.833667	0.677867	0.756000
std	1.387826e+10	2.496315e+10	0.356733	0.123036	0.090927	0.080611	0.134130	0.091137

Resumo estatístico do Cluster 1:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.840000e+02	1.840000e+02	184.000000	184.000000	184.000000	184.000000	184.000000	184.000000
mean	2.531150e+10	4.861361e+10	0.173358	0.650966	0.877415	0.918252	0.820833	0.860356
std	5.377683e+10	8.795234e+10	0.323879	0.109721	0.067500	0.080480	0.082652	0.079497

Resumo estatístico do Cluster 2:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.030000e+02	1.030000e+02	103.000000	103.000000	103.000000	103.000000	103.000000	103.000000
mean	2.089001e+09	3.829683e+09	0.216986	0.433118	0.591424	0.641990	0.519417	0.566073
std	2.585099e+09	6.093260e+09	0.597280	0.125765	0.065623	0.112598	0.134757	0.118115

Melhor cluster
de 2019: 2

Descrição estatística dos clusters em 2020

Resumo estatístico do Cluster 2:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	4.250000e+02	4.250000e+02	425.000000	425.000000	425.000000	425.000000	425.000000	425.000000
mean	4.516746e+09	8.899557e+09	0.092116	0.467451	0.730327	0.823235	0.662118	0.741307
std	9.110240e+09	1.582124e+10	0.426116	0.111621	0.077281	0.092464	0.127487	0.100203

Resumo estatístico do Cluster 1:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	3.300000e+02	3.300000e+02	330.000000	330.000000	330.000000	330.000000	330.000000	330.000000
mean	1.753285e+10	3.411578e+10	0.096653	0.629040	0.869529	0.917803	0.812626	0.841751
std	4.200570e+10	6.625173e+10	0.404347	0.118679	0.072638	0.082466	0.097133	0.084072

Resumo estatístico do Cluster 0:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.990000e+02	1.990000e+02	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000
mean	7.916466e+08	1.481754e+09	0.183863	0.400893	0.578029	0.630653	0.480067	0.527778
std	1.068026e+09	2.552974e+09	0.398945	0.119376	0.067140	0.112575	0.118204	0.113230

Melhor cluster
de 2020: 0

Descrição estatística dos clusters em 2021

Resumo estatístico do Cluster 2:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	4.840000e+02	4.840000e+02	484.000000	484.000000	484.000000	484.000000	484.000000	484.000000
mean	3.374642e+09	8.439469e+09	0.155366	0.483930	0.750057	0.837896	0.670248	0.758953
std	5.991467e+09	1.587573e+10	0.424468	0.115469	0.073471	0.091931	0.125131	0.091961

=====

Resumo estatístico do Cluster 0:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	3.760000e+02	3.760000e+02	376.000000	376.000000	376.000000	376.000000	376.000000	376.000000
mean	1.929121e+10	3.841910e+10	0.189622	0.620493	0.892878	0.938830	0.840426	0.862072
std	4.486150e+10	7.383200e+10	0.449338	0.122459	0.065894	0.071803	0.086358	0.072953

=====

Resumo estatístico do Cluster 1:

	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	2.010000e+02	2.010000e+02	201.000000	201.000000	201.000000	201.000000	201.000000	201.000000
mean	9.011930e+08	1.885289e+09	0.230655	0.413626	0.589414	0.631841	0.492040	0.543809
std	1.150390e+09	2.743490e+09	0.448055	0.122500	0.079455	0.109671	0.122441	0.116220

Melhor cluster
de 2021: 1

Descrição estatística dos clusters em 2022

Resumo estatístico do Cluster 2:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	5.840000e+02	5.840000e+02	584.000000	584.000000	584.000000	584.000000	584.000000	584.000000
mean	4.641575e+09	8.658350e+09	-0.163684	0.491581	0.767837	0.855808	0.688699	0.769216
std	9.491246e+09	1.635611e+10	0.424894	0.120502	0.072269	0.082274	0.115014	0.088296
=====								
Resumo estatístico do Cluster 0:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	4.390000e+02	4.390000e+02	439.000000	439.000000	439.000000	439.000000	439.000000	439.000000
mean	1.770474e+10	2.875809e+10	-0.212416	0.647051	0.901987	0.948462	0.851329	0.869590
std	4.846099e+10	6.507229e+10	0.429655	0.113240	0.066303	0.055794	0.090898	0.073984
=====								
Resumo estatístico do Cluster 1:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.830000e+02	1.830000e+02	183.000000	183.000000	183.000000	183.000000	183.000000	183.000000
mean	1.076001e+09	2.214909e+09	-0.023817	0.411050	0.594718	0.635929	0.497814	0.566940
std	2.122711e+09	4.936087e+09	0.460828	0.129662	0.085400	0.106248	0.126935	0.119959

Melhor cluster
de 2022: 1

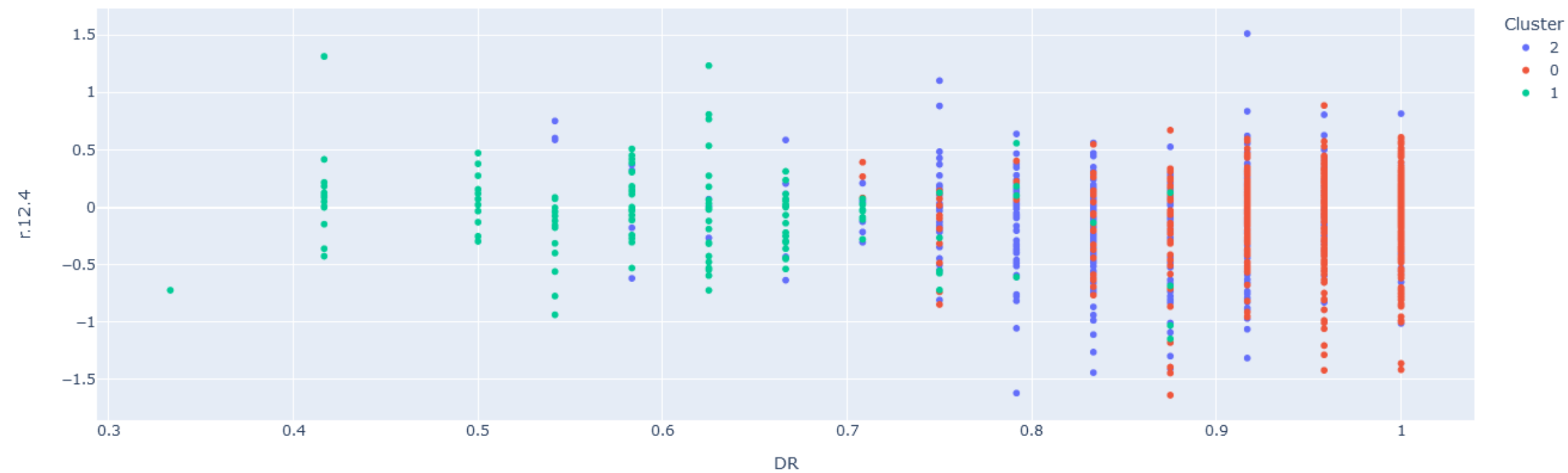
Descrição estatística dos clusters em 2023

Resumo estatístico do Cluster 2:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	5.010000e+02	5.010000e+02	501.000000	501.000000	501.000000	501.000000	501.000000	501.000000
mean	4.942729e+09	7.333973e+09	-0.112413	0.476325	0.778499	0.866434	0.698536	0.774895
std	1.013834e+10	1.498352e+10	0.387969	0.120166	0.073458	0.086237	0.117351	0.087846
=====								
Resumo estatístico do Cluster 0:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	5.550000e+02	5.550000e+02	555.000000	555.000000	555.000000	555.000000	555.000000	555.000000
mean	1.574944e+10	2.261709e+10	-0.085607	0.646396	0.906807	0.949024	0.853033	0.867017
std	4.359905e+10	5.583523e+10	0.379534	0.108815	0.063697	0.059987	0.100900	0.076510
=====								
Resumo estatístico do Cluster 1:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.300000e+02	1.300000e+02	130.000000	130.000000	130.000000	130.000000	130.000000	130.000000
mean	1.029041e+09	1.518532e+09	-0.046057	0.354274	0.568162	0.603846	0.499744	0.566667
std	2.356662e+09	3.237081e+09	0.414930	0.100426	0.077888	0.111544	0.123507	0.122558

Melhor cluster
de 2023: 1

2023

Primeira visualização



Olhando mais de perto...

- Os clusters 2:2019, 0:2020, 1:2021, 1:2022, 1:2023 tem as somas ponderadas bem parecidas...
- Os demais clusters tem as somas ponderadas (excluindo AC) bem diferentes dos clusters com maiores rendimentos...
- De fato, empresas com alto DR tem retornos menores comparados aos demais

Resumo estatístico do Cluster 2:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.030000e+02	1.030000e+02	103.000000	103.000000	103.000000	103.000000	103.000000	103.000000
mean	2.089001e+09	3.829683e+09	0.216986	0.433118	0.591424	0.641990	0.519417	0.566073
std	2.585099e+09	6.093260e+09	0.597280	0.125765	0.065623	0.112598	0.134757	0.118115
Resumo estatístico do Cluster 0:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.990000e+02	1.990000e+02	199.000000	199.000000	199.000000	199.000000	199.000000	199.000000
mean	7.916466e+08	1.481754e+09	0.183863	0.400893	0.578029	0.630653	0.480067	0.527778
std	1.068026e+09	2.552974e+09	0.398945	0.119376	0.067140	0.112575	0.118204	0.113230
Resumo estatístico do Cluster 1:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	2.010000e+02	2.010000e+02	201.000000	201.000000	201.000000	201.000000	201.000000	201.000000
mean	9.011930e+08	1.885289e+09	0.230655	0.413626	0.589414	0.631841	0.492040	0.543809
std	1.150390e+09	2.743490e+09	0.448055	0.122500	0.079455	0.109671	0.122441	0.116220
Resumo estatístico do Cluster 1:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.830000e+02	1.830000e+02	183.000000	183.000000	183.000000	183.000000	183.000000	183.000000
mean	1.076001e+09	2.214909e+09	-0.023817	0.411050	0.594718	0.635929	0.497814	0.566940
std	2.122711e+09	4.936087e+09	0.460828	0.129662	0.085400	0.106248	0.126935	0.119959
Resumo estatístico do Cluster 1:								
	BE	ME	r.12.4	AC	CA	DR	OFC	ECI
count	1.300000e+02	1.300000e+02	130.000000	130.000000	130.000000	130.000000	130.000000	130.000000
mean	1.029041e+09	1.518532e+09	-0.046057	0.354274	0.568162	0.603846	0.499744	0.566667
std	2.356662e+09	3.237081e+09	0.414930	0.100426	0.077888	0.111544	0.123507	0.122558

Temos uma espécie de padrão!