



Università degli Studi di Milano Bicocca

**Scuola di Scienze**

**Dipartimento di Informatica, Sistemistica e Comunicazione**

**Corso di laurea in Informatica**

# **Business Intelligence per i Servizi Finanziari**

**Alessandro Zanotti - 885892**

## Sommario

Sommario dei dati utilizzati: .....	4
Funzioni utilizzate per caricare i dati finanziari .....	4
Presentazione dei dati con un grafico e le prime righe del DataFrame .....	5
Statistiche descrittive .....	6
Rendimenti cumulati .....	6
Rendimenti composti .....	6
Rendimenti semplici .....	7
Rendimenti logaritmici .....	10
Informazioni sui rendimenti .....	13
Grafici diagnostici a 3 sezioni (istogramma, boxplot e qq-plot) .....	14
Broadcom Inc. ....	14
Advanced Micro Devices Inc. ....	15
Eli Lilly .....	16
Pfizer .....	17
Berkshire Hathaway Class B .....	18
Blackrock .....	19
Statistiche descrittive univariate .....	20
Matrici di covarianza e di correlazione dei rendimenti .....	21
Matrice di covarianza .....	21
Matrice di correlazione .....	21
Andamento nel tempo delle correlazioni tra gli asset e le correlazioni medie .....	22
Correlazione media dei titoli del settore tecnologico .....	25
Correlazione media dei titoli del settore sanitario .....	26
Correlazione media dei titoli del settore finanziario .....	27
Analisi di previsione .....	28
Analisi AVGO .....	28
Analisi AMD .....	29
Analisi LLY .....	30
Analisi PFE .....	31
Analisi BRK.B .....	32
Analisi BLK .....	33
Strategie di trading e backtesting .....	34
Confronto con la strategia Buy and Hold .....	37
Capital Asset Pricing Model .....	40
Costruzione di un portafoglio .....	44

Conclusioni ..... 46

## Sommario dei dati utilizzati:

Breve descrizione di ciascun titolo e motivazione della scelta:

- Settore tecnologico:
  - Broadcom, Inc (AVGO): leader tecnologico nel settore microchip e nel settore software. Recentemente ha acquisito VMware Inc (VMW) per 69 miliardi.
  - Advanced Micro Devices, Inc (AMD): AMD sviluppa processori per PC e tecnologie collegate per consumatori e imprese. Scelto perché nel 2024 ha un YTD return alto grazie all'aumento della richiesta per soluzioni computazionali ad alte performance.
- Settore sanitario:
  - Eli Lilly and Company (LLY): azienda farmaceutica globale che sviluppa, produce e commercia nei settori endocrinologi, oncologici, immunologici. Scelta per la serie di farmaci disponibili, specialmente per diabete e pazienti oncologici, e recenti trial clinici di successo.
  - Pfizer Inc. (PFE): azienda biofarmaceutica globale che sviluppa farmaci innovativi e vaccini per diverse aree terapeutiche come immunologia, cardiologia, endocrinologia e neurologia. Scelta per il suo contributo durante la pandemia di Covid-19 del 2020 grazie al vaccino Pfizer-BioNTech.
- Settore finanziario:
  - Berkshire Hathaway Inc. Class B (BRK.B): conglomerato multinazionale amministrato da Warren Buffet. L'azienda possiede un business molto diversificato che include assicurazioni, servizi, ferrovie, manifatturiero e possiede posizioni di equity importanti in Apple, Bank of America e Coca-Cola. Scelto per lo storico record di crescita e il portafoglio altamente diversificato le cui componenti sono scelte oculatamente.
  - BlackRock, Inc. (BLK): è l'asset manager più grande al mondo e fornisce servizi di management degli investimenti, dei rischi e servizi di advisory. Scelta per la posizione dominante nel mercato e per il suo approccio innovativo alle soluzioni di investimento.

## Funzioni utilizzate per caricare i dati finanziari

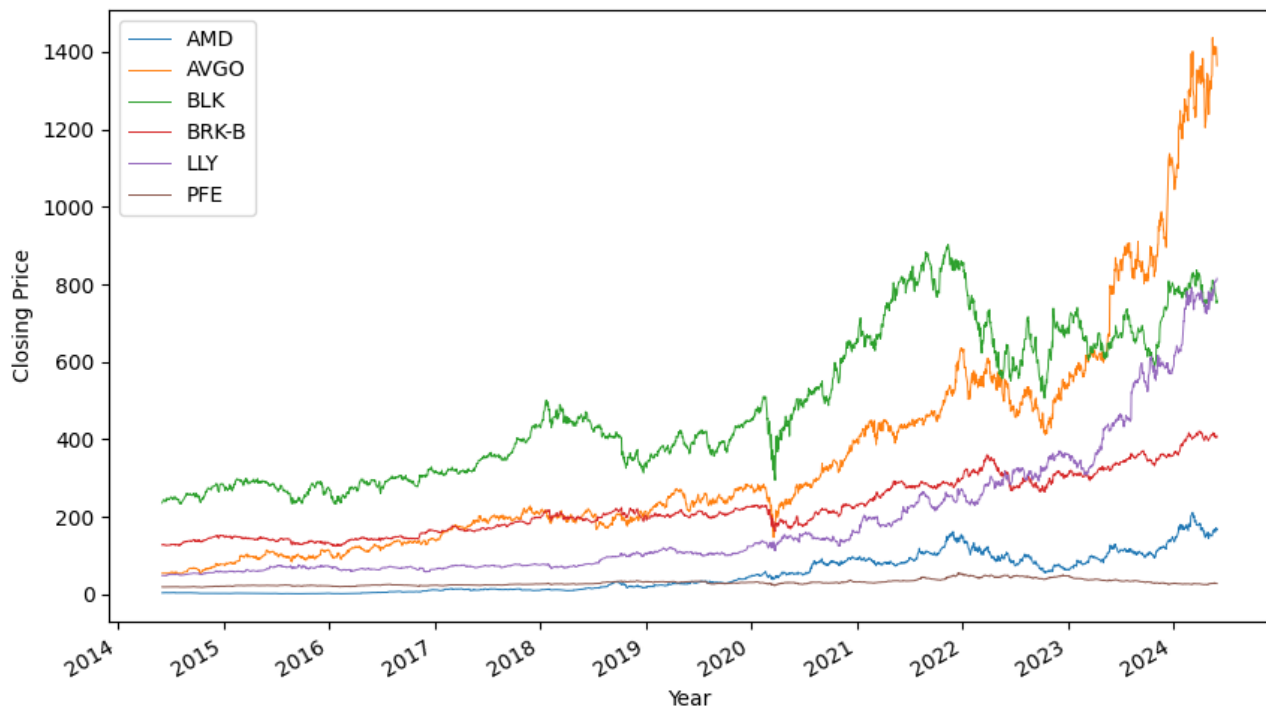
Per scaricare i dati ho utilizzato la funzione **download** della libreria **yfinance** (Yahoo! Finance).

L'interfaccia della funzione è `yf.download(tickers, ...)` -> `pd.DataFrame` e possiede diversi parametri:

<b>Tickers</b>	<b>Interval</b>	<b>Period</b>	<b>Start</b>
<b>End</b>	Prepost	Actions	Auto_adjust
<b>Back_adjust</b>	Repair	Keepna	Rounding
<b>Group_by</b>	Threads	Proxy	Session
<b>Timeout</b>	Progress		

I dati Fama-French vengono scaricati tramite la funzione **request.urlretrieve** di **urllib** dalla Data Library di Ken French (Dartmouth University).

## Presentazione dei dati con un grafico e le prime righe del DataFrame



DATA\TICKER	AVGO	AMD	LLY	PFE	BRK.B	BLK
2014-06-02	54.481522	3.97	48.665128	19.234695	127.879997	235.882156
2014-06-03	54.337021	3.94	48.540630	19.157009	126.930000	237.870789
2014-06-04	54.458710	4.04	48.434299	19.189383	127.230003	239.520187
2014-06-05	54.656429	4.08	48.728752	19.267069	128.199997	240.381165
2014-06-06	54.473927	4.06	48.769634	19.046946	128.649994	243.506805

# Statistiche descrittive

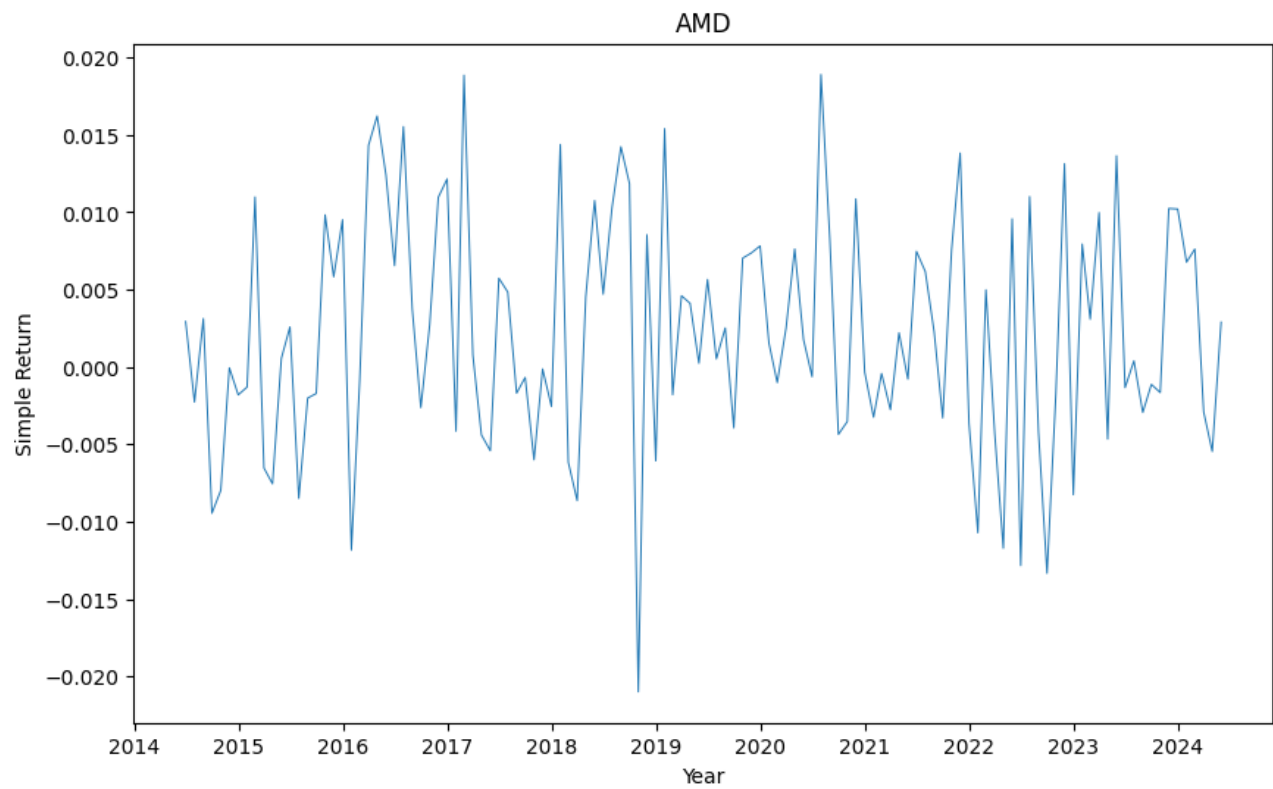
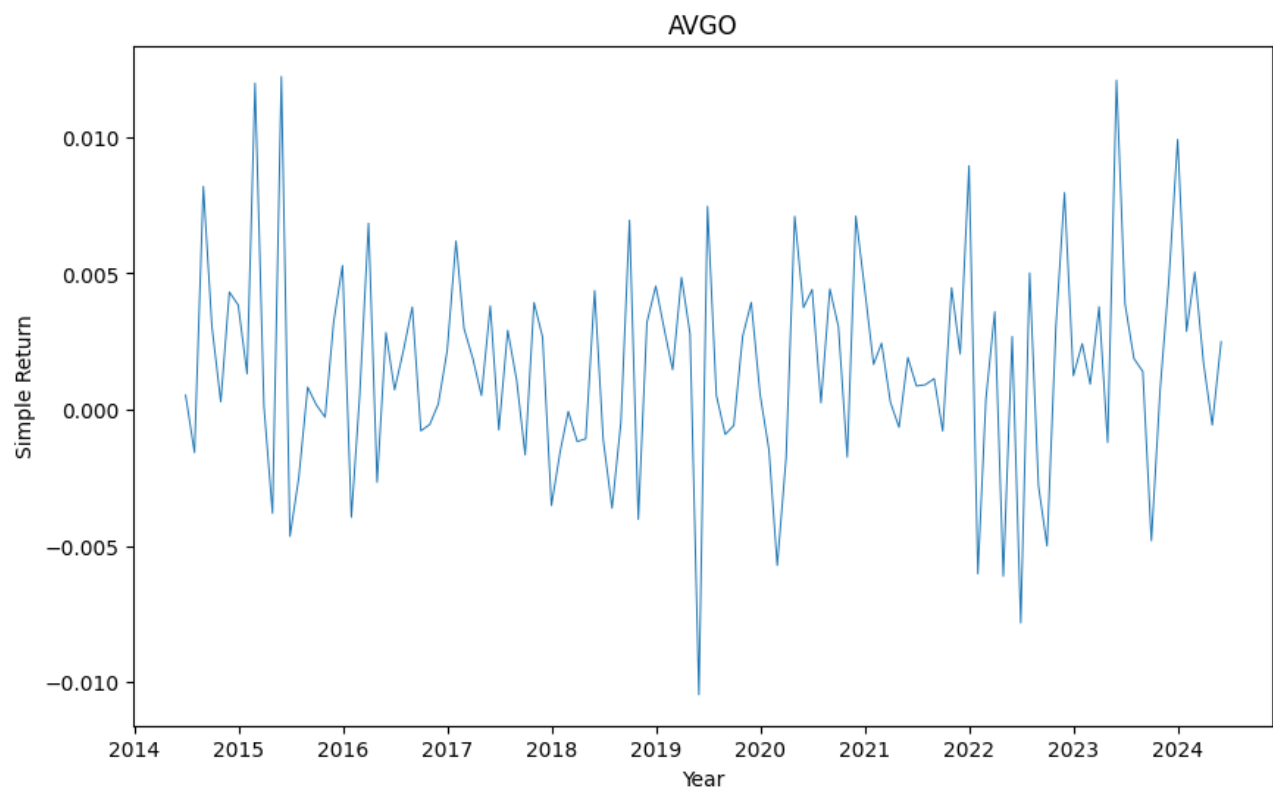
## Rendimenti cumulati

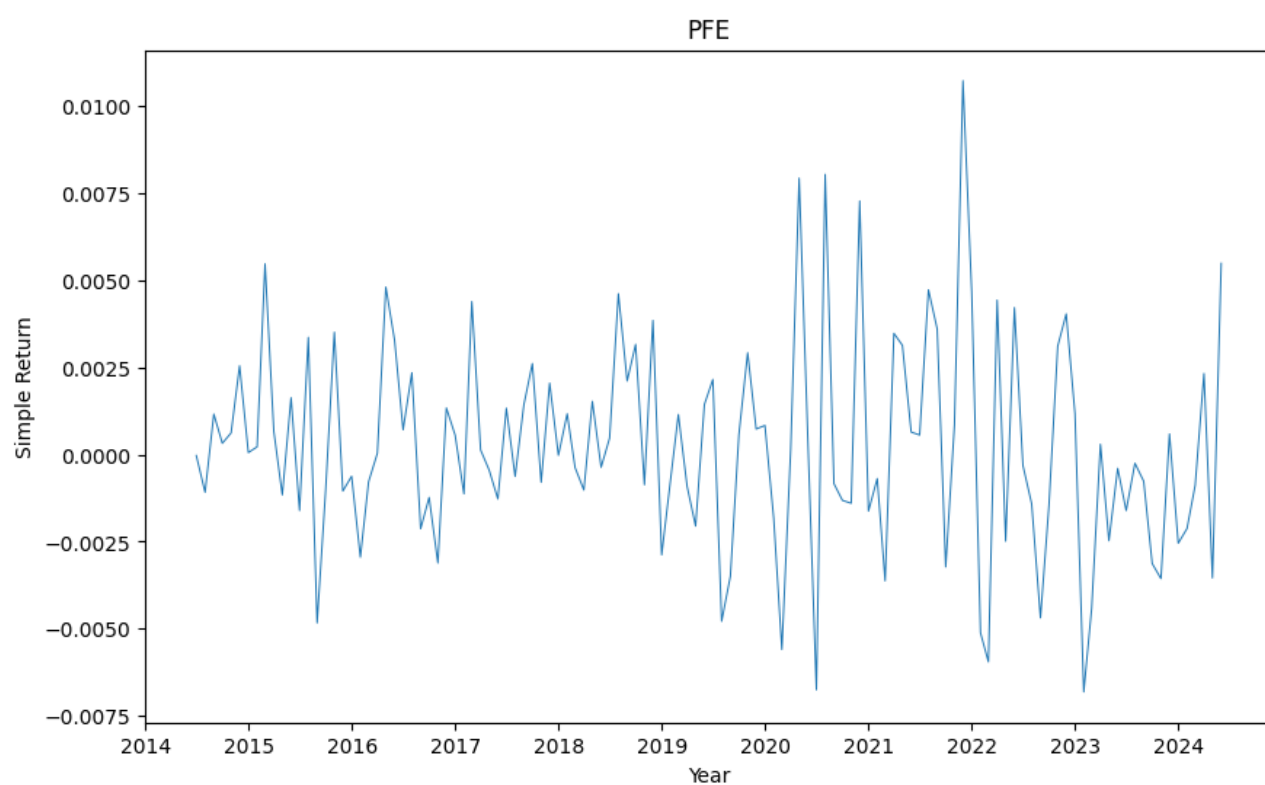
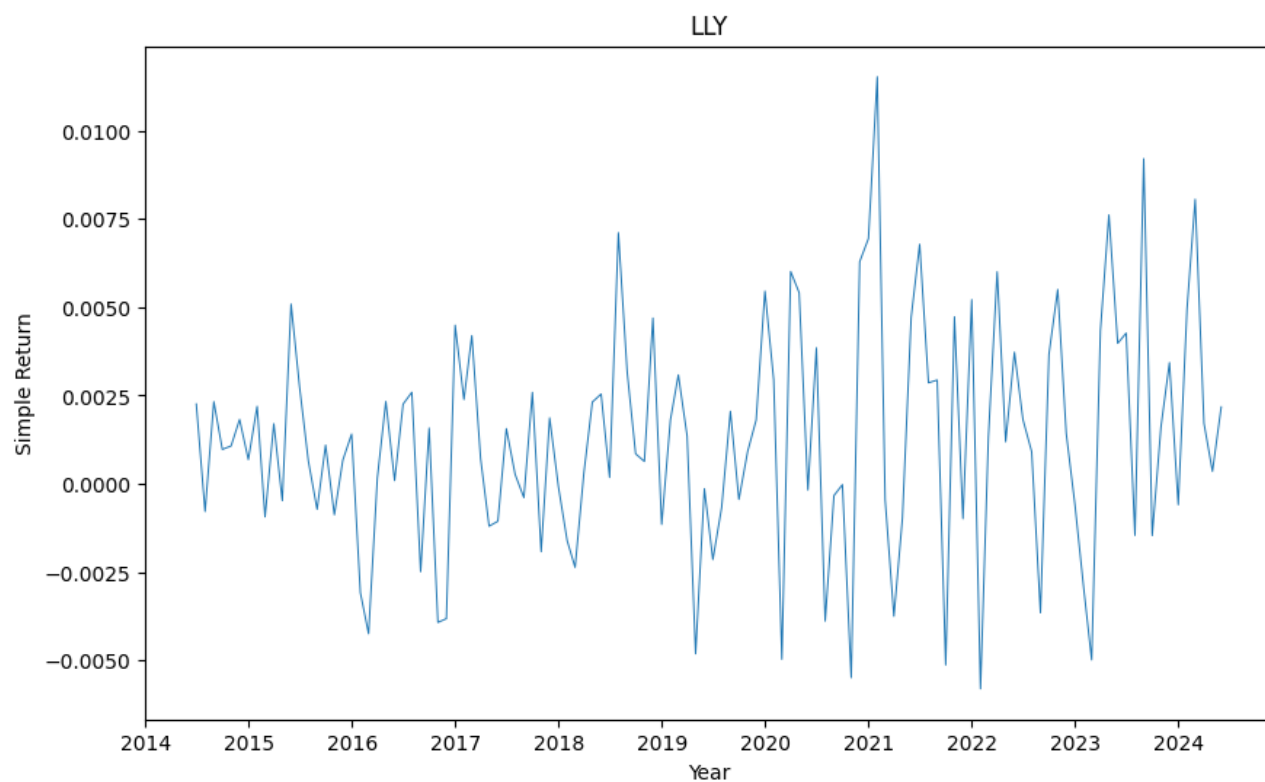
DATA\TICKER	AVGO	AMD	LLY	PFE	BRK-B	BLK
2014-06-03	0.997348	0.992443	0.997647	0.995961	0.992571	1.008431
2014-06-04	0.999581	1.017632	0.995461	0.997644	0.994917	1.015423
2014-06-05	1.003210	1.027708	1.001513	1.001683	1.002502	1.019073
2014-06-06	0.999861	1.022670	1.002353	0.990239	1.006021	1.032324
...	...	...	...	...	...	...
2024-05-23	25.573626	40.410577	16.615926	1.491576	3.173913	3.304237
2024-05-24	25.840687	41.904282	16.594962	1.501453	3.185877	3.293076
2024-05-28	25.925303	43.226700	16.603799	1.471300	3.158430	3.248896
2024-05-29	25.525536	41.596977	16.679229	1.446345	3.159916	3.186438
2024-05-30	25.037479	42.002519	16.751780	1.466101	3.195261	3.208675

## Rendimenti composti

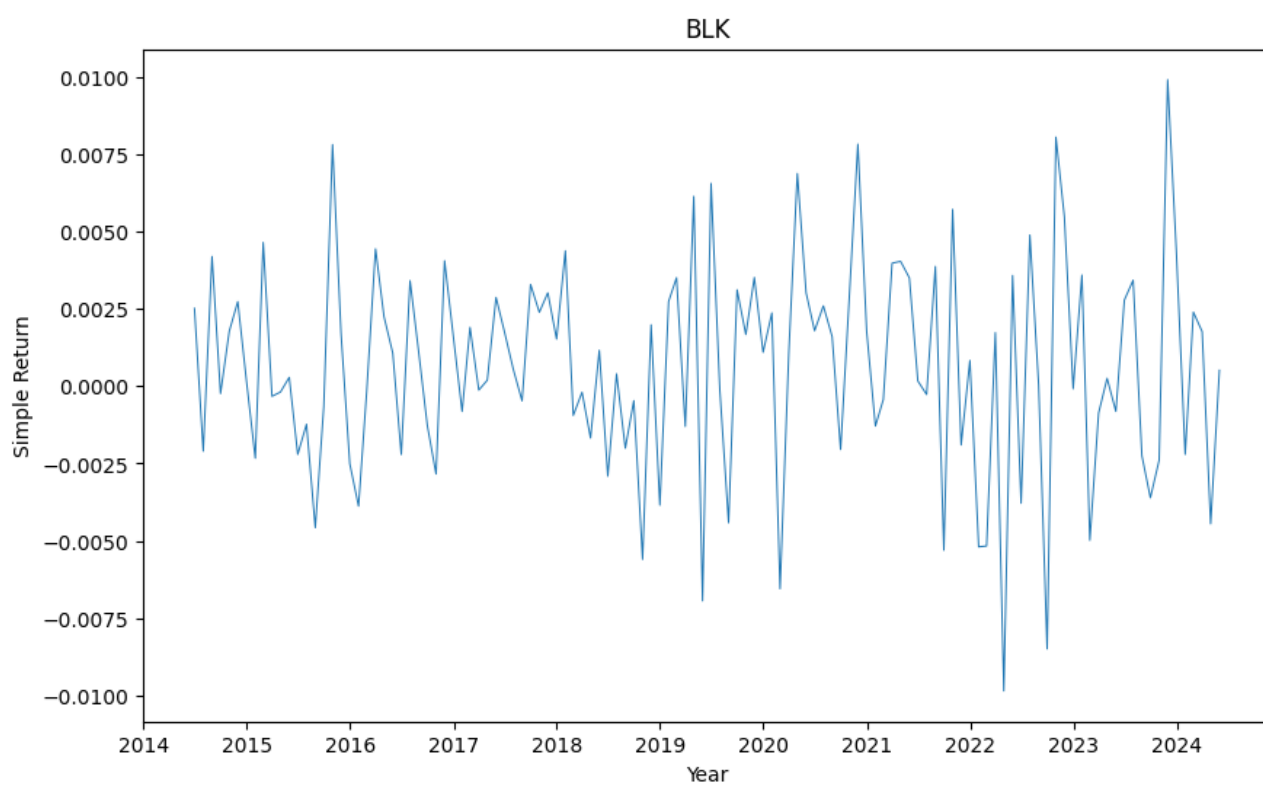
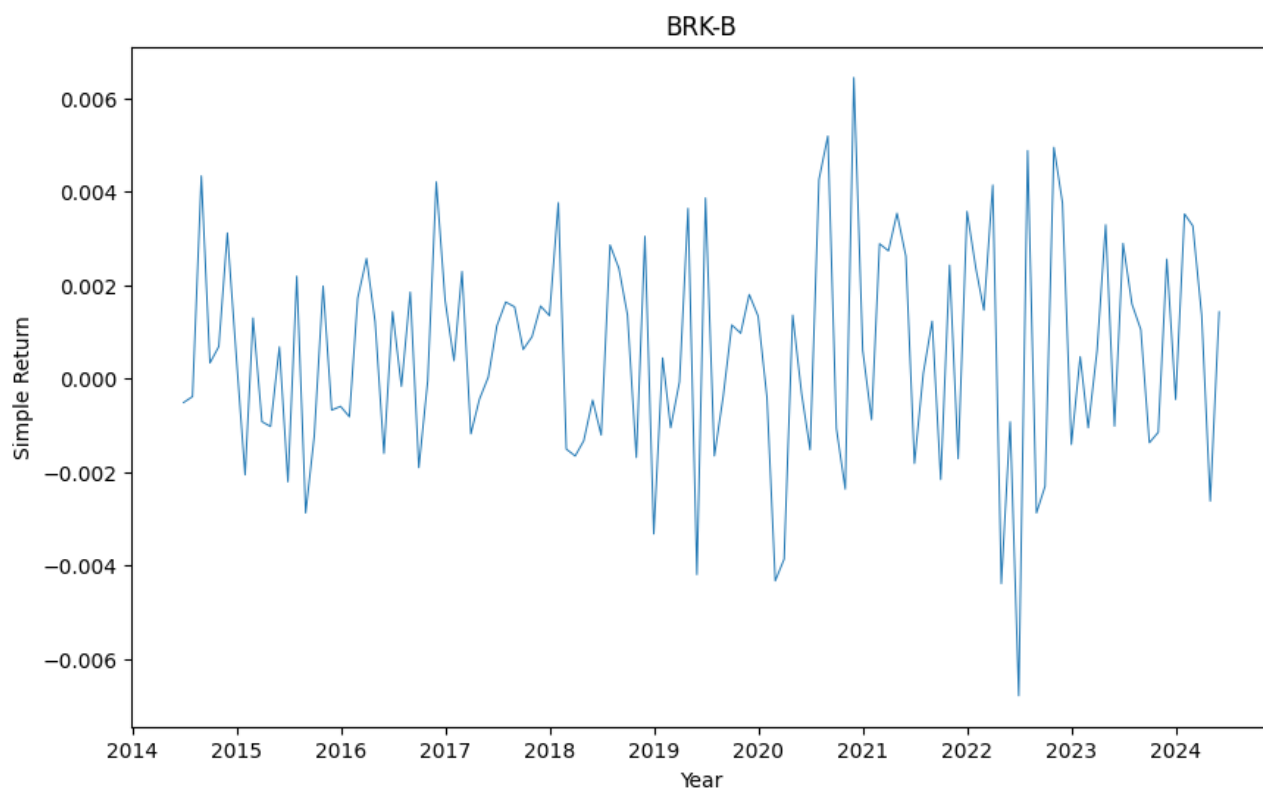
TICKER	AVGO	AMD	LLY	PFE	BRK.B	BLK
VALORE (%)	38.0%	45.3%	32.5%	3.9%	12.3%	12.4%

Rendimenti semplici

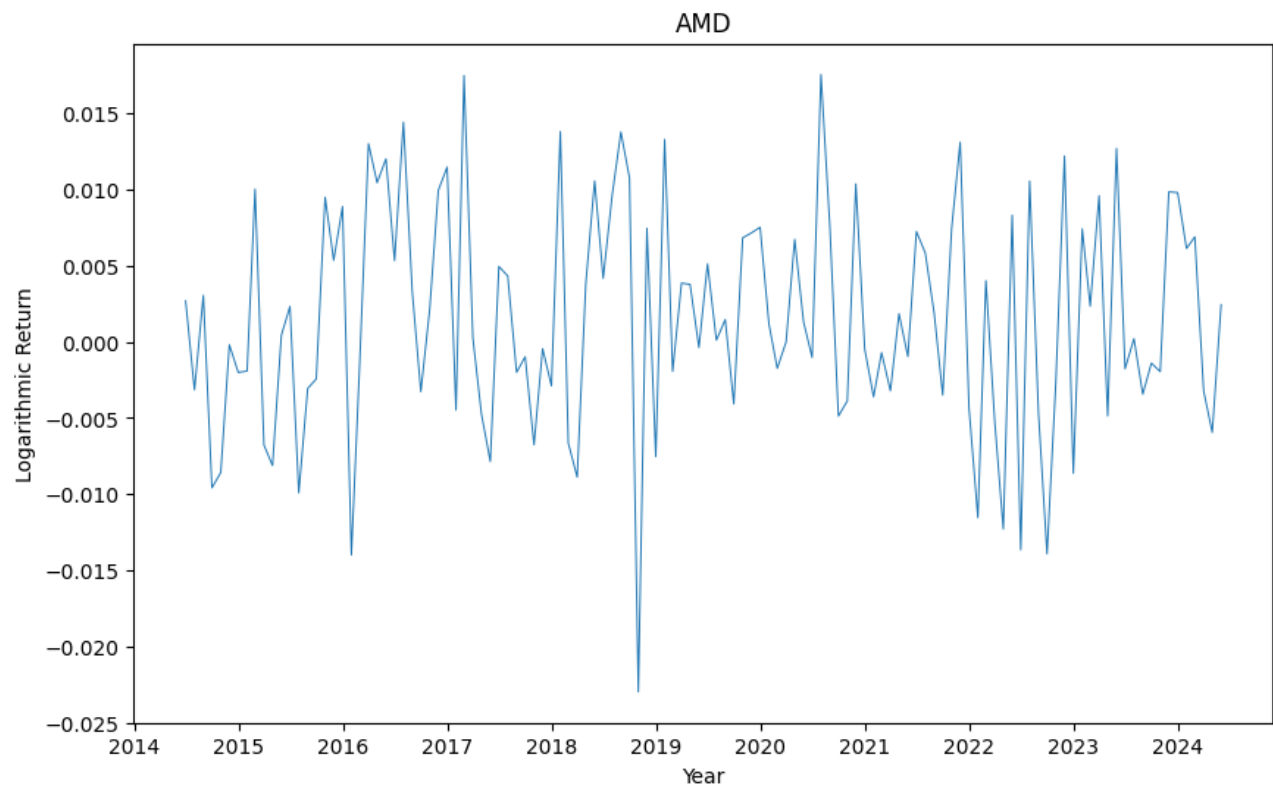
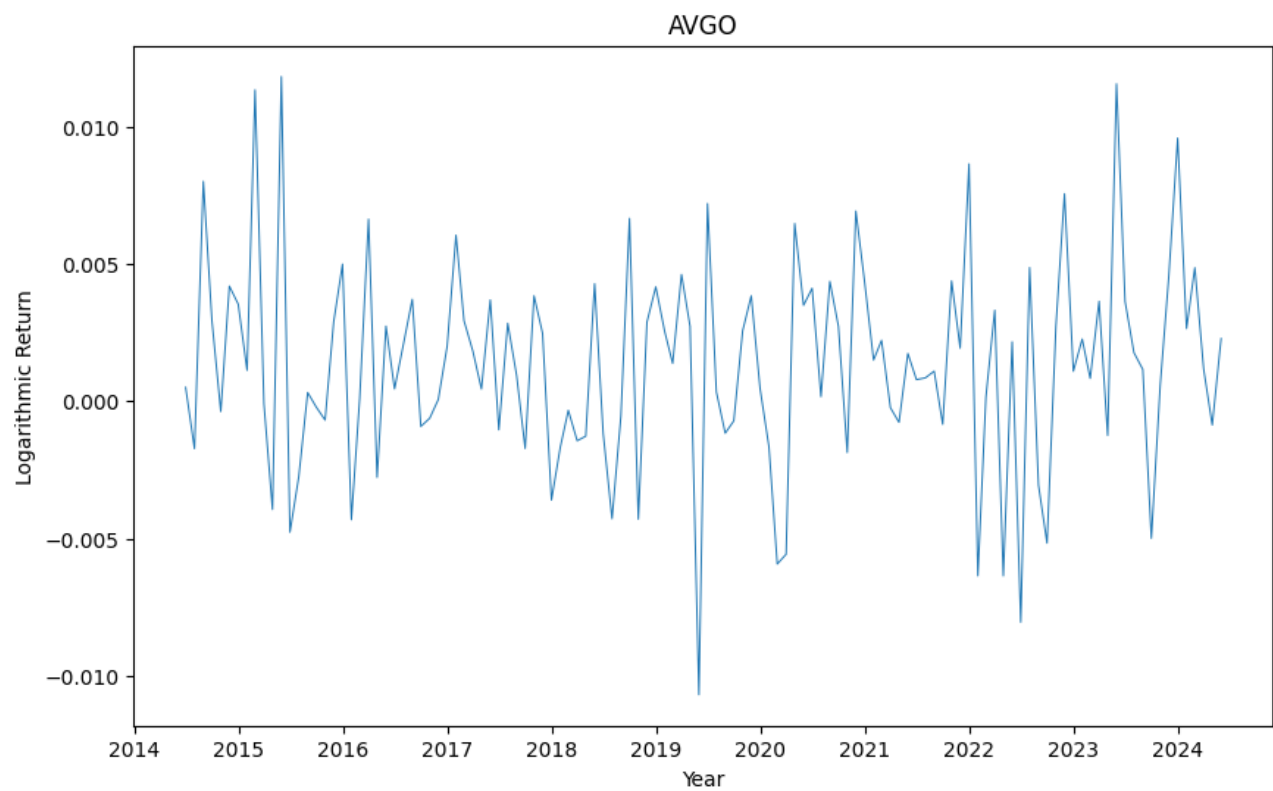


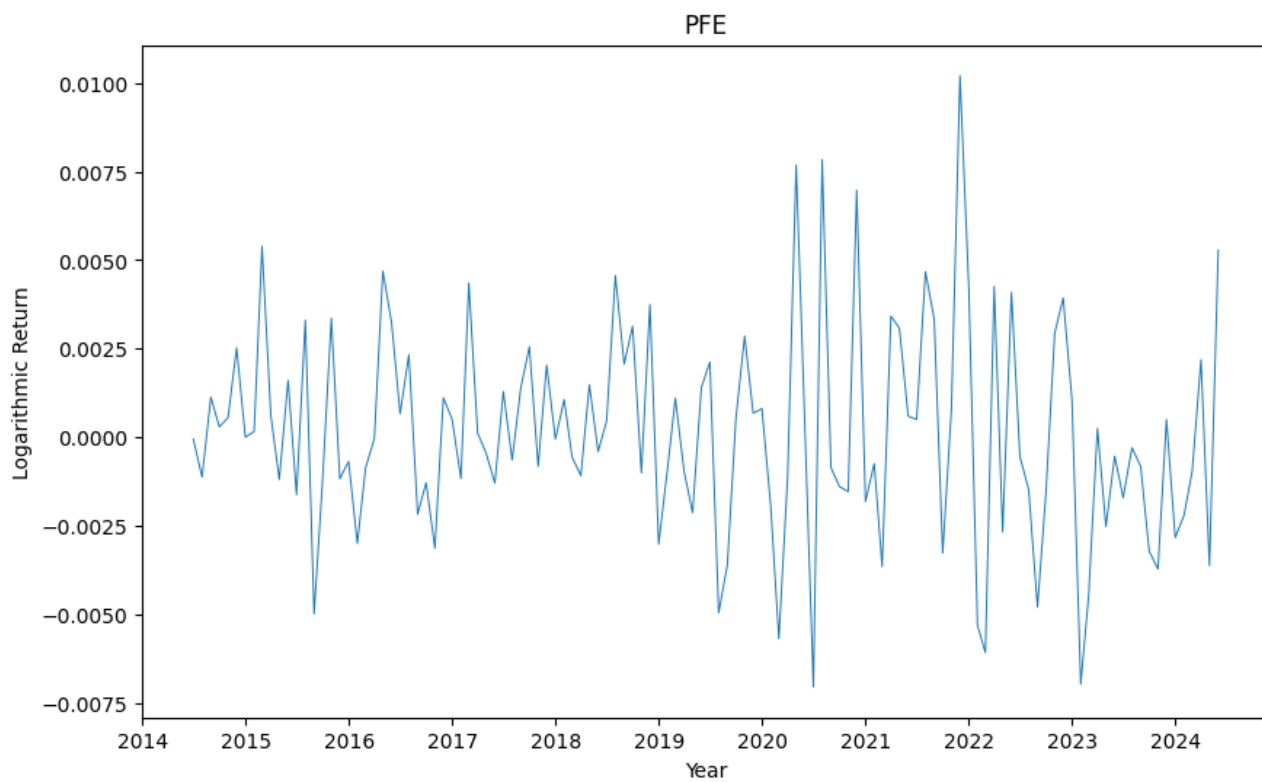
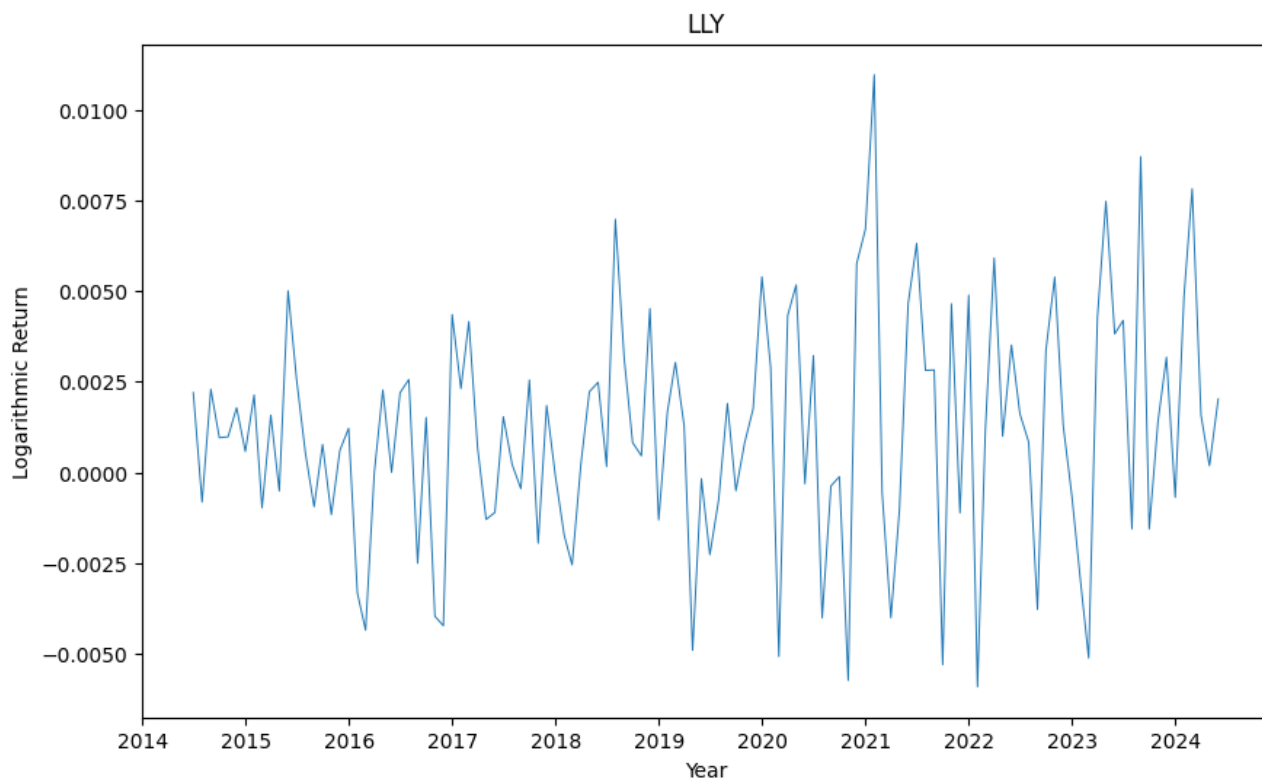


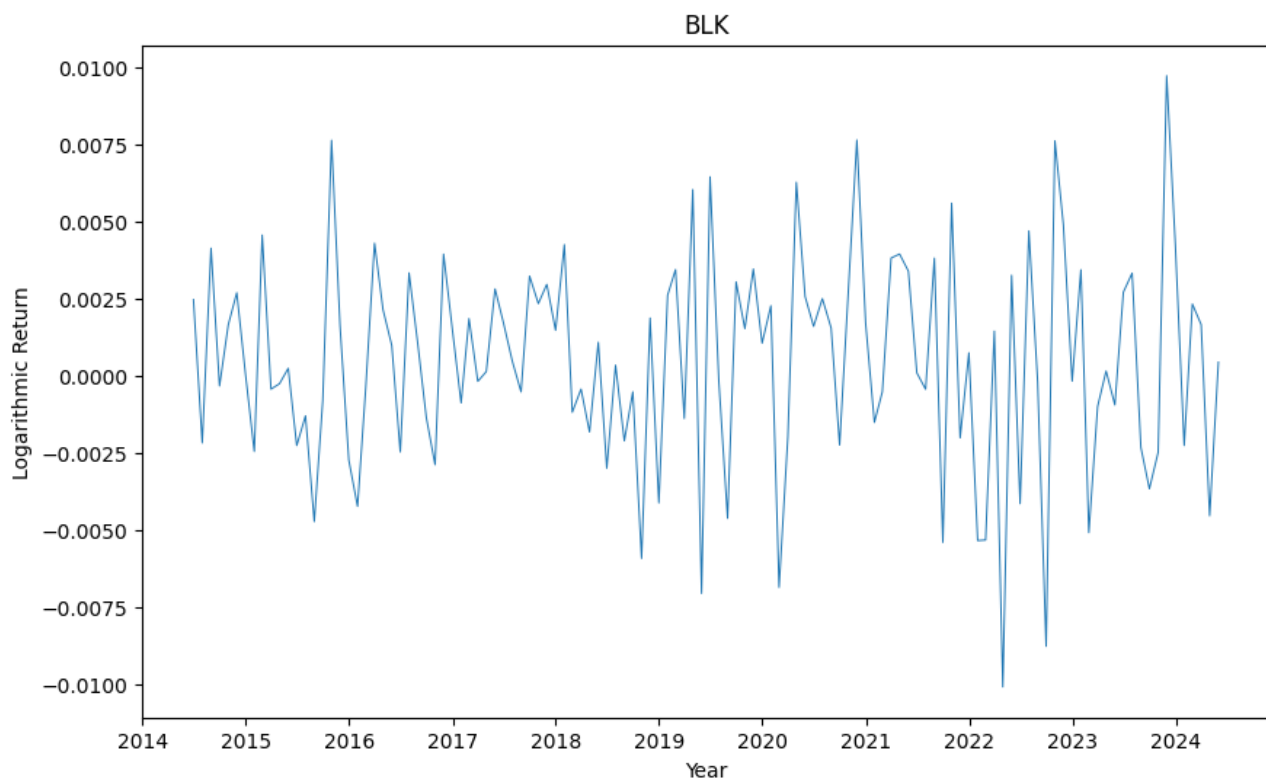
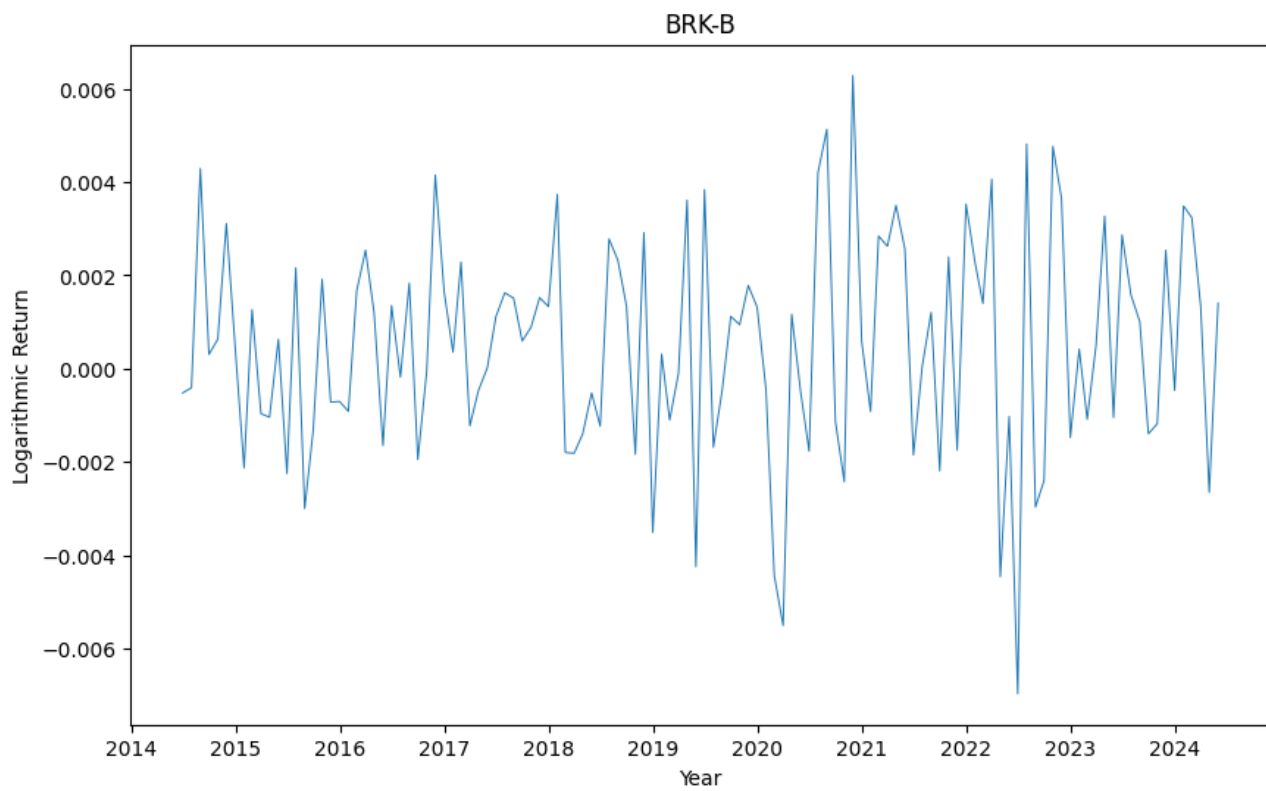




Rendimenti logaritmici







Analizzando i grafici possiamo notare alcuni andamenti e correlazioni tra i titoli scelti. Ovviamente la prima informazione che si nota è che l'andamento del mercato finanziario influenza allo stesso modo tutti i titoli, ad esempio nel 2020 a causa del Covid Crash.

Inoltre si nota anche una correlazione positiva tra le aziende che operano nello stesso settore:

- AVGO e AMD producono entrambe chip
- LLY e PFE producono farmaci per settori sanitari che spesso sono trasversali
- BRK.B e BLK spesso investono nelle stesse aziende

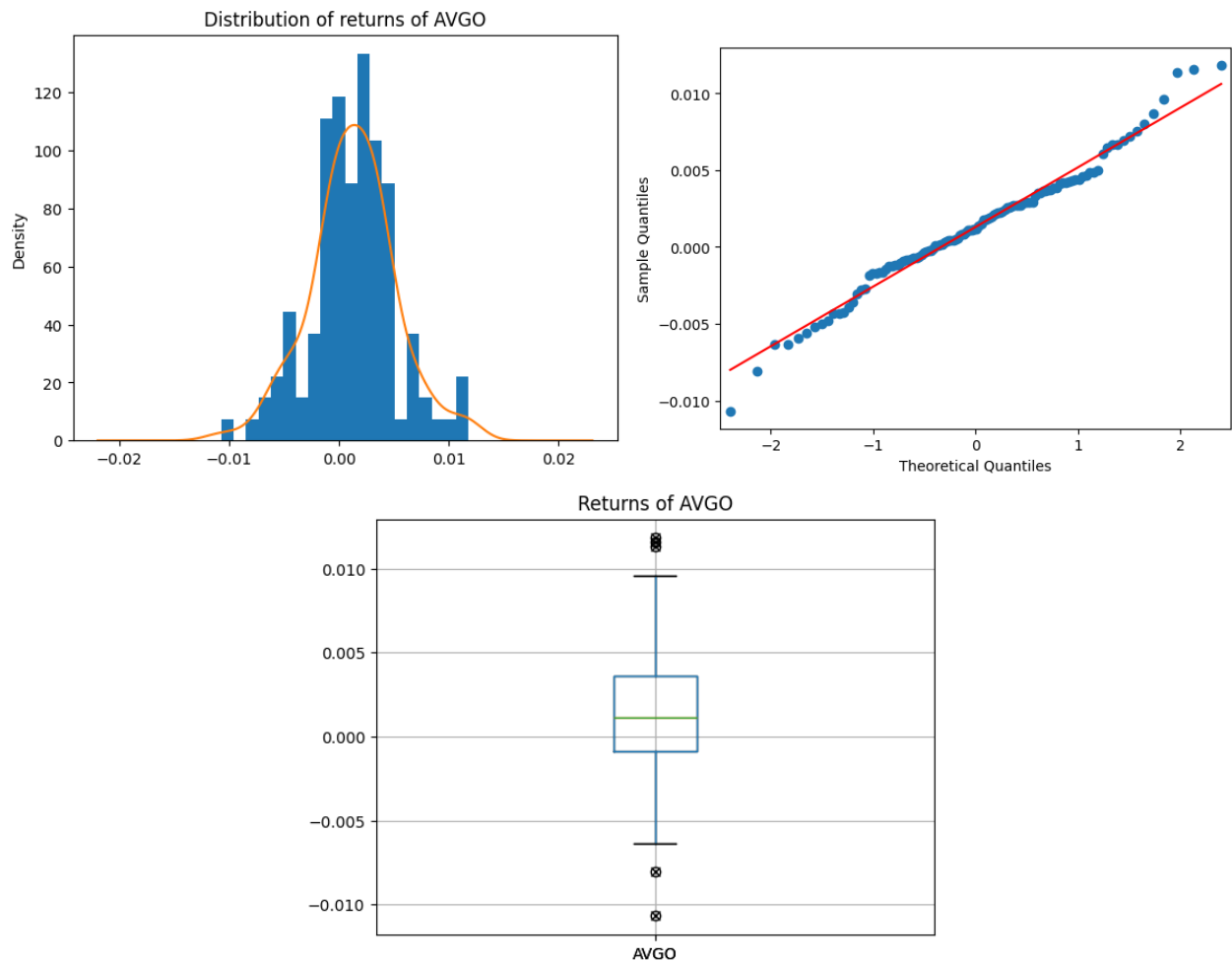
#### Informazioni sui rendimenti

TICKER	AMD	AVGO	BLK	BRK-B	LLY	PFE
COUNT	2516.000000	2516.000000	2516.000000	2516.000000	2516.000000	2516.000000
MEAN	0.001486	0.001280	0.000463	0.000462	0.001120	0.000152
STD	0.036220	0.022179	0.017147	0.012110	0.016900	0.014355
MIN	-0.277456	-0.222055	-0.146815	-0.100838	-0.111094	-0.080502
25%	-0.016697	-0.009964	-0.007751	-0.005321	-0.006760	-0.006726
50%	0.000586	0.001316	0.000778	0.000481	0.000961	0.000000
75%	0.019693	0.012688	0.008739	0.006540	0.008839	0.007011
MAX	0.420617	0.146985	0.126843	0.109840	0.145656	0.103055

Dalla deviazione standard dei rendimenti logaritmici possiamo notare che la volatilità per i titoli dei settori Sanitari (LLY, PFE) e Finanziari (BRK-B, BLK) è relativamente bassa, mentre per i titoli del settore tecnologico (AVGO, AMD) è più alta. In particolare, la volatilità di AMD è significativamente maggiore rispetto agli altri titoli, quasi il doppio rispetto alla maggior parte dei titoli sanitari e finanziari.

## Grafici diagnostici a 3 sezioni (istogramma, boxplot e qq-plot)

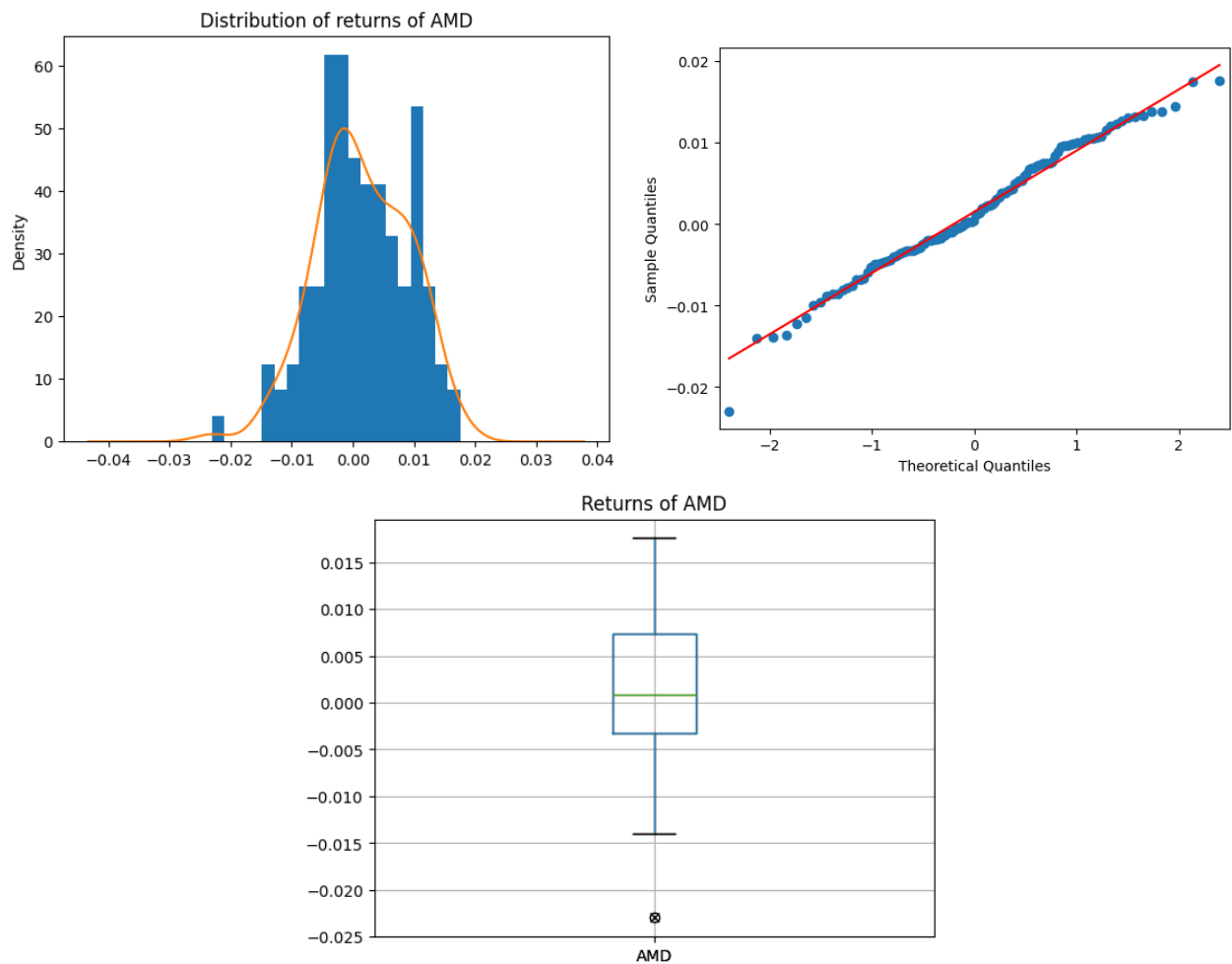
Broadcom Inc.



Dai grafici diagnostici vediamo che i dati sono distribuiti normalmente, la maggior parte dei punti si trova vicino alla linea rossa (che ha inclinazione positiva) indicando che i rendimenti seguono una distribuzione normale. Ci sono alcune deviazioni agli estremi ma non sono eccessive.

La variabilità dei rendimenti è relativamente bassa con la maggior parte dei rendimenti che rientrano in un intervallo ristretto. Ci sono alcuni outlier che è tipico per i dati finanziari e possono indicare periodi di maggiore volatilità o eventi di mercato significativi.

## Advanced Micro Devices Inc.

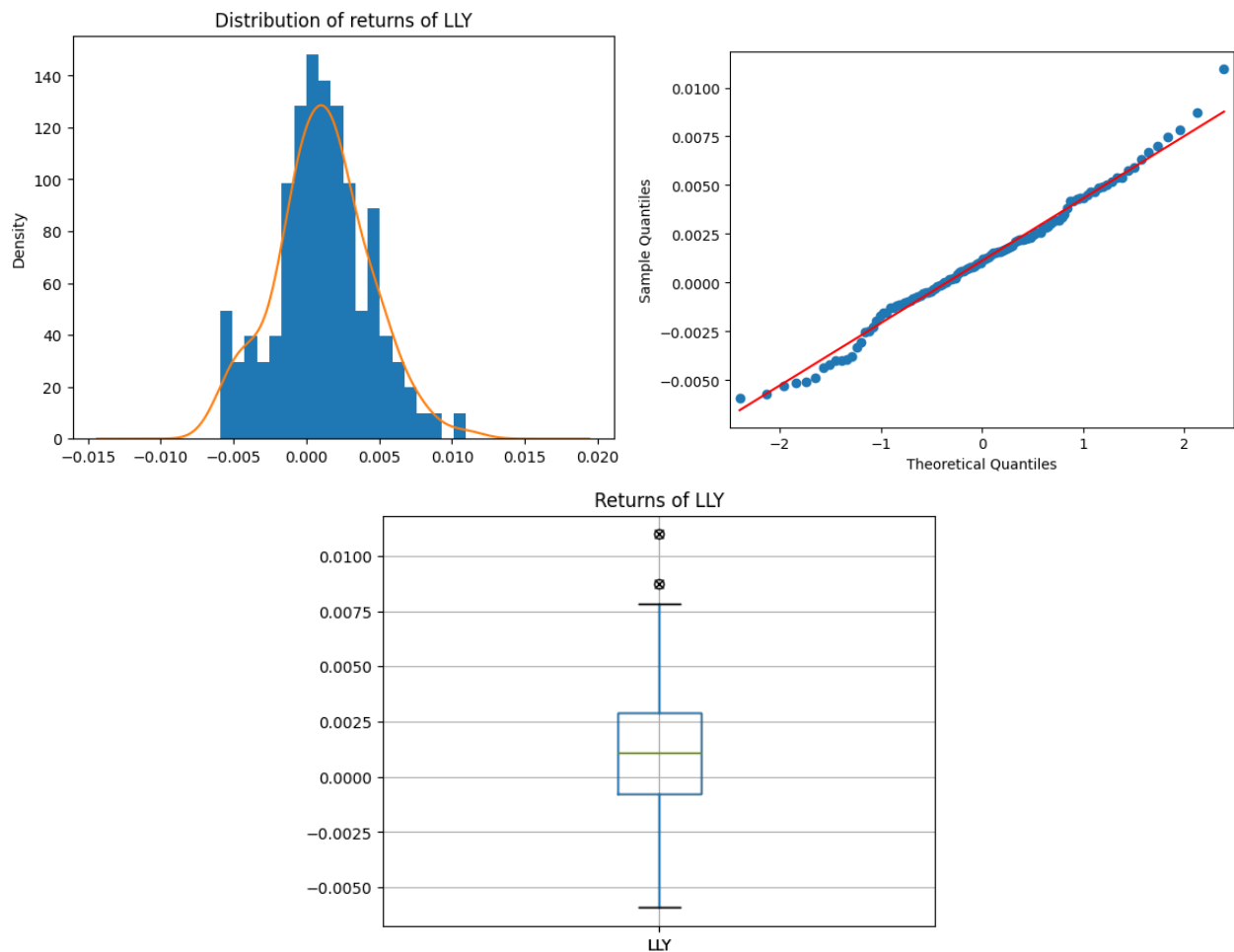


La distribuzione appare abbastanza simmetrica e centrata intorno a zero. La curva di densità sovrapposta conferma questa simmetria e suggerisce una distribuzione normale dei rendimenti con una leggera inclinazione negativa.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale. Ci sono alcune deviazioni agli estremi ma non sono eccessive.

Il boxplot indica che la maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier sotto la scatola. La presenza di outlier è tipico per i dati finanziari e può indicare periodi di maggiore volatilità o eventi di mercato significativi.

## Eli Lilly



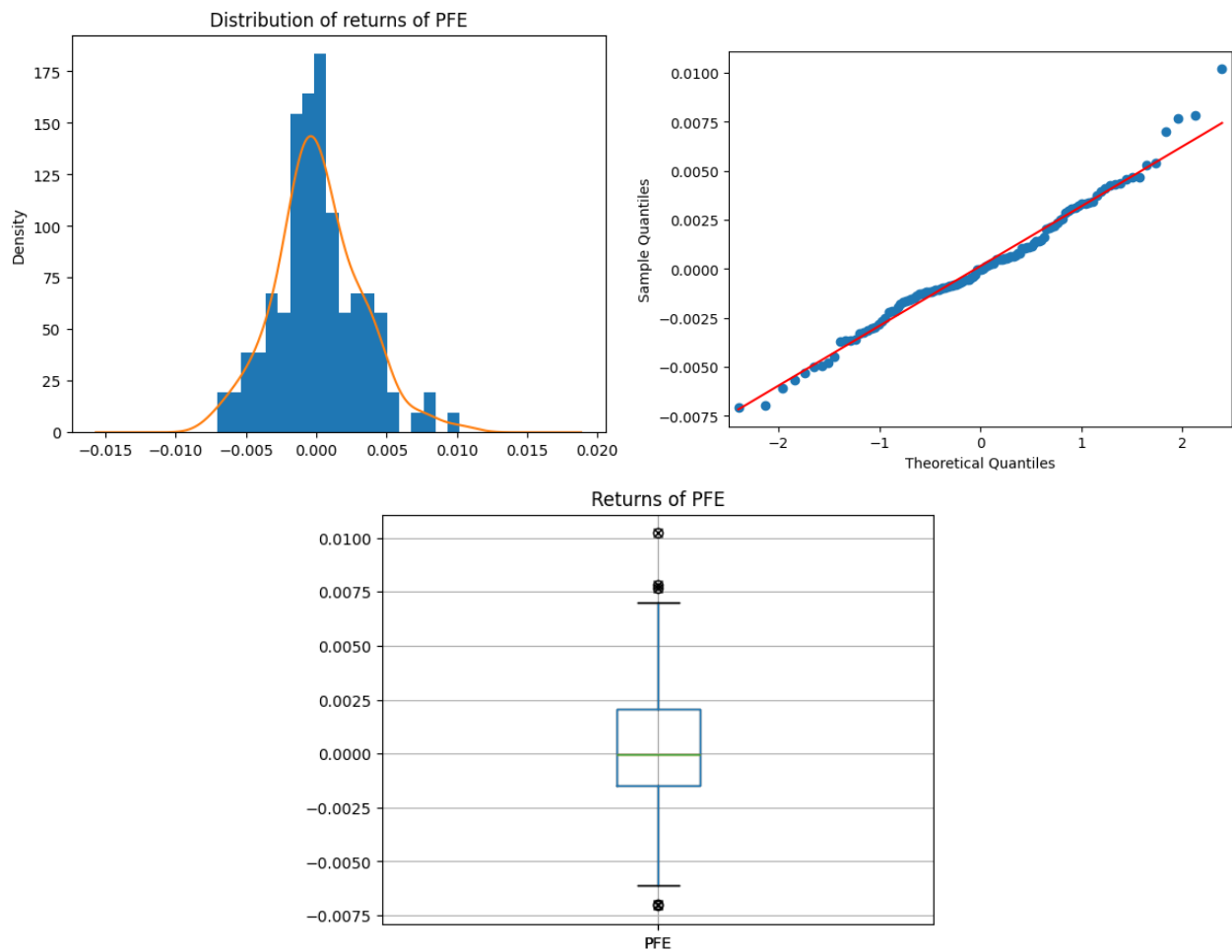
L'istogramma mostra una distribuzione abbastanza simmetrica e centrata attorno a zero. La curva di densità conferma questa simmetria e suggerisce una distribuzione dei rendimenti con una leggera inclinazione positiva.

Il grafico QQ plot mostra che la maggior parte dei punti si trova vicino alla linea rossa indicando che i rendimenti seguono una distribuzione normale. Ci sono alcune deviazioni agli estremi ma non sono eccessive e confermano la tendenza generale alla normalità.

Infine il boxplot mostra che la maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier, sopra la scatola. La presenza di outlier è comune e può essere dovuta a periodi di maggiore volatilità o eventi di mercato significativi.



## Pfizer

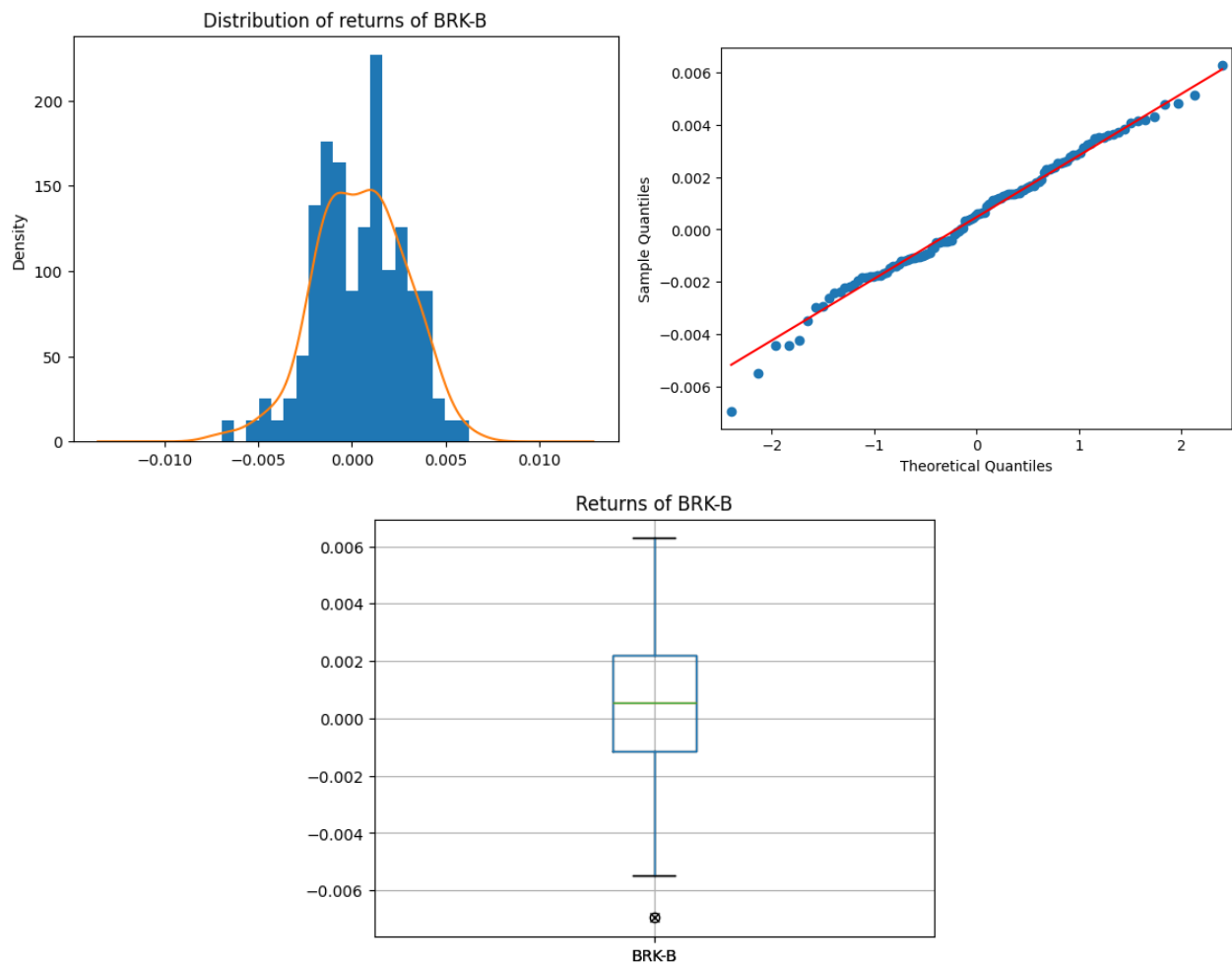


L'istogramma mostra una distribuzione abbastanza simmetrica e centrata attorno a zero. La curva di densità conferma la simmetria e suggerisce una distribuzione normale dei rendimenti con una leggera inclinazione positiva.

Il grafico QQ plot indica che la maggior parte dei punti si trova vicino alla linea rossa indicando che i rendimenti seguono una distribuzione normale. Ci sono alcune deviazioni agli estremi ma non sono eccessive e confermo la tendenza generale alla normalità.

Infine il boxplot mostra che la maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier presenti sopra e sotto la scatola.

## Berkshire Hathaway Class B

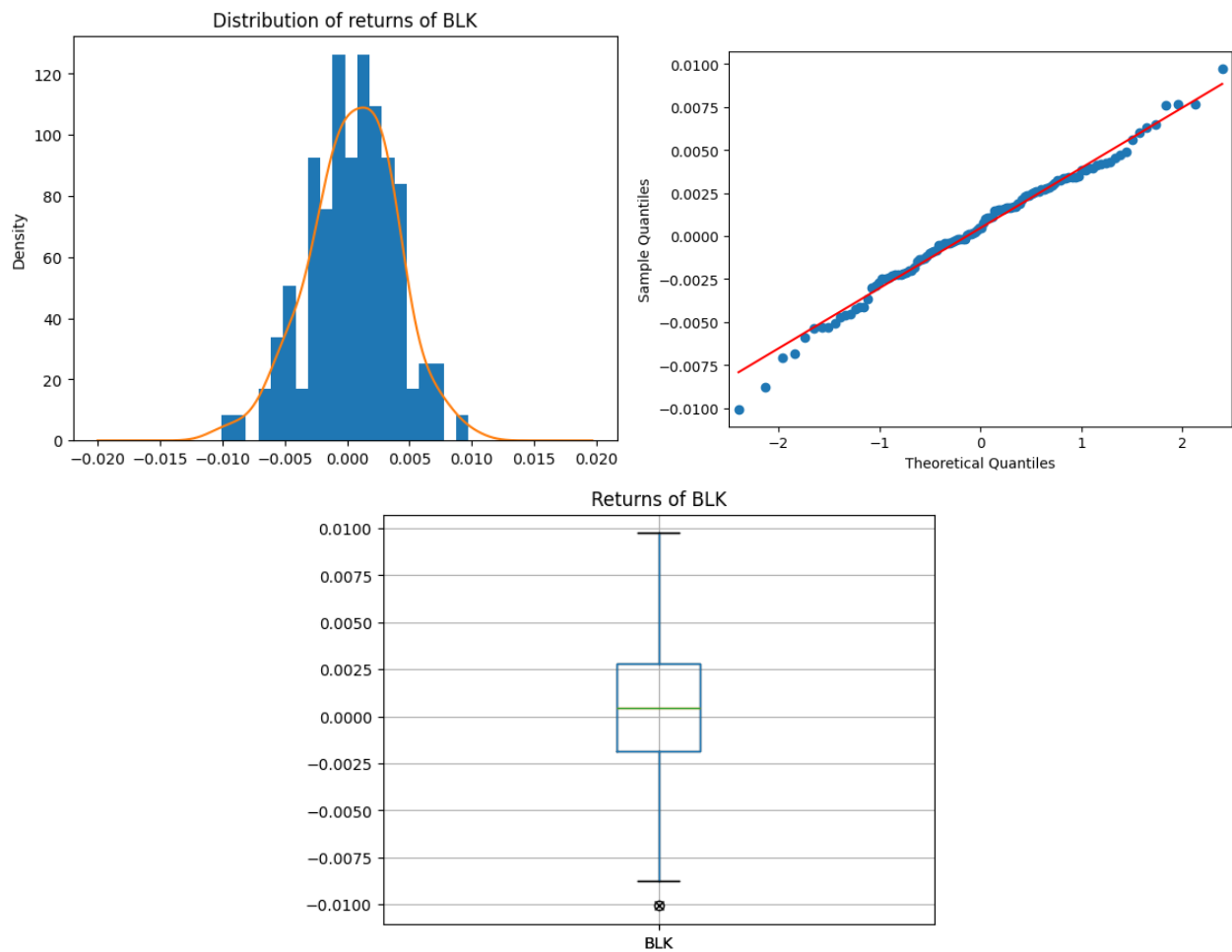


L'istogramma mostra una distribuzione che appare abbastanza simmetrica e centrata intorno a zero. La curva di densità sovrapposta conferma la simmetria e suggerisce una distribuzione normale dei rendimenti con una leggera inclinazione positiva.

Il grafico QQ plot indica che la maggior parte dei punti si trova vicino alla linea rossa quindi i rendimenti seguono una distribuzione normale. Ci sono deviazioni agli estremi ma non sono eccessivi, confermando la tendenza generale alla normalità.

Infine il boxplot mostra che la maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier sotto la scatola.

## Blackrock



L'istogramma mostra una distribuzione abbastanza simmetrica e centrata intorno allo zero. La curva di densità sovrapposta conferma la simmetria e suggerisce una distribuzione normale con una leggera inclinazione negativa.

Il grafico QQ plot indica che la maggior parte dei punti si trova vicino alla linea rossa e quindi i rendimenti seguono una distribuzione normale. Ci sono alcune deviazioni agli estremi ma non sono eccessive.

Il boxplot mostra che la maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con outlier sotto la scatola.

## Statistiche descrittive univariate

Utilizzando le statistiche descrittive univariate, possiamo ottenere preziose informazioni sui vari aspetti chiave degli asset finanziari scelti. Le statistiche di maggiore interesse sono:

- **Media:** rappresenta il rendimento medio dell'asset nel periodo di tempo considerato e ci consente di individuare la tendenza degli investimenti nel tempo.
- **Varianza e Deviazione Standard:** indicano la dispersione dei rendimenti dell'asset attorno alla media; valori più elevati suggeriscono una maggiore volatilità e quindi un rischio di investimento più alto.
- **Asimmetria:** misura l'asimmetria della distribuzione dei rendimenti dell'asset. Un valore positivo indica un'asimmetria verso destra, con una predominanza di valori superiori alla media, mentre un valore negativo segnala un'asimmetria verso sinistra, con una prevalenza di valori inferiori alla media.
- **Curtosi:** valuta la distribuzione dei rendimenti rispetto a una distribuzione normale. Un valore elevato di curtosi indica una maggiore probabilità di eventi estremi rispetto alla media.

Queste statistiche forniscono una visione completa e dettagliata del comportamento degli asset finanziari, aiutandoci a prendere decisioni informate sugli investimenti.

TICKER	AVGO	AMD	LLY	PFE	BRK-B	BLK
MEDIA	0.0013	0.0015	0.0011	0.0002	0.0005	0.0005
VARIANZA	0.0005	0.0013	0.0003	0.0002	0.0001	0.0003
DEVIAZIONE STANDARD	0.0222	0.0362	0.0169	0.0144	0.0121	0.0171
ASIMMETRIA	-0.4277	0.4757	0.7395	0.1503	-0.2482	-0.1606
CURTOSI	9.3448	10.6656	10.7300	5.1215	11.2449	8.7427

L'asset con il rendimento maggiore (0.0015) è AMD ma possiede anche la deviazione standard maggiore (0.0362)

PFE ha rendimento minore ma volatilità molto contenuta.

Dalla varianza e dalla deviazione standard notiamo che la volatilità dei prezzi rispetto alla media è bassa e quindi gli asset scelti potrebbero essere considerati asset sicuri.

LLY ha l'asimmetria positiva maggiore quindi ha spesso ritorni positivi mentre AVGO ha asimmetria negativa che indica ritorni negativi frequenti.

BRK.B ha la curtosi maggiore quindi ha valori più estremi nella distribuzione dei guadagni.

Infine l'asset che più si avvicina alla normale è PFE che ha deviazione minore dell'asimmetria 0 (0.1503) e una curtosi meno alta (5.1215) rispetto alle altre. La curtosi della normale è 3.

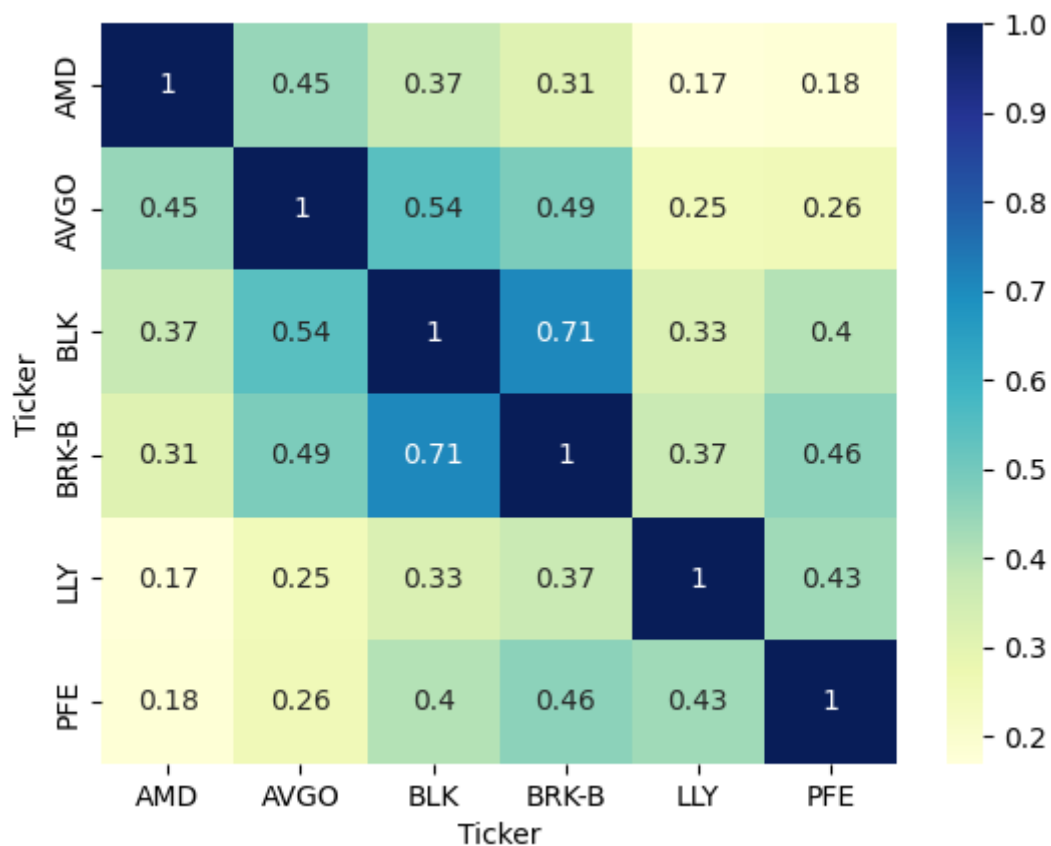
## Matrici di covarianza e di correlazione dei rendimenti

La matrice di covarianza fornisce informazioni sulla relazione lineare tra i rendimenti di due asset mentre la matrice di correlazione ci fornisce informazioni riguardo alla correlazione tra i vari asset.

### Matrice di covarianza

TICKER	AMD	AVGO	BLK	BRK-B	LLY	PFE
AMD	0.001312	0.000359	0.000233	0.000137	0.000104	0.000093
AVGO	0.000359	0.000492	0.000207	0.000130	0.000094	0.000081
BLK	0.000233	0.000207	0.000294	0.000147	0.000094	0.000099
BRK-B	0.000137	0.000130	0.000147	0.000147	0.000076	0.000081
LLY	0.000104	0.000094	0.000094	0.000076	0.000286	0.000105
PFE	0.000093	0.000081	0.000099	0.000081	0.000105	0.000206

### Matrice di correlazione



I titoli maggiormente correlati sono BLK e BRK.B (0.706424) che appartengono entrambi al settore finance.

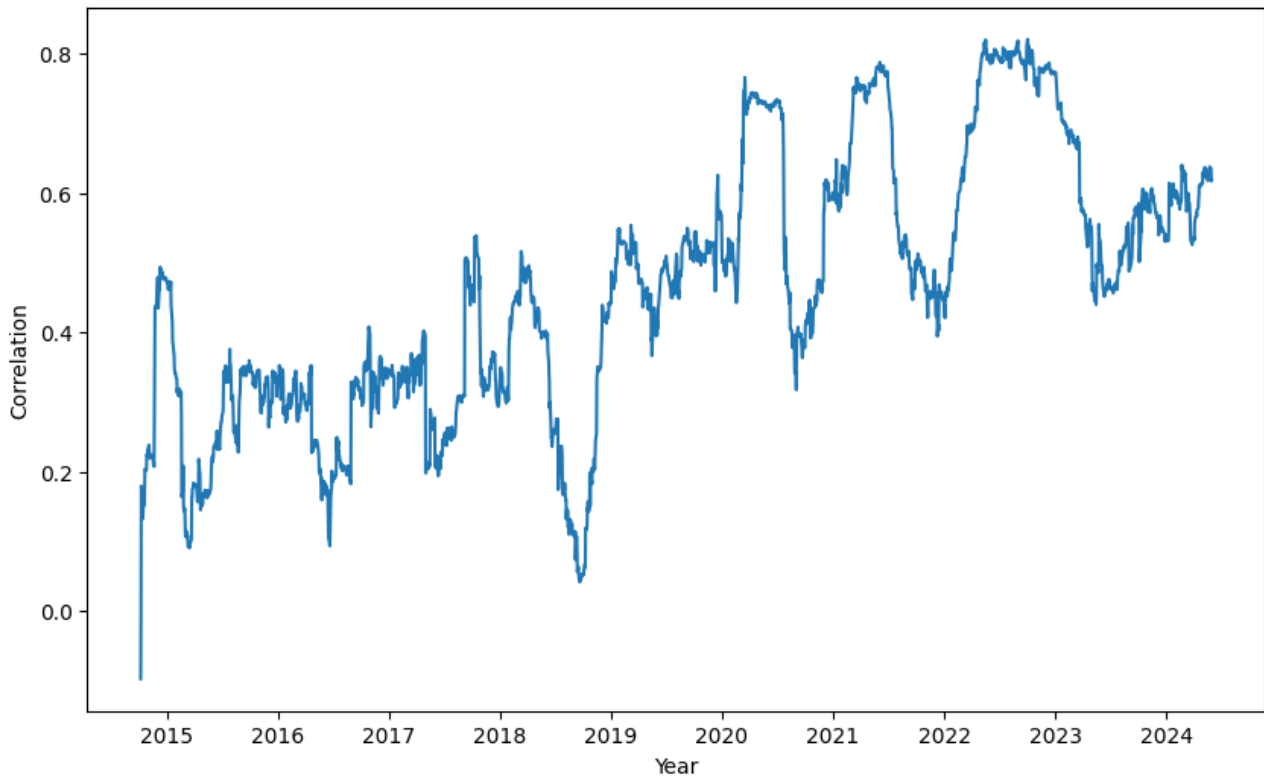
Invece i titoli a correlazione minore sono AMD e LLY (0.169330) e AMD con PFE (0.178133) che appartengono rispettivamente al settore IT e al settore Healthcare.

Bisogna tenere in considerazione che diversi fattori possono influenzare la correlazione tra i titoli:

- Settore di appartenenza
- Condizioni globale
- Scelte aziendali
- Cambiamenti nella tecnologia
- Politiche economiche

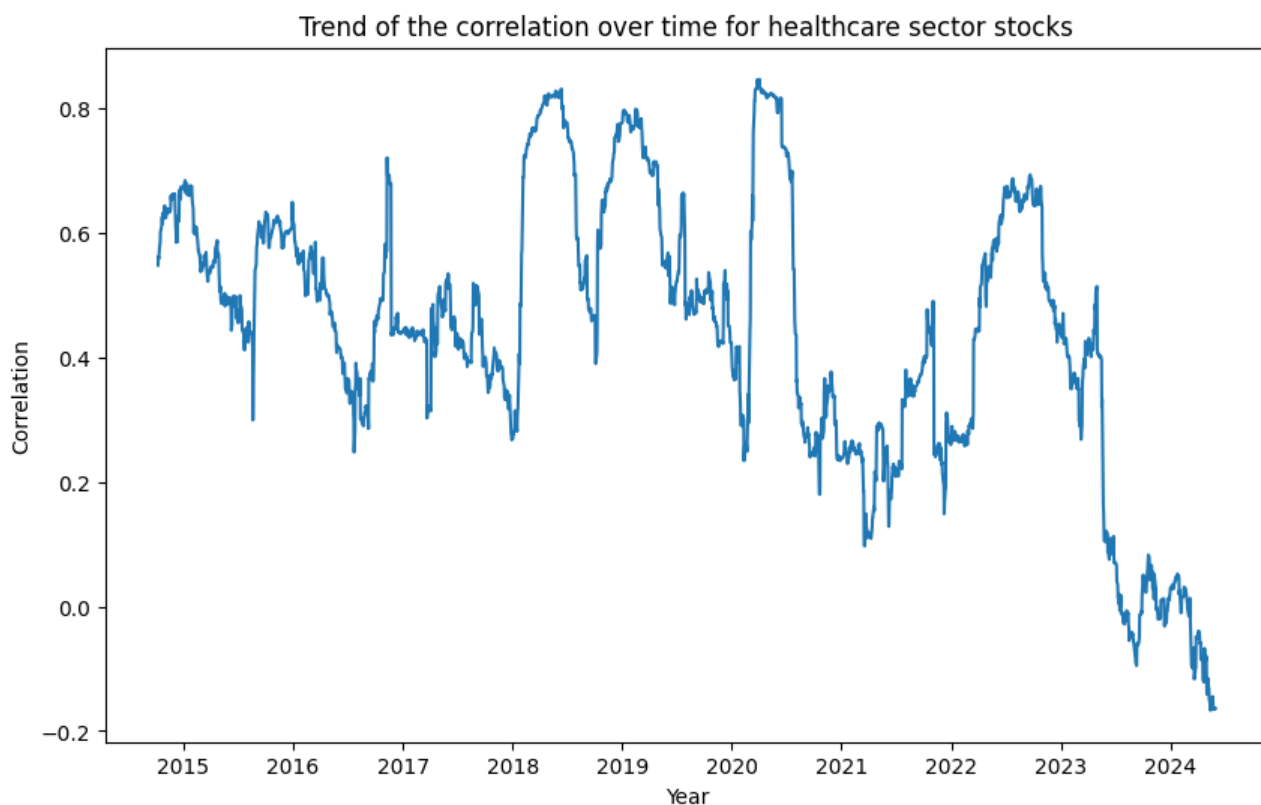
## Andamento nel tempo delle correlazioni tra gli asset e le correlazioni medie

Trend of the correlation over time for IT sector stocks

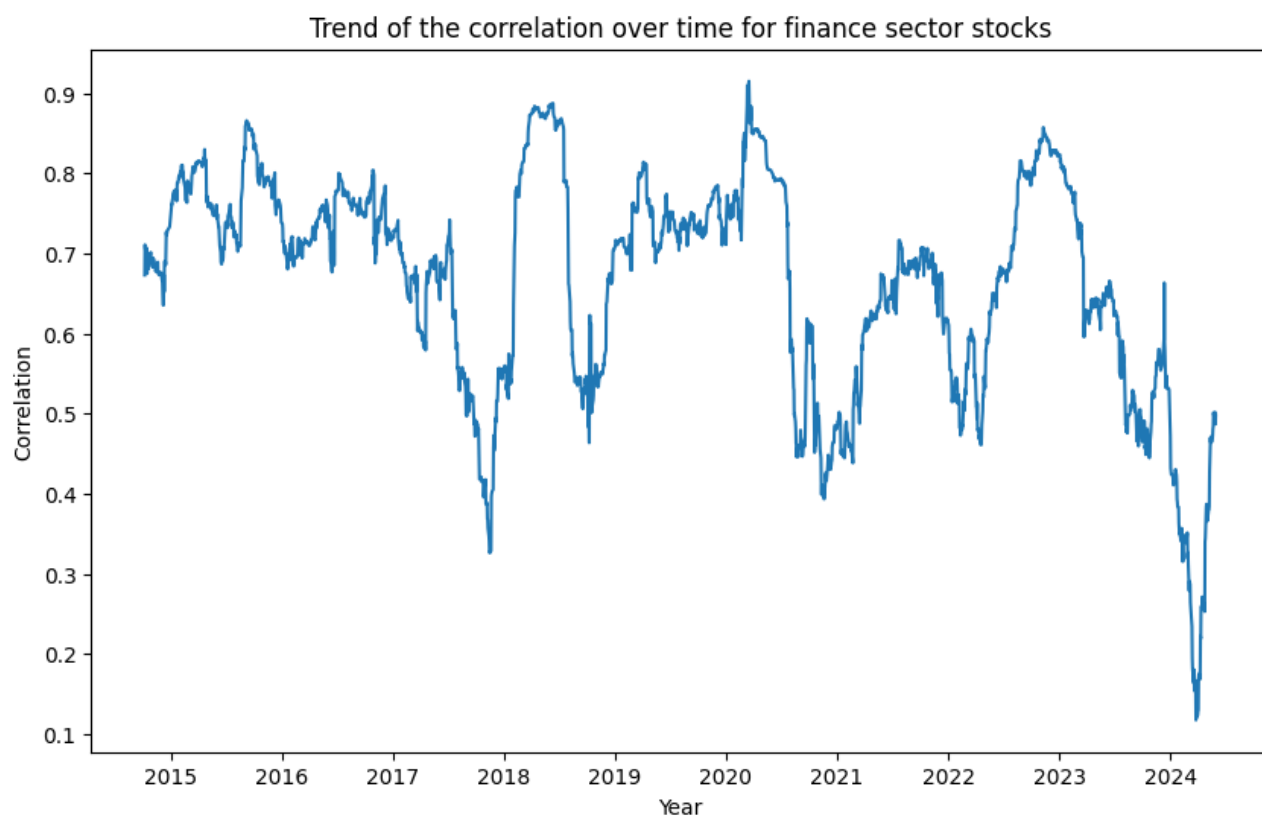


Nei primi due anni (2014-2016) la correlazione inizia bassa ma aumenta gradualmente. Dal 2017 al 2018 la correlazione si mantiene stabile mentre dal 2019 al 2020 si osservano oscillazioni marcate. Dal 2021 la correlazione raggiunge valori più alti indicando una sincronizzazione nelle performance delle due aziende.

In generale il trend di correlazione è positivo ma variabile.



Per il settore sanitario la correlazione inizia a livelli media e mostra fluttuazioni moderate. Nel 2019-2020 la correlazione mostra fluttuazioni più ampie probabilmente dovute alla pandemia di Covid-19 raggiungendo un picco di 0.8. Successivamente la correlazione tende a diminuire indicando che i titoli potrebbero muoversi in direzioni opposte. Questa correlazione negativa può essere interessante per investitori che cercano di diversificare.

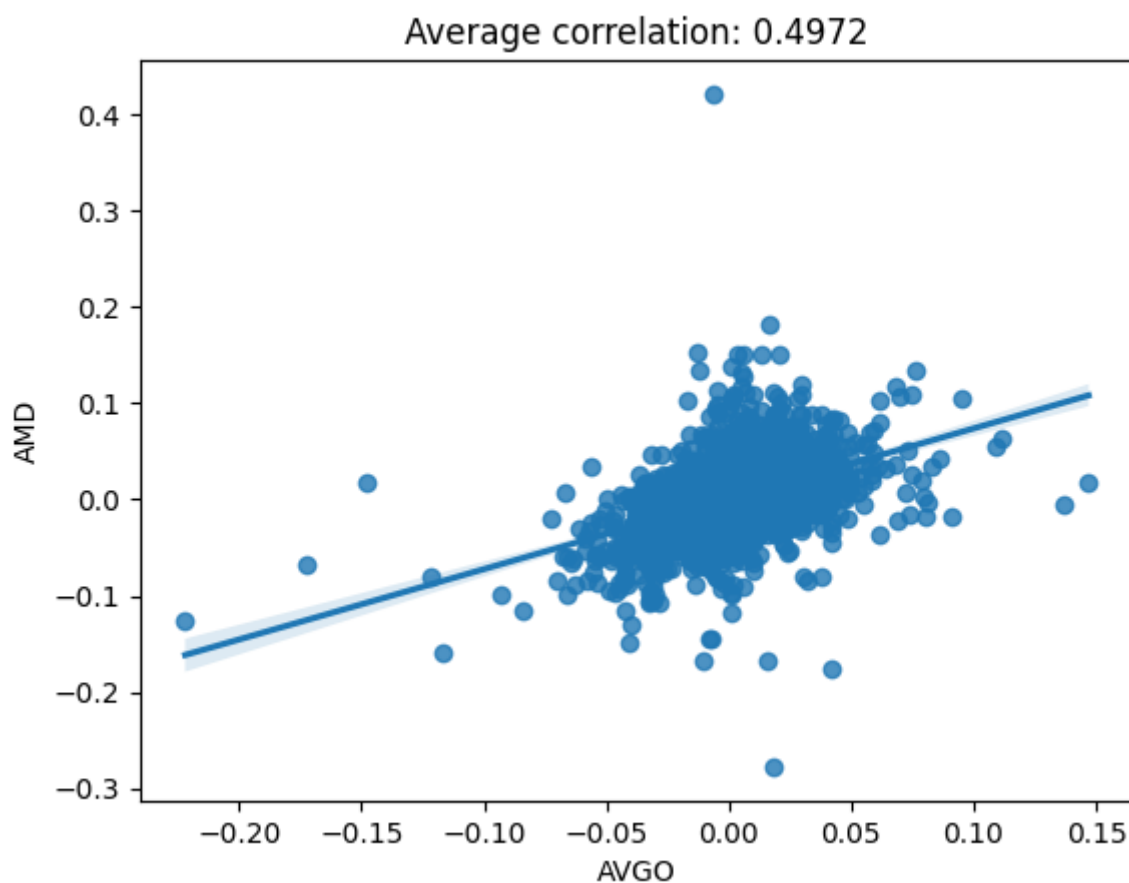


La correlazione tra BRK.B e BLK è generalmente alta, spesso sopra 0.6, con alcuni periodi che raggiunge valori vicini a 0.9. Questo indica una relazione forte e positiva tra i due titoli.

Ci sono significative fluttuazioni nel tempo indicando periodi di sincronizzazione più stretta e altri di leggera divergenza.

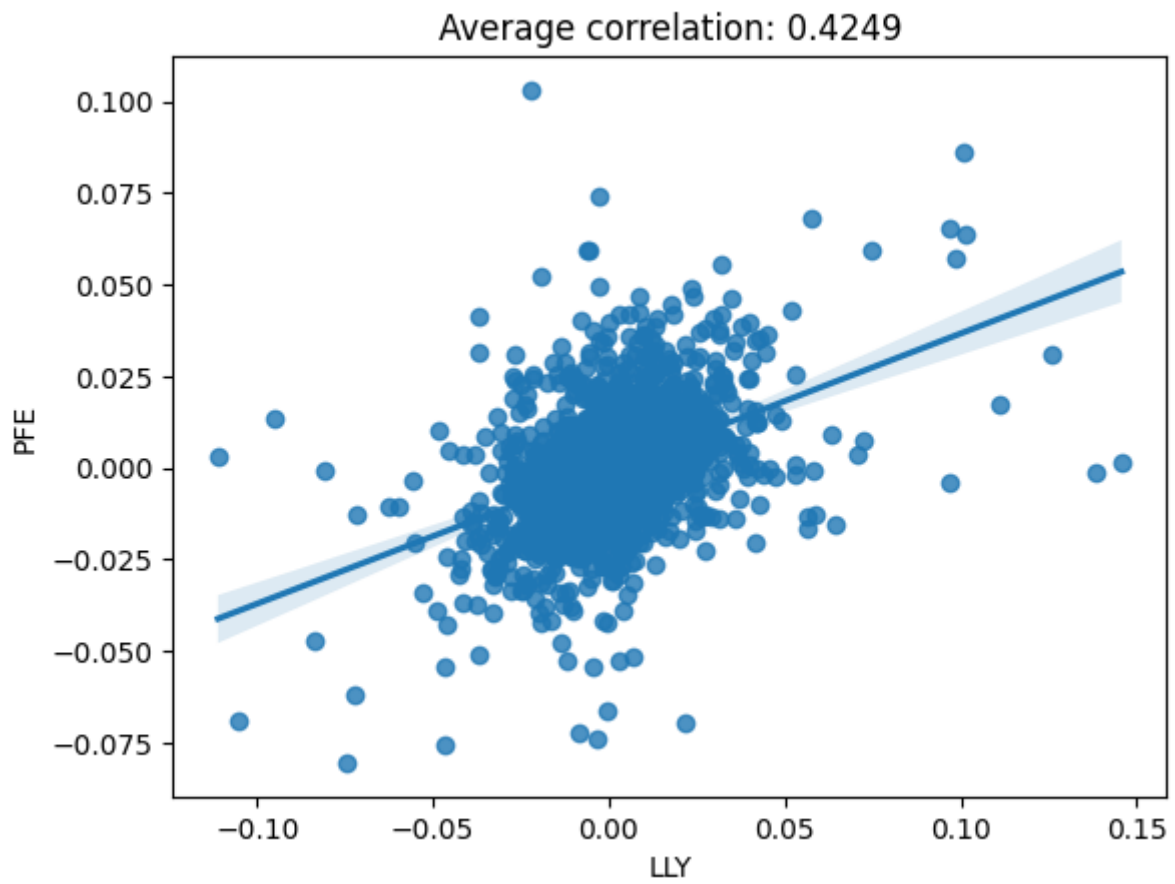


### Correlazione media dei titoli del settore tecnologico



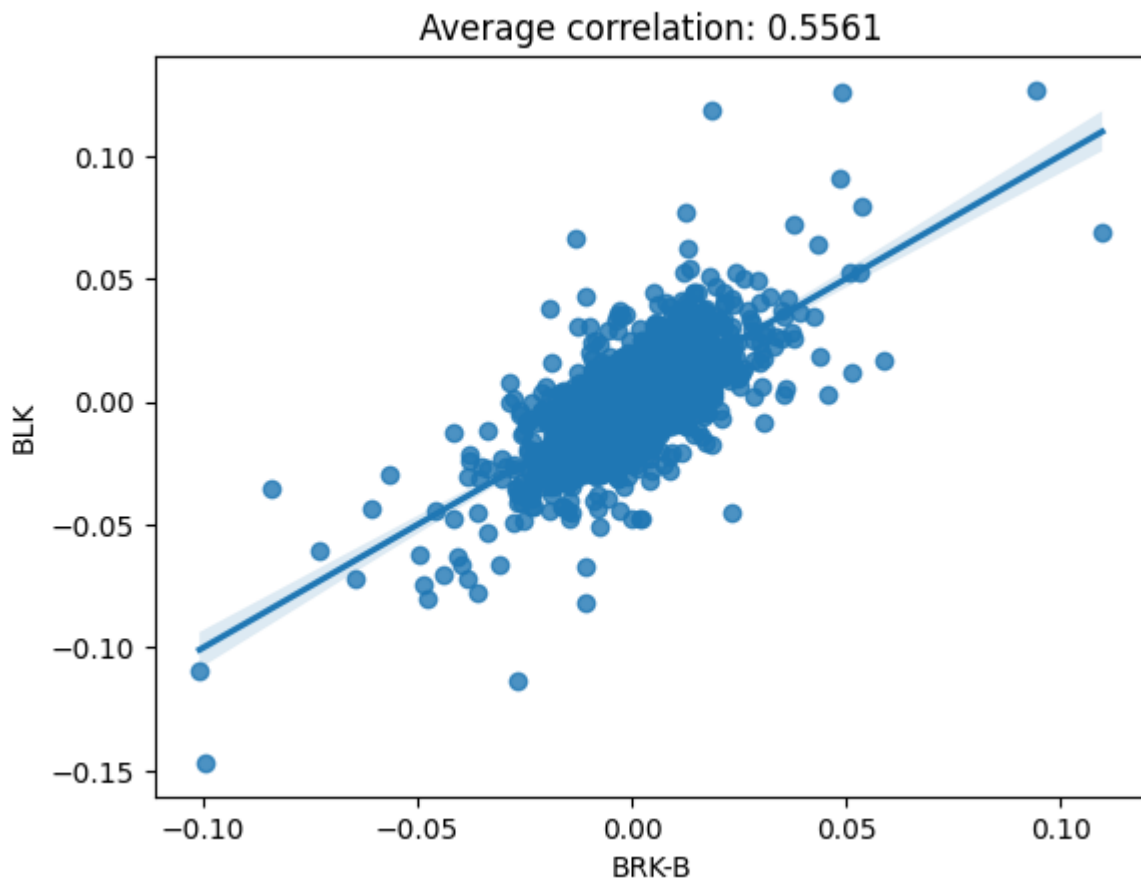
La correlazione è positiva quindi i titoli tendono a muoversi nella stessa direzione ma non in modo perfettamente allineato. Si nota una linea di regressione positiva che indica una relazione lineare tra i rendimenti dei due titoli ma la dispersione attorno ad essa indica che ci sono altri fattori che influenzano i rendimenti di ogni titolo.

### Correlazione media dei titoli del settore sanitario



La correlazione positiva è moderata quindi i rendimenti dei titoli tendono a muoversi nella stessa direzione ma non in modo perfettamente allineato. Si nota una linea di regressione positiva che indica una relazione lineare tra i rendimenti dei due titoli ma la dispersione attorno ad essa indica che ci sono altri fattori che influenzano i rendimenti.

### Correlazione media dei titoli del settore finanziario



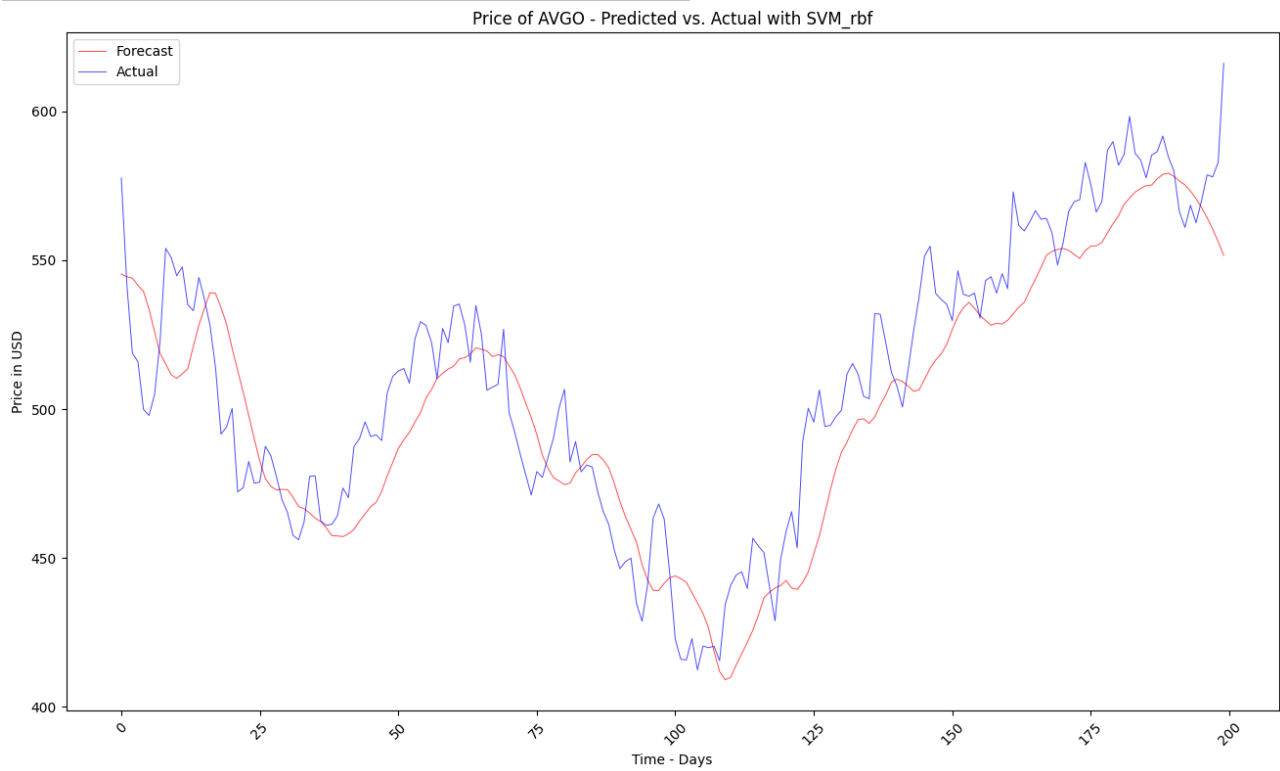
La correlazione è positiva quindi i rendimenti dei titoli tendono a muoversi nella stessa direzione ma non in modo perfettamente allineato. Si nota una linea di regressione positiva che indica una relazione lineare tra i rendimenti dei due titoli ma la dispersione attorno ad essa indica che ci sono altri fattori che influenzano i rendimenti.

# Analisi di previsione

Per l'analisi di previsione ho scelto di usare il modello Support Vector Machine (SVM) che è un algoritmo di apprendimento supervisionato.

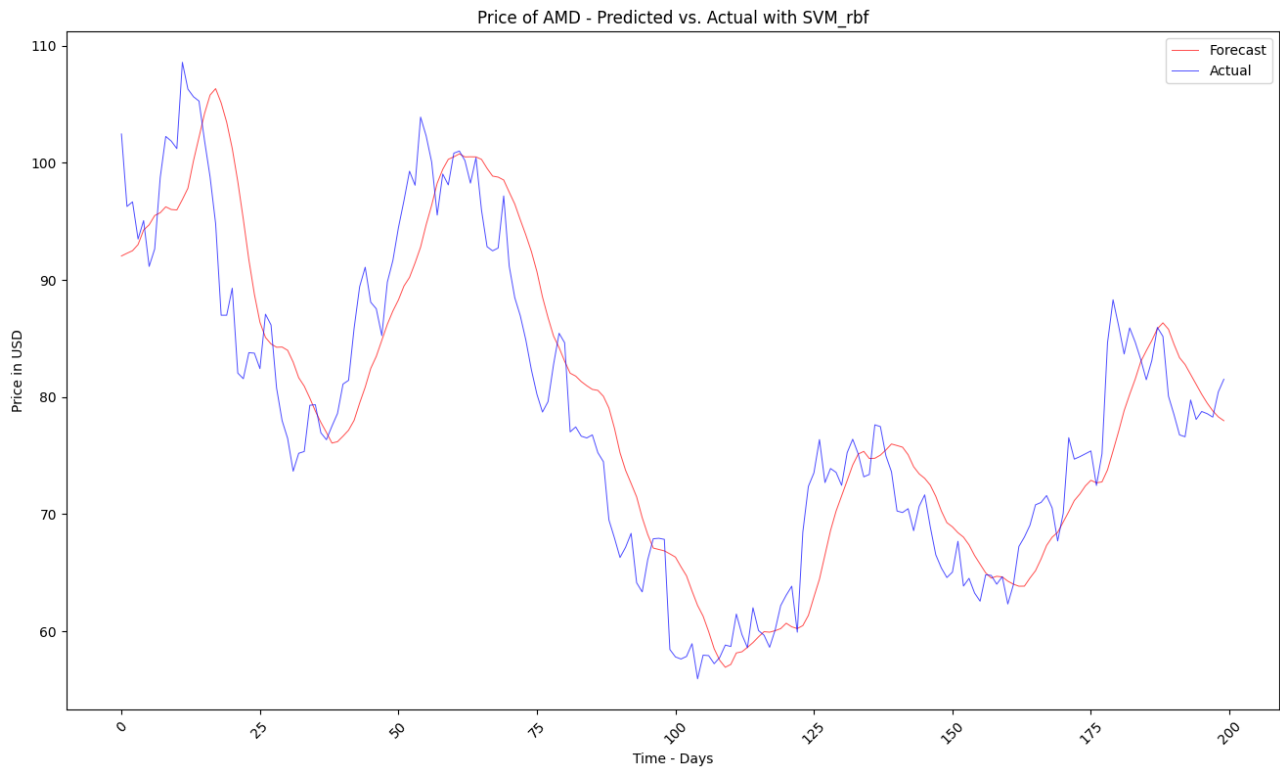
## Analisi AVGO

SVM Model Confidence Score	0.930938856499875
Mean Absolute Error	17.098031327448936
Mean Squared Error	468.48412871934084
R2 Score	0.902008957982275



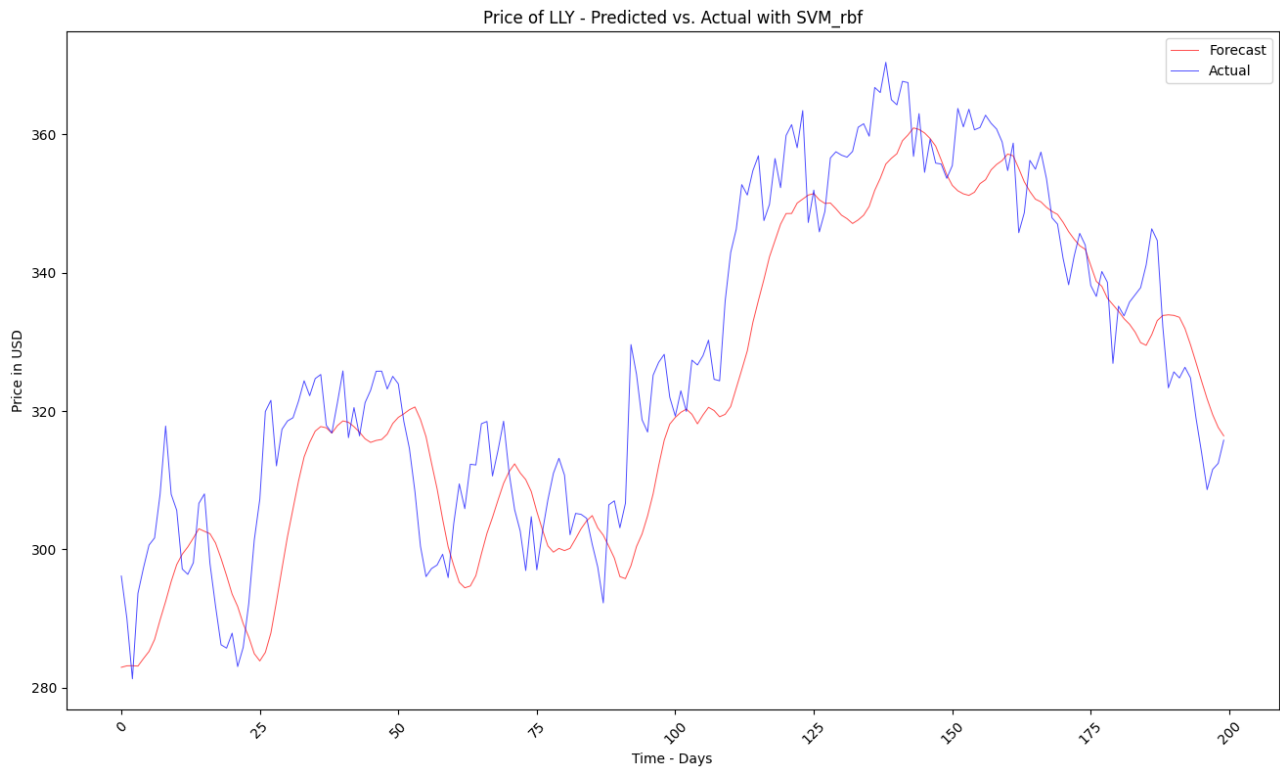
Analisi AMD

SVM Model Confidence Score	0.9295317141583804
Mean Absolute Error	5.252784451862971
Mean Squared Error	47.571228788361566
R2 Score	0.9002271303214141



Analisi LLY

<b>SVM Model Confidence Score</b>	0.9203884116186143
<b>Mean Absolute Error</b>	8.309602241364045
<b>Mean Squared Error</b>	115.89843251791758
<b>R2 Score</b>	0.9720322006292014



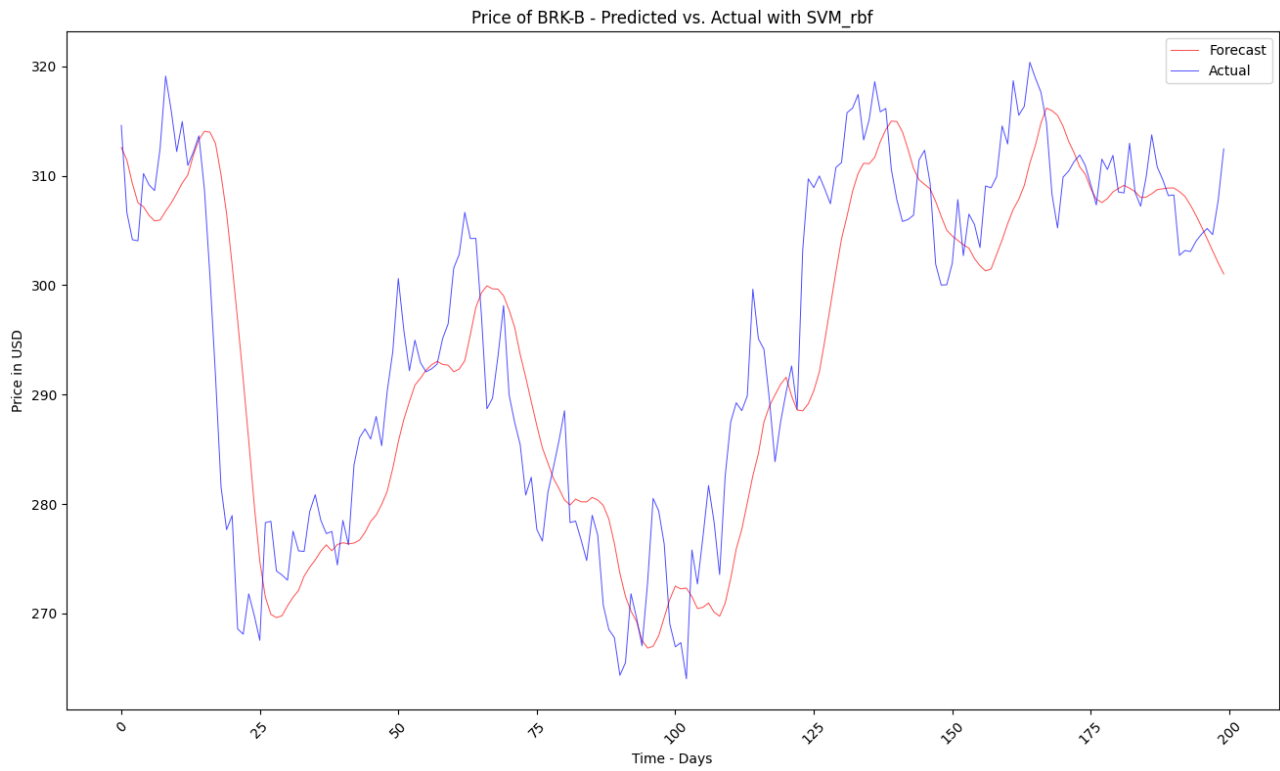
Analisi PFE

SVM Model Confidence Score	0.9453439652681435
Mean Absolute Error	1.234388776903777
Mean Squared Error	2.751768244502588
R2 Score	0.9334186843056151



Analisi BRK.B

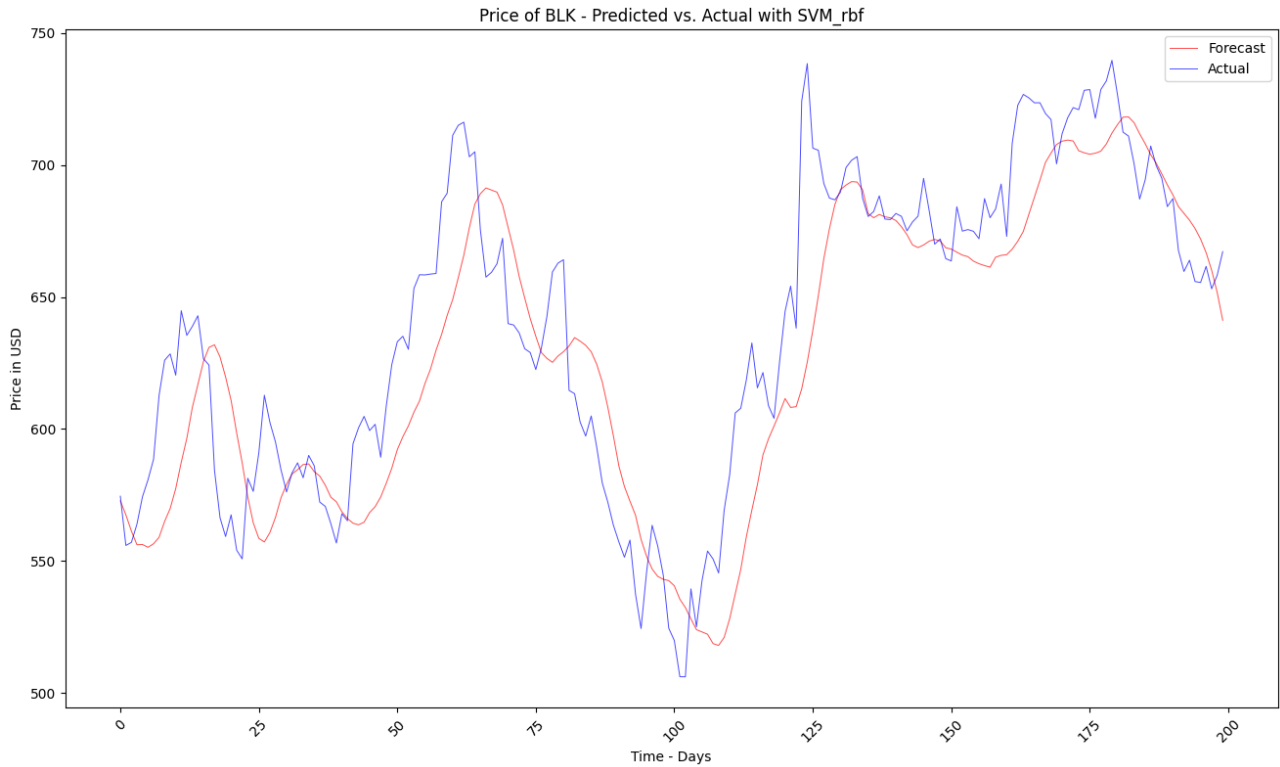
SVM Model Confidence Score	0.9418194634492808
Mean Absolute Error	6.057625538634839
Mean Squared Error	63.76416645169634
R2 Score	0.9372949005767744





## Analisi BLK

<b>SVM Model Confidence Score</b>	0.8632983112061992
<b>Mean Absolute Error</b>	24.349094264309915
<b>Mean Squared Error</b>	923.5901432447761
<b>R2 Score</b>	0.8993167283953558



PFE mostra le performance migliori complessive con confidenza elevata, MAE e MSE molto bassi e un ottimo  $R^2$ .

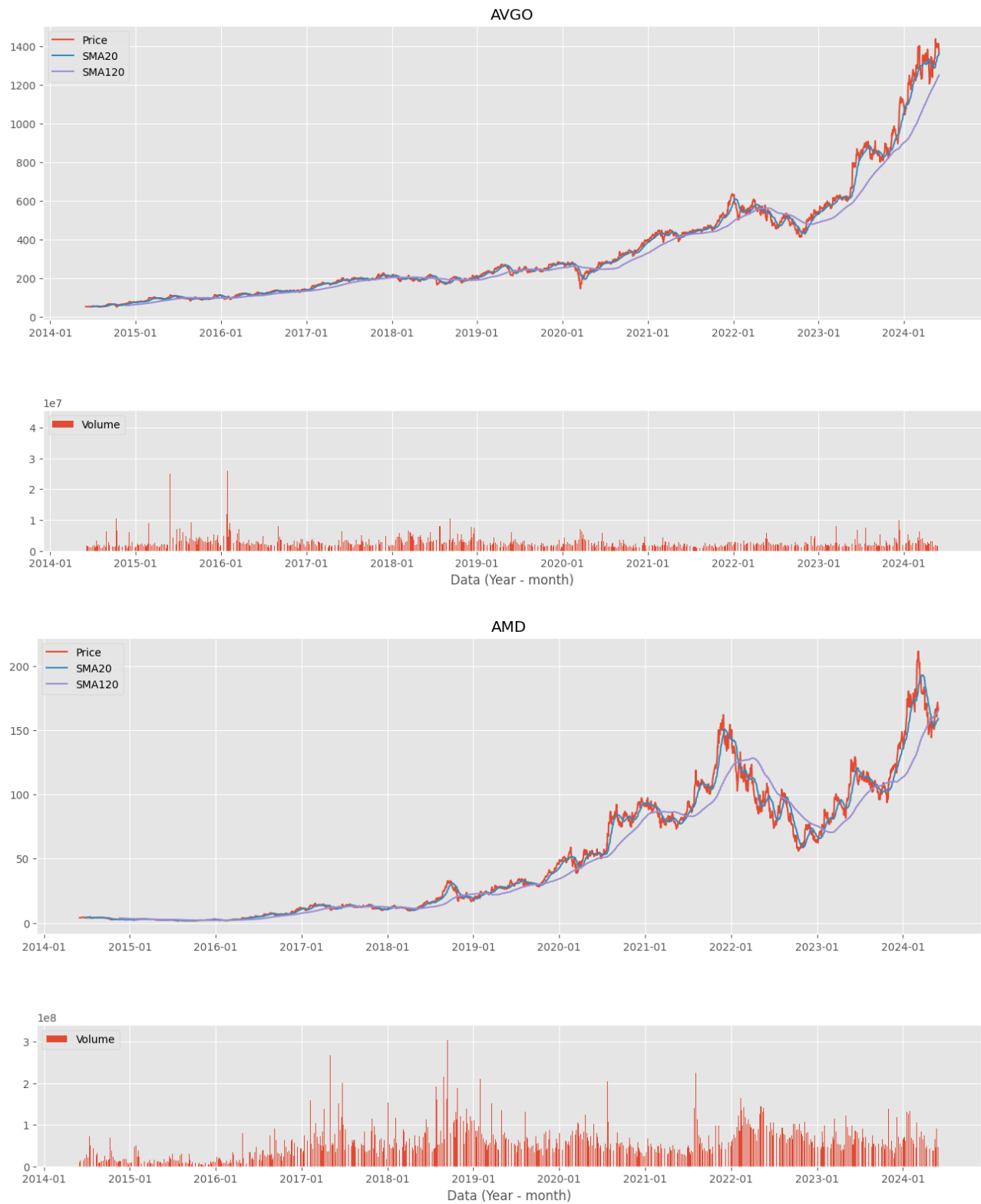
AMD e BRK.B hanno ottime performance con elevata confidenza e buoni valori di MAE e MSE.

LLY e AVGO hanno buone performance ma con errori significativi come indicato dai valori di MAE e MSE.

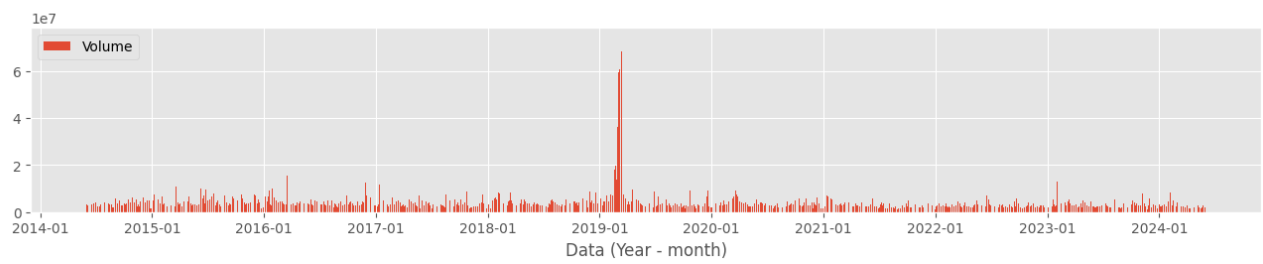
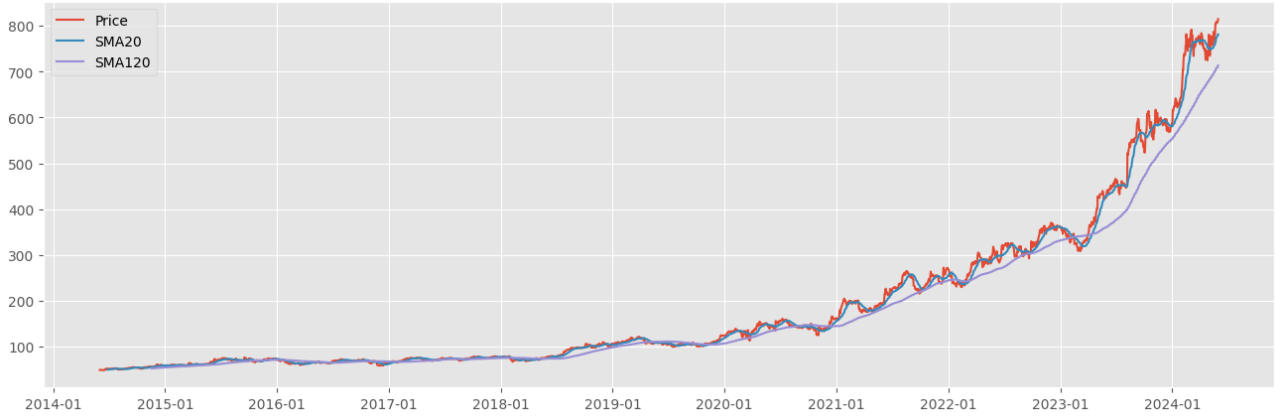
Infine BLK ha la performance più debole tra i titoli analizzati con la confidenza minore e alti valori di MAE e MSE suggerendo che il modello ha difficoltà a prevedere accuratamente i rendimenti di questo titolo.

# Strategie di trading e backtesting

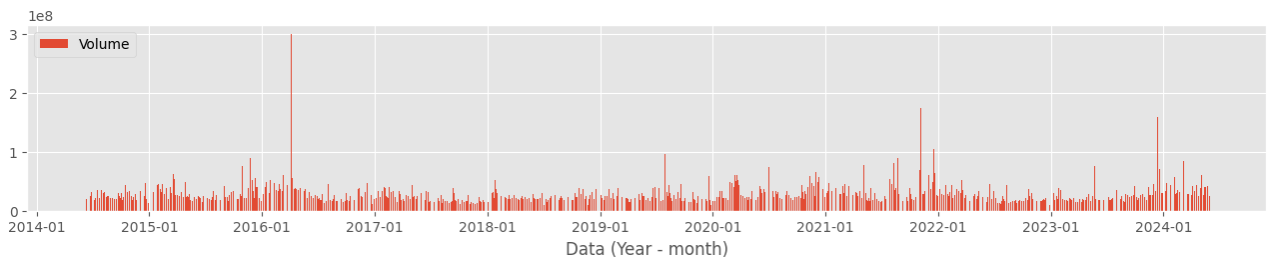
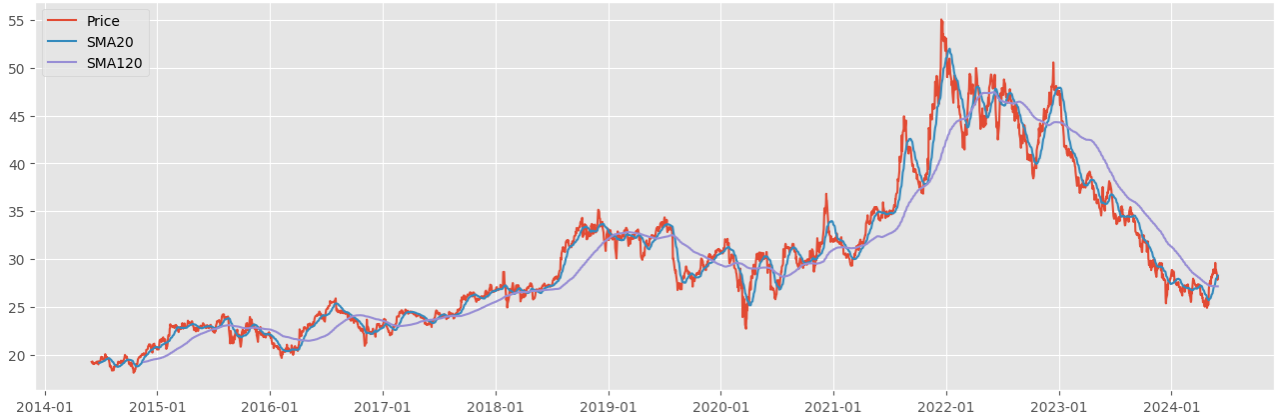
Ho utilizzato Moving Average Convergence Divergence (MACD) come indicatore per identificare i segnali di acquisto/vendita di titoli sui mercati finanziari per confrontarlo con la strategia base Buy and Hold.

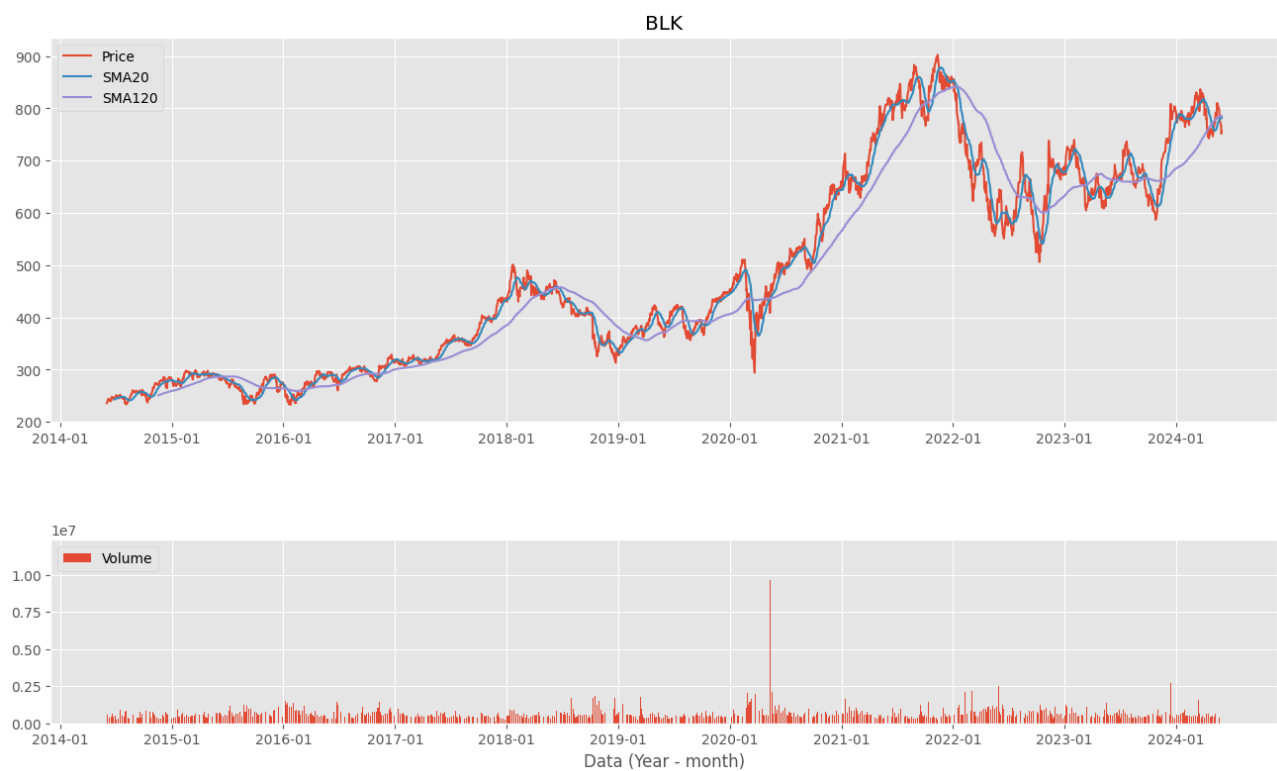
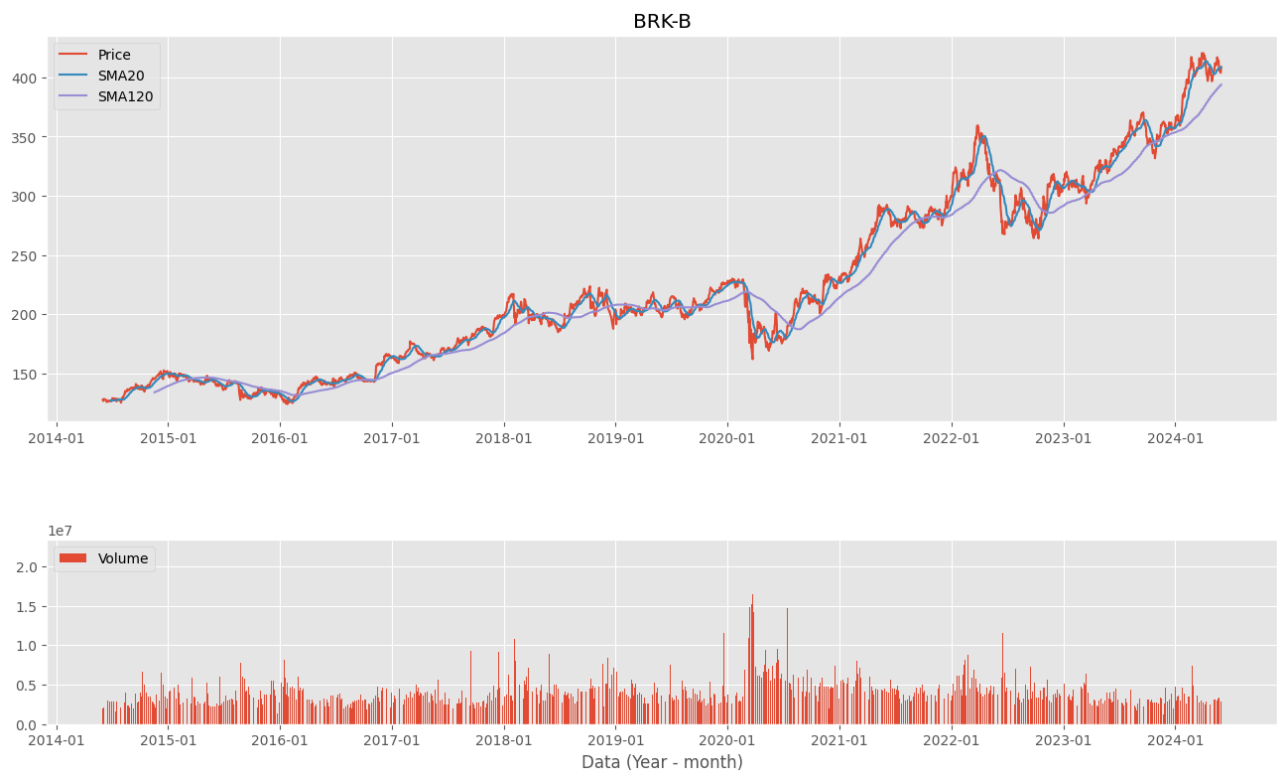


# LLY



# PFE

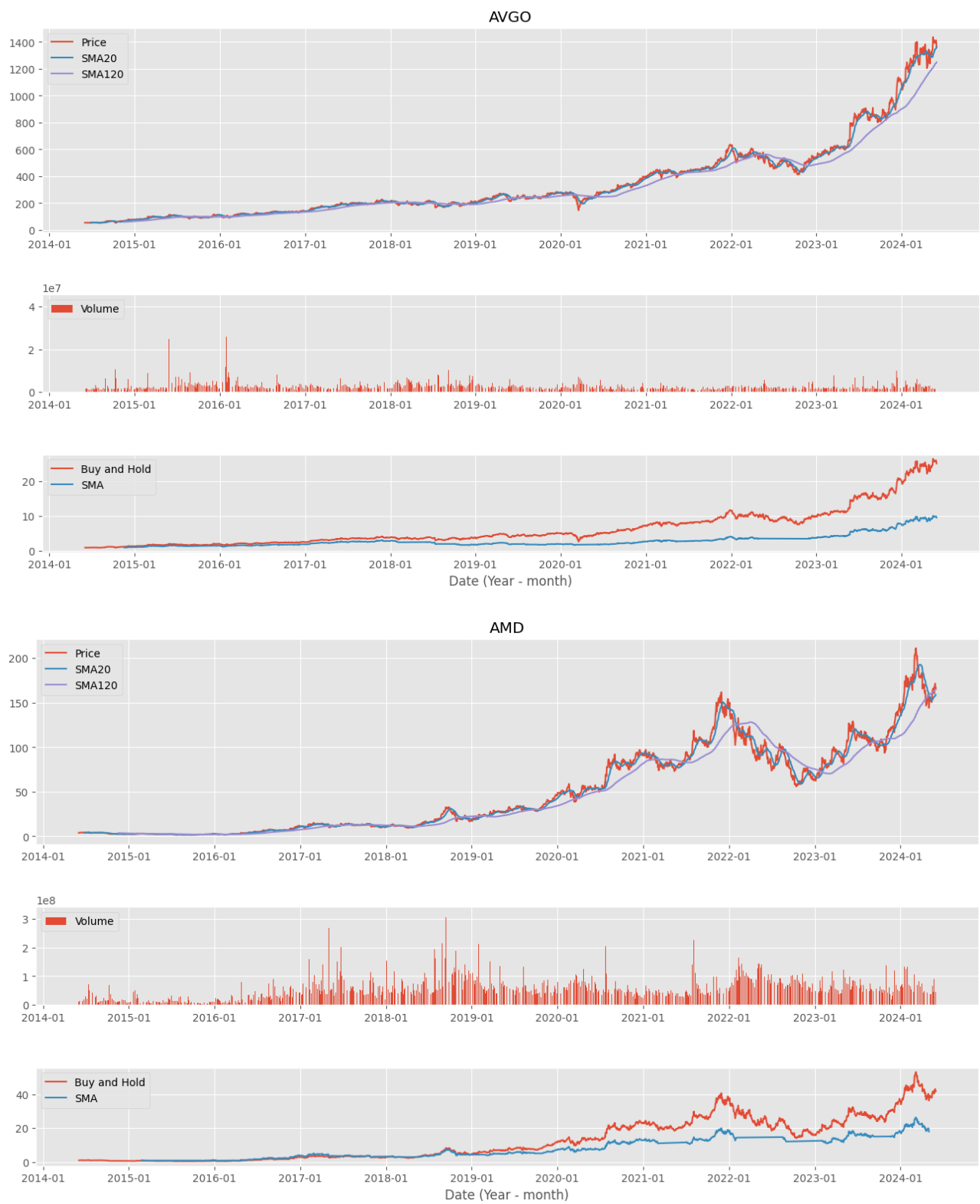




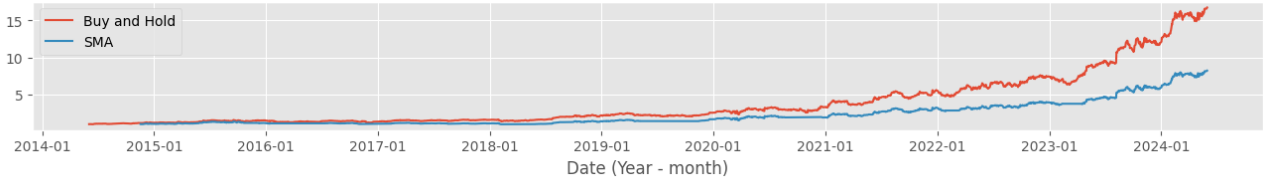
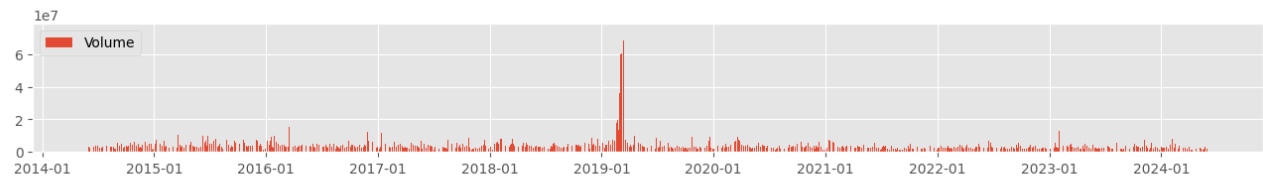
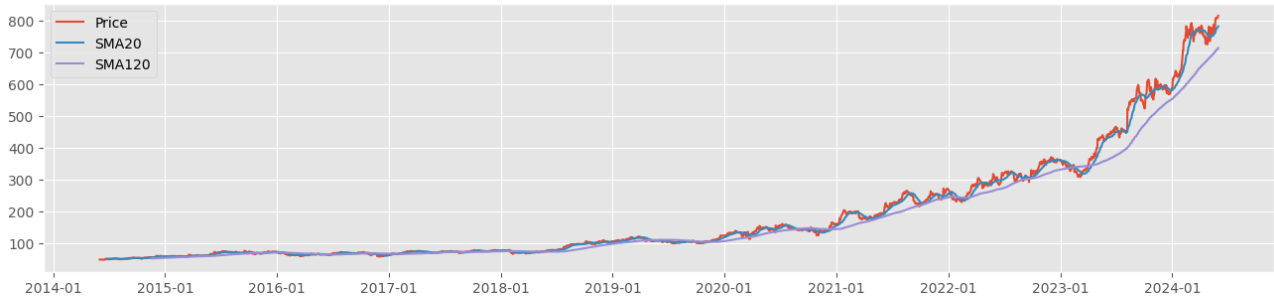
La strategia è stata implementata su due periodi:

- SMA20: calcolata sulla base degli ultimi 20 periodi prezzo
- SMA120: calcolata sulla base degli ultimi 120 periodi di prezzo

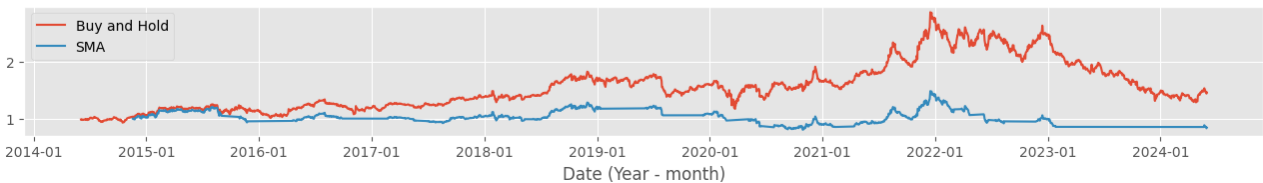
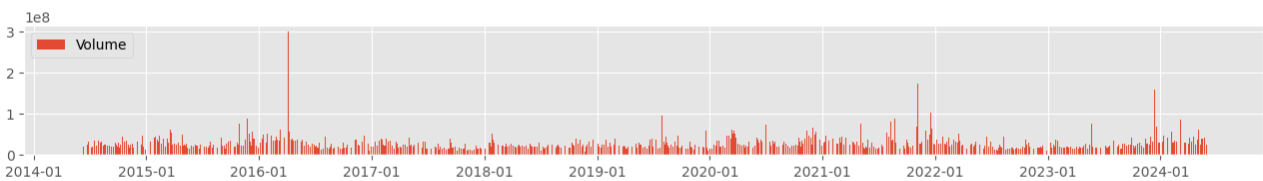
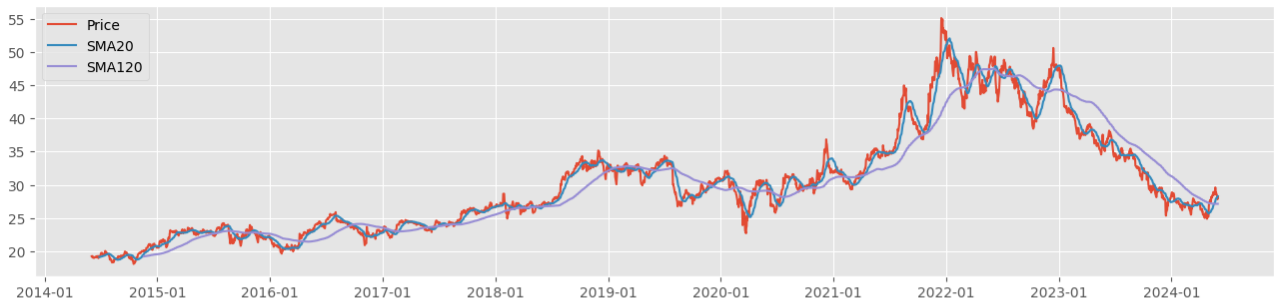
# Confronto con la strategia Buy and Hold

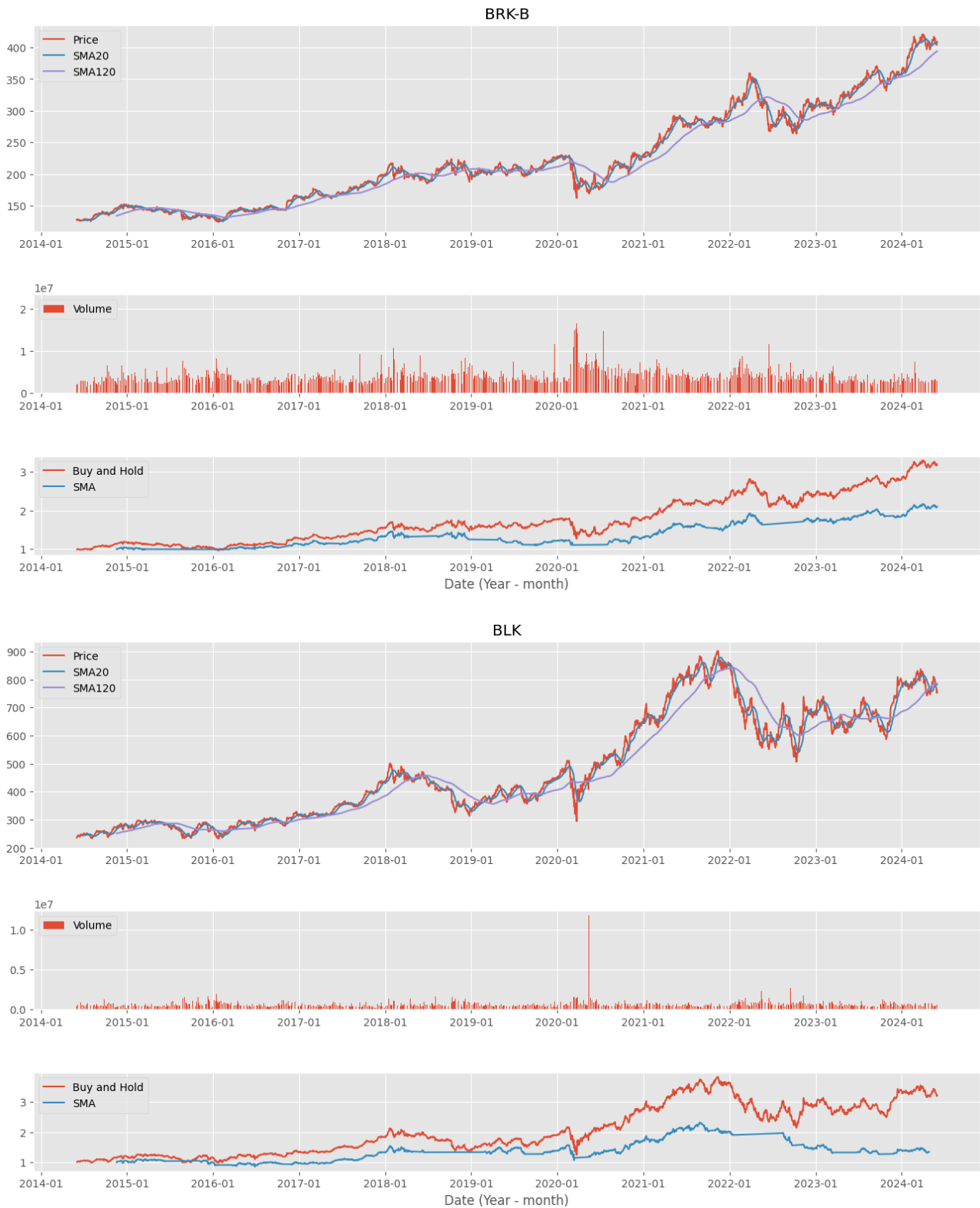


# LLY



# PFE





È evidente che la strategia Buy and Hold è una strategia migliore e più sicura rispetto alla strategia basata su media mobile.

## Capital Asset Pricing Model

Procederemo con l'utilizzo del modello Capital Asset Pricing Model (CAPM) per analizzare l'utile atteso dei nostri asset.

### AVGO OLS Regression Results

Dep. Variable:	excess_rtn	R-squared:	0.344			
Model:	OLS	Adj. R-squared:	0.327			
Method:	Least Squares	F-statistic:	19.92			
Date:	Tue, 18 Jun 2024	Prob (F-statistic):	1.88e-10			
Time:	10:08:44	Log-Likelihood:	149.73			
No. Observations:	118	AIC:	-291.5			
Df Residuals:	114	BIC:	-280.4			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0197	0.007	2.999	0.003	0.007	0.033
mkt	1.0351	0.146	7.097	0.000	0.746	1.324
smb	0.1124	0.243	0.463	0.644	-0.369	0.593
hml	-0.2747	0.171	-1.610	0.110	-0.613	0.063
=====						
Omnibus:	18.349	Durbin-Watson:	2.284			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	25.397			
Skew:	0.797	Prob(JB):	3.06e-06			
Kurtosis:	4.620	Cond. No.	39.1			
=====						

## AMD OLS Regression Results

Dep. Variable:	excess_rtn	R-squared:	0.363			
Model:	OLS	Adj. R-squared:	0.346			
Method:	Least Squares	F-statistic:	21.62			
Date:	Tue, 18 Jun 2024	Prob (F-statistic):	3.72e-11			
Time:	10:08:44	Log-Likelihood:	71.960			
No. Observations:	118	AIC:	-135.9			
Df Residuals:	114	BIC:	-124.8			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0226	0.013	1.788	0.076	-0.002	0.048
mkt	2.0822	0.282	7.385	0.000	1.524	2.641
smb	-0.0252	0.469	-0.054	0.957	-0.955	0.905
hml	-0.8083	0.330	-2.450	0.016	-1.462	-0.155
=====						
Omnibus:	2.232	Durbin-Watson:	1.970			
Prob(Omnibus):	0.328	Jarque-Bera (JB):	2.218			
Skew:	0.279	Prob(JB):	0.330			
Kurtosis:	2.625	Cond. No.	39.1			



# LLY OLS Regression Results

```

=====
Dep. Variable:          excess_rtn      R-squared:                0.104
Model:                  OLS             Adj. R-squared:           0.080
Method:                 Least Squares   F-statistic:             4.412
Date:                  Tue, 18 Jun 2024 Prob (F-statistic):       0.00565
Time:                  10:08:44         Log-Likelihood:          154.41
No. Observations:      118             AIC:                     -300.8
Df Residuals:          114             BIC:                     -289.7
Df Model:               3
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0210	0.006	3.328	0.001	0.008	0.033
mkt	0.3181	0.140	2.269	0.025	0.040	0.596
smb	0.0590	0.233	0.253	0.801	-0.403	0.521
hml	-0.4501	0.164	-2.744	0.007	-0.775	-0.125

```

=====
Omnibus:                7.412      Durbin-Watson:           1.765
Prob(Omnibus):          0.025      Jarque-Bera (JB):        7.084
Skew:                   0.511      Prob(JB):                0.0290
Kurtosis:               3.628      Cond. No.:               39.1
=====

```

# PFE OLS Regression Results

```

=====
Dep. Variable:          excess_rtn      R-squared:                0.220
Model:                  OLS             Adj. R-squared:           0.199
Method:                 Least Squares   F-statistic:             10.70
Date:                  Tue, 18 Jun 2024 Prob (F-statistic):       2.98e-06
Time:                  10:08:44         Log-Likelihood:          170.02
No. Observations:      118             AIC:                     -332.0
Df Residuals:          114             BIC:                     -321.0
Df Model:               3
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-0.0039	0.006	-0.709	0.480	-0.015	0.007
mkt	0.6917	0.123	5.632	0.000	0.448	0.935
smb	-0.4709	0.204	-2.303	0.023	-0.876	-0.066
hml	0.0015	0.144	0.010	0.992	-0.283	0.286

```

=====
Omnibus:                14.588      Durbin-Watson:           1.909
Prob(Omnibus):          0.001      Jarque-Bera (JB):        27.226
Skew:                   0.506      Prob(JB):                1.22e-06
Kurtosis:               5.125      Cond. No.:               39.1
=====

```

## BRK.B OLS Regression Results

```

=====
Dep. Variable:          excess_rtn      R-squared:                0.695
Model:                  OLS             Adj. R-squared:           0.687
Method:                 Least Squares   F-statistic:             86.41
Date:                   Tue, 18 Jun 2024 Prob (F-statistic):       3.14e-29
Time:                   10:08:44        Log-Likelihood:          256.01
No. Observations:      118             AIC:                     -504.0
Df Residuals:          114             BIC:                     -492.9
Df Model:               3
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0011	0.003	0.422	0.674	-0.004	0.006
mkt	0.9105	0.059	15.365	0.000	0.793	1.028
smb	-0.5782	0.099	-5.860	0.000	-0.774	-0.383
hml	0.3232	0.069	4.662	0.000	0.186	0.461

```

=====
Omnibus:                0.413      Durbin-Watson:           1.950
Prob(Omnibus):          0.813      Jarque-Bera (JB):         0.477
Skew:                   -0.137     Prob(JB):                 0.788
Kurtosis:               2.851      Cond. No.:                39.1
=====

```

## BLK OLS Regression Results

```

=====
Dep. Variable:          excess_rtn      R-squared:                0.709
Model:                  OLS             Adj. R-squared:           0.701
Method:                 Least Squares   F-statistic:             92.56
Date:                   Tue, 18 Jun 2024 Prob (F-statistic):       2.02e-30
Time:                   10:08:44        Log-Likelihood:          213.17
No. Observations:      118             AIC:                     -418.3
Df Residuals:          114             BIC:                     -407.3
Df Model:               3
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	-0.0011	0.004	-0.284	0.777	-0.009	0.006
mkt	1.3475	0.085	15.817	0.000	1.179	1.516
smb	-0.0377	0.142	-0.266	0.791	-0.319	0.243
hml	0.1791	0.100	1.797	0.075	-0.018	0.377

```

=====
Omnibus:                14.863     Durbin-Watson:           2.050
Prob(Omnibus):          0.001      Jarque-Bera (JB):         18.562
Skew:                   0.707      Prob(JB):                 9.32e-05
Kurtosis:               4.332      Cond. No.:                39.1
=====

```

Grazie al modello CAPM possiamo calcolare i Beta dei titoli rispetto al mercato:

TICKER	BETA	VOLATILITÀ
AVGO	1.2122270420392738	Maggiore rispetto al mercato
AMD	2.140762152610683	Maggiore rispetto al mercato
LLY	0.3319090954248763	Minore rispetto al mercato
PFE	0.6544597446384158	Minore rispetto al mercato
BRK.B	0.7956732278617329	Minore rispetto al mercato
BLK	1.5144790850240815	Maggiore rispetto al mercato

Grazie ai Beta possiamo stimare i rendimenti attesi per ogni asset:

TICKER	RENDIMENTO ATTESO	RENDIMENTO ATTESO PERCENTUALE
AVGO	0.14871095747641178	14.9%
AMD	0.25222311526498586	25.2%
LLY	0.0505740005034152	5.1%
PFE	0.0865316216165817	8.7%
BRK.B	0.10227395832616595	10.2%
BLK	0.18240571043638204	18.2%

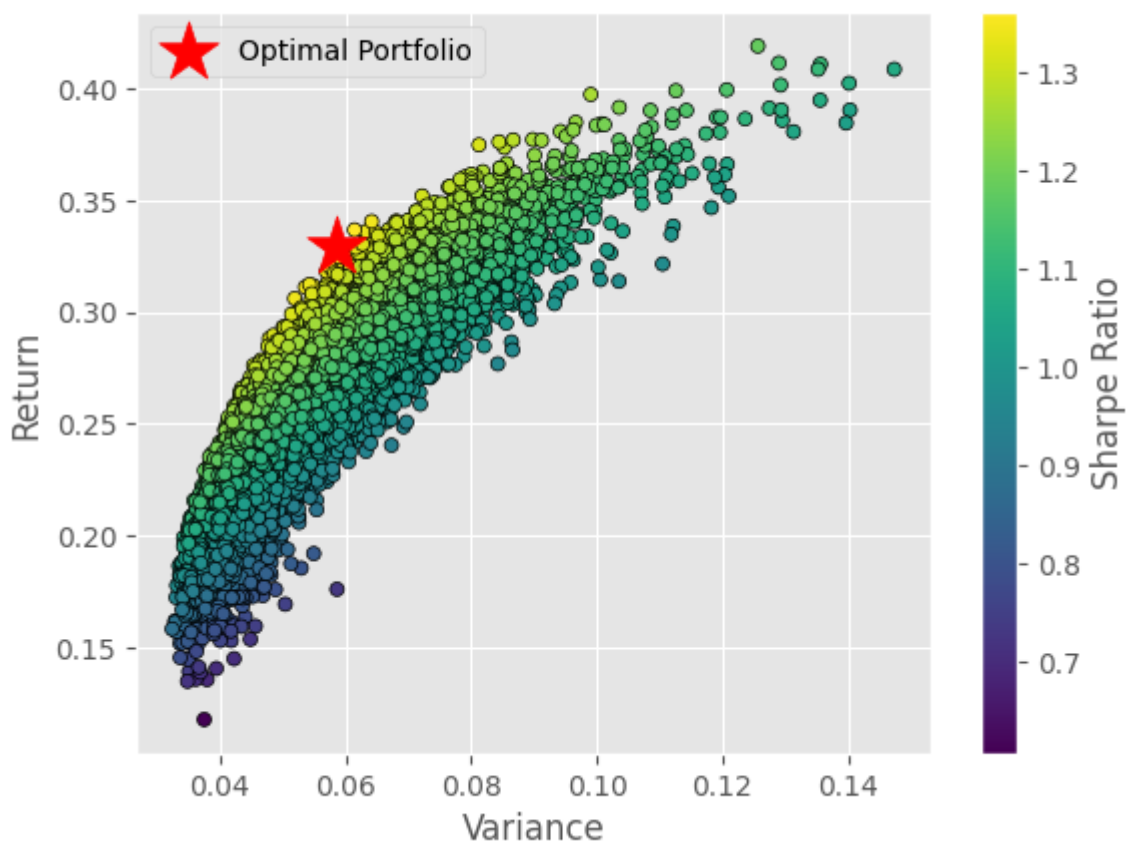
Vediamo che il range dei rendimenti è compreso tra il 5.1% e il 25.2%, ovviamente si tratta di una stima e quindi per piccole variazioni mercato queste stime potrebbero venire invalidate.

## Costruzione di un portafoglio

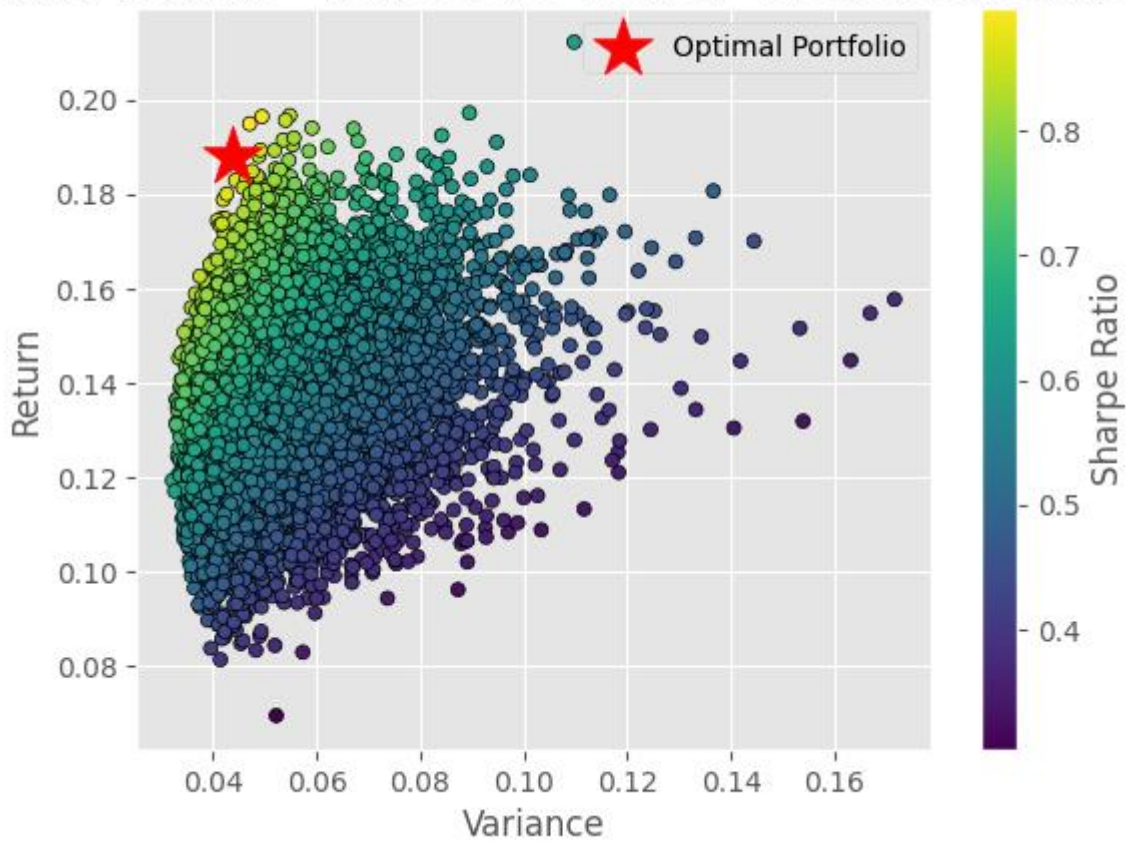
In questa sezione viene costruito un portafoglio in termini di media-varianza sugli ultimi 108 mesi di dati.

Per il metodo analitico usiamo la frontiera efficiente per visualizzare il portafoglio ottimale basato sui rendimenti passati e su quelli attesi.

### Optimal Portfolio - Monte Carlo Method - Historical Returns



## Optimal Portfolio - Monte Carlo Method - Predicted Returns



Il Beta calcolato per il portafoglio ottimale è **1.5655037174429673** superiore al Beta standard per SP500 pari a 1.0. Di conseguenza il portafoglio è più volatile rispetto al mercato.

Invece il Beta del portafoglio effettivo (dove tutti gli asset hanno lo stesso peso) è **0.13711989393732374** che è inferiore rispetto al Beta di SP500 (1.0). Di conseguenza il rischio del portafoglio è minore.

## Conclusioni

In base ai dati esaminati, possiamo concludere che essi seguono una distribuzione normale, come evidenziato dalle statistiche descrittive. Inoltre, la deviazione standard dei rendimenti logaritmici mostra una bassa variabilità rispetto alla media, suggerendo stabilità nei dati. Gli asset appartenenti allo stesso settore mostrano una forte correlazione tra di loro.

Il modello di previsione che abbiamo costruito ci consente di stimare l'andamento del mercato con un certo grado di approssimazione, tenendo conto esclusivamente dei dati utilizzati per l'addestramento e non considerando eventi straordinari nei mercati finanziari.

Per quanto riguarda il Beta dei singoli titoli, abbiamo scoperto che metà degli asset presenta una volatilità superiore a quella del mercato. La strategia di previsione analizzata ha mostrato risultati peggiori rispetto alla tradizionale strategia di "Buy and Hold".

Nella costruzione del portafoglio, abbiamo identificato i portafogli ottimali in termini di rapporto rendimento/varianza, sia basandoci sui rendimenti passati che su quelli previsti.