

Business Intelligence per i Servizi Finanziari

Presentazione del progetto
Alessandro Zanotti - 885892

Agenda

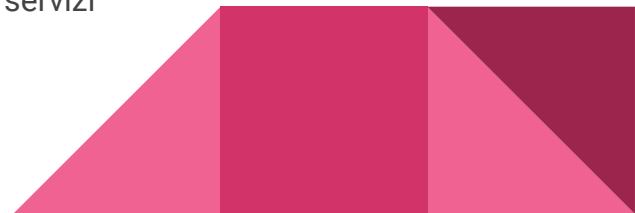
- ❑ Sommario dei dati utilizzati
- ❑ Statistiche descrittive
- ❑ Analisi di previsione
- ❑ Strategie di trading e backtesting
- ❑ Capital Asset Pricing Model
- ❑ Costruzione di un portafoglio
- ❑ Conclusioni



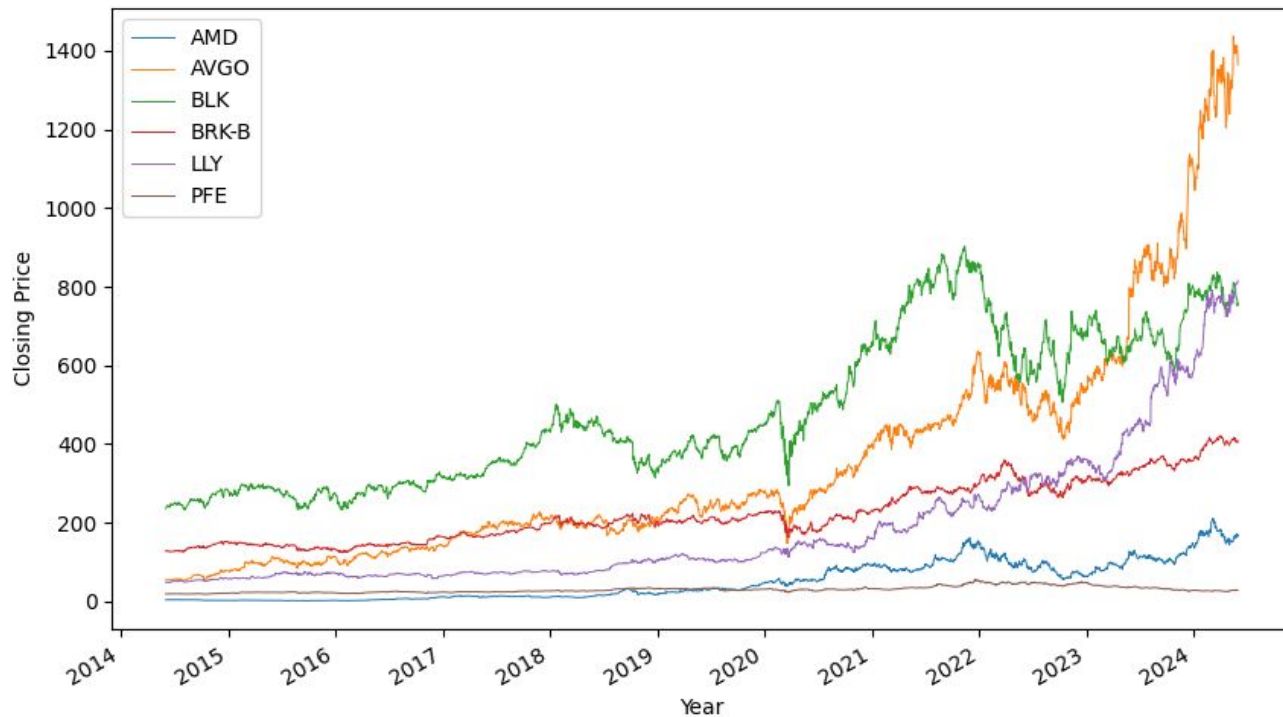
Sommario dei dati utilizzati

Descrizione di ciascun titolo scelto

- ❑ Settore tecnologico:
 - ❑ Broadcom, Inc (AVGO): leader tecnologico del settore microchip e nel settore software enterprise. Recentemente ha acquisito VMware Inc (VMW) per 69 miliardi.
 - ❑ Advanced Micro Devices, Inc (AMD): AMD sviluppa processi per PC e tecnologie collegate. Scelto per gli alti ritorni.
- ❑ Settore sanitario:
 - ❑ Eli Lilly and Company (LLY): azienda farmaceutica globale che sviluppa, produce e commercia in vari settori sanitari. Possiede una serie di farmaci di successo e ha superato trial clinici rilevanti.
 - ❑ Pfizer Inc. (PFE): azienda biofarmaceutica globale che sviluppa farmaci innovativi e vaccini. Scelto per il contributo durante la pandemia Covid-19 del 2020.
- ❑ Settore finanziario:
 - ❑ Berkshire Hathaway Inc. Class B (BRK.B): conglomerato multinazionale amministrato da Warren Buffett. Scelto per lo storico record di crescita e il portafoglio selezionato.
 - ❑ BlackRock, Inc (BLK): asset manager più grande al mondo che fornisce servizi di consulenza. Scelto l'approccio innovativo alla gestione degli asset.



Andamento dei prezzi di chiusura



Questo grafico mostra l'andamento dei prezzi di chiusura dal 31-05-2014 al 31-05-2024.

Statistiche descrittive

Rendimenti cumulati annui

Data\Ticker	AVGO	AMD	LLY	PFE	BRK-B	BLK
2015-12-31	1.547897	0.665537	1.245654	1.159322	1.021807	1.073686
2016-12-31	1.962241	1.499882	1.238859	1.169407	1.050108	1.111680
2017-12-31	3.001753	3.542592	1.368636	1.261269	1.276536	1.379621
2018-12-31	3.137660	4.913436	1.601322	1.504659	1.483872	1.638927
2019-12-31	3.930686	8.545907	2.036873	1.587515	1.517006	1.535193
2020-12-31	4.604207	18.729458	2.656382	1.523711	1.494680	1.969557
2021-12-31	7.311707	28.869481	4.083911	1.914985	1.988267	3.049261
2022-12-31	8.239187	25.941535	5.670284	2.336701	2.219957	2.535283
2023-12-31	12.293508	29.407074	8.669035	1.793123	2.444005	2.626449
2024-12-31	20.324127	48.227003	13.773212	1.391353	2.912060	3.067892

Il rendimento cumulato mostra la crescita totale, in percentuale, dell'asset nel tempo.

Un rendimento cumulato superiore ad uno indica un rendimento positivo
Un rendimento cumulato pari ad uno indica nessun rendimenti.
Infine un rendimento cumulato inferiore ad uno indica una perdita.

L'asset con crescita maggiore è AMD con il 4800%.

L'asset con crescita minore è PFE con crescita del 139%.

Rendimenti composti annui

Ticker	AVGO	AMD	LLY	PFE	BRK.B	BLK
CAGR	38.0%	45.3%	32.5%	3.9%	12.3%	12.4%

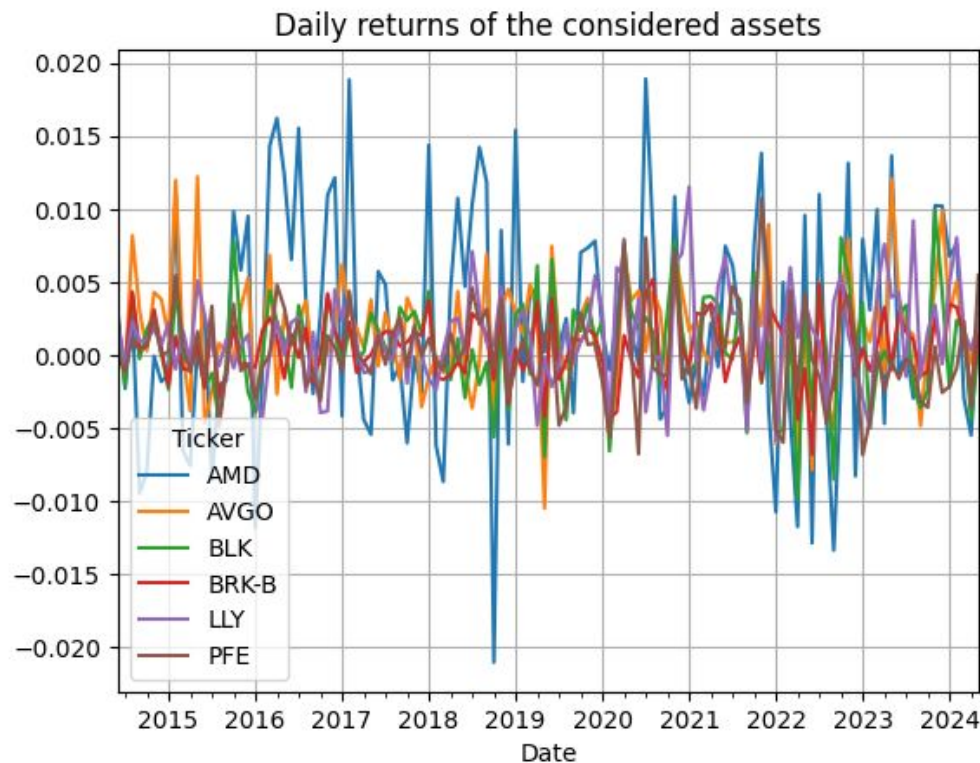
Il rendimento composto annuo o CAGR (Compound Annual Growth Rate) è una misura della crescita annuale media di un investimento su un periodo di tempo, tenendo conto del compounding.

È una metrica utile perché consente di capire la crescita di un investimento come se fosse cresciuto a un tasso costante ogni anno.

CAGR permette di identificare la crescita, fare confronti con altri asset e valutazione di performance.



Rendimenti semplici

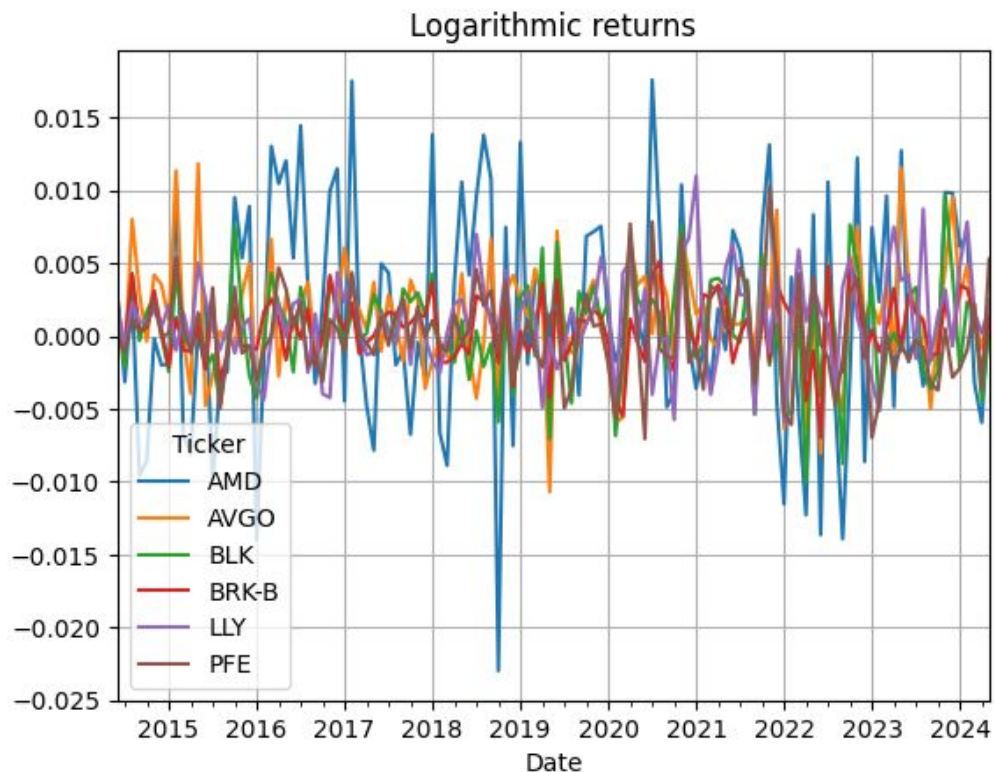


Da questo grafico vediamo che AMD ha una volatilità molto alta.

Gli altri asset hanno una volatilità decisamente inferiore.

In certi periodi si vedono movimenti sincronizzati degli asset suggerendo che reagiscono in modo simile a fattori di mercato comuni.

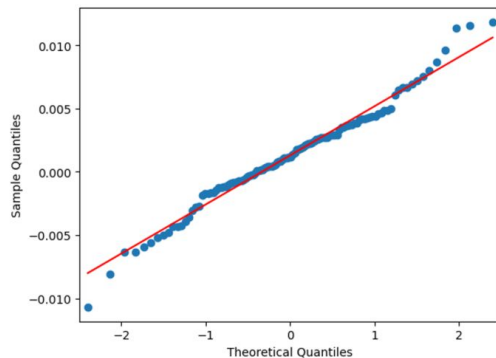
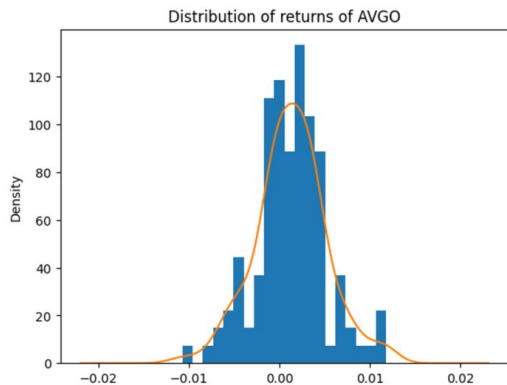
Rendimenti logaritmici



Anche dal grafico dei rendimenti logaritmici vediamo che AMD ha una volatilità molto alta. Gli altri asset hanno una volatilità decisamente inferiore.

I rendimenti logaritmici tendono a mostrare oscillazioni che appaiono più regolari e simmetriche rispetto ai rendimenti semplici. Questo è dovuto al modo in cui i rendimenti logaritmici trattano le percentuali di cambiamento, riducendo l'impatto di variazioni molto alte o molto basse.

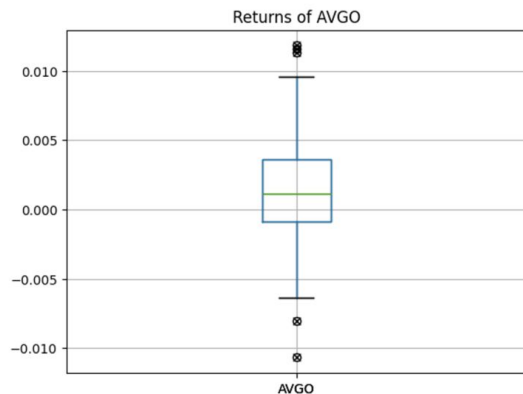
Grafici diagnostici a tre sezioni



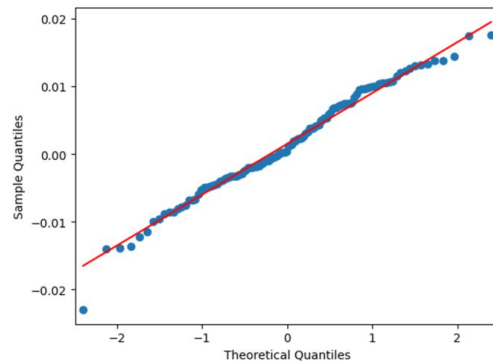
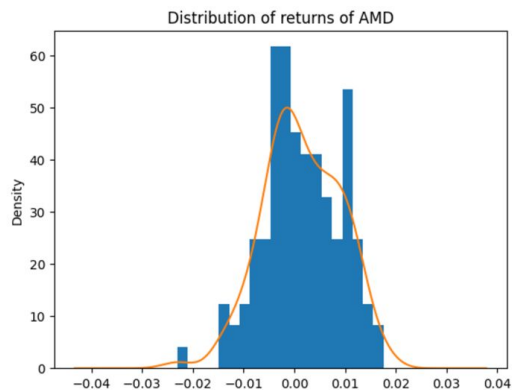
I rendimenti seguono una distribuzione normale dato che la maggior parte dei punti nel QQ plot seguono la linea rossa.

Ci sono alcune deviazioni agli estremi ma non sono eccessive.

La variabilità dei rendimenti è relativamente bassa con alcuni outlier.



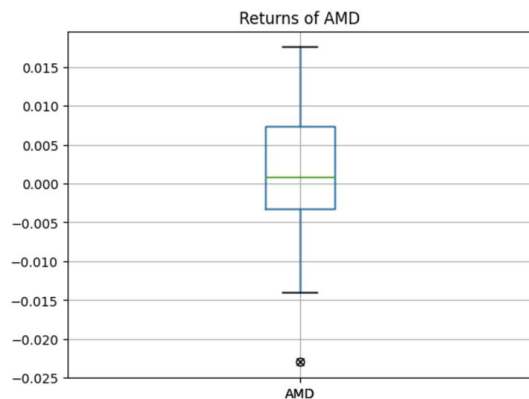
Grafici diagnostici a tre sezioni



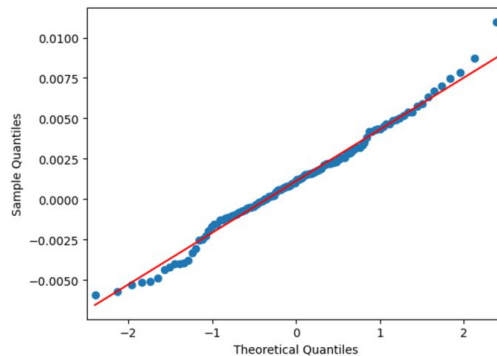
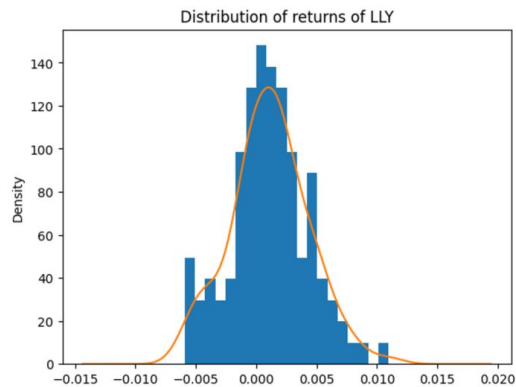
La distribuzione appare simmetrica e centrata attorno a zero.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale.

La maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier sotto la scatola.



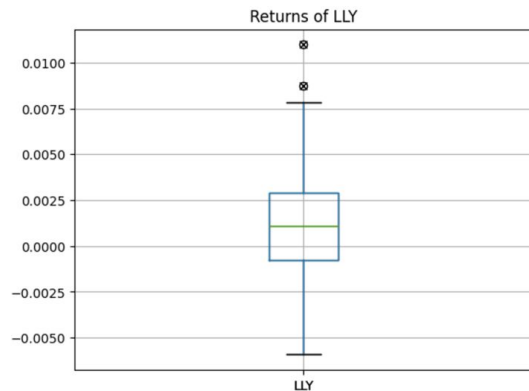
Grafici diagnostici a tre sezioni



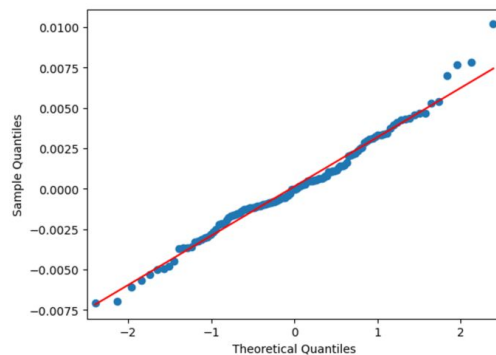
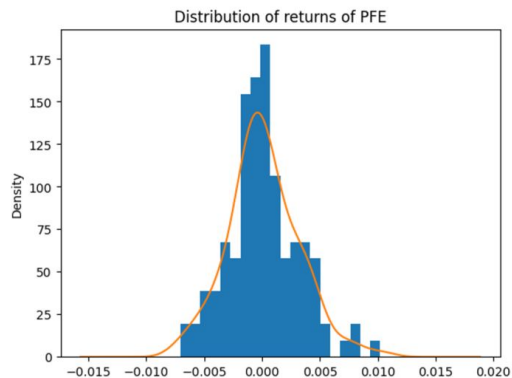
La distribuzione appare simmetrica e centrata attorno a zero.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale.

La maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con pochi outlier sopra la scatola.



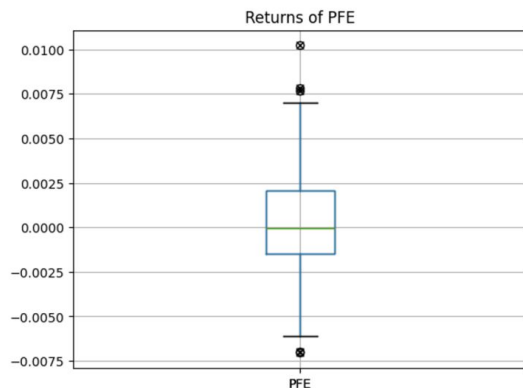
Grafici diagnostici a tre sezioni



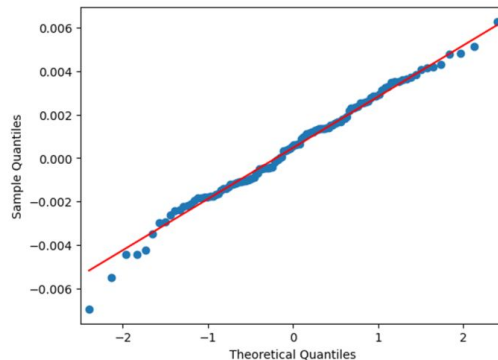
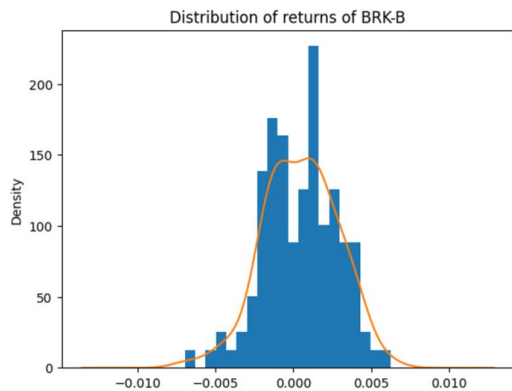
La distribuzione appare simmetrica e centrata attorno a zero.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale.

La maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con diversi outlier.



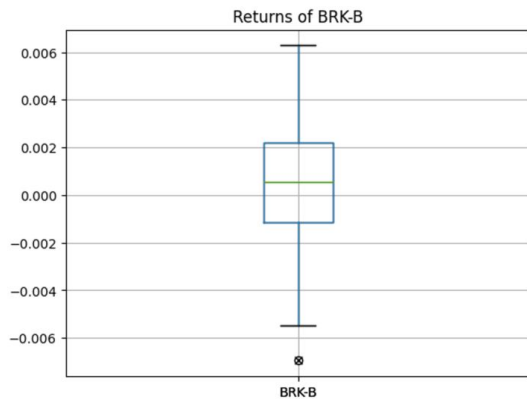
Grafici diagnostici a tre sezioni



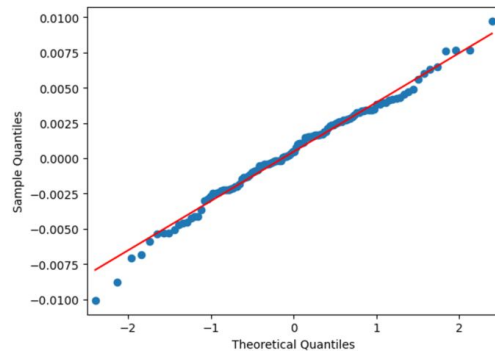
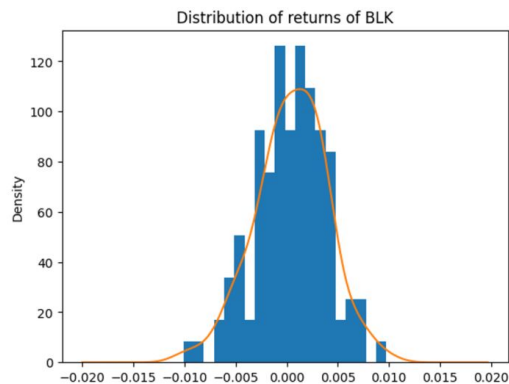
La distribuzione appare abbastanza simmetrica intorno a zero con una tendenza positiva.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale.

La maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con un outlier sotto la scatola.



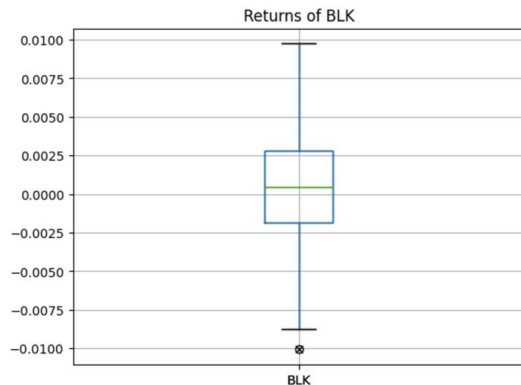
Grafici diagnostici a tre sezioni



La distribuzione appare simmetrica intorno alla zero.

Dal grafico QQ plot vediamo che la maggior parte dei punti si trova vicino alla linea rossa indicando una distribuzione normale.

La maggior parte dei rendimenti rientra in un intervallo ristretto intorno a zero con un outlier sotto la scatola.

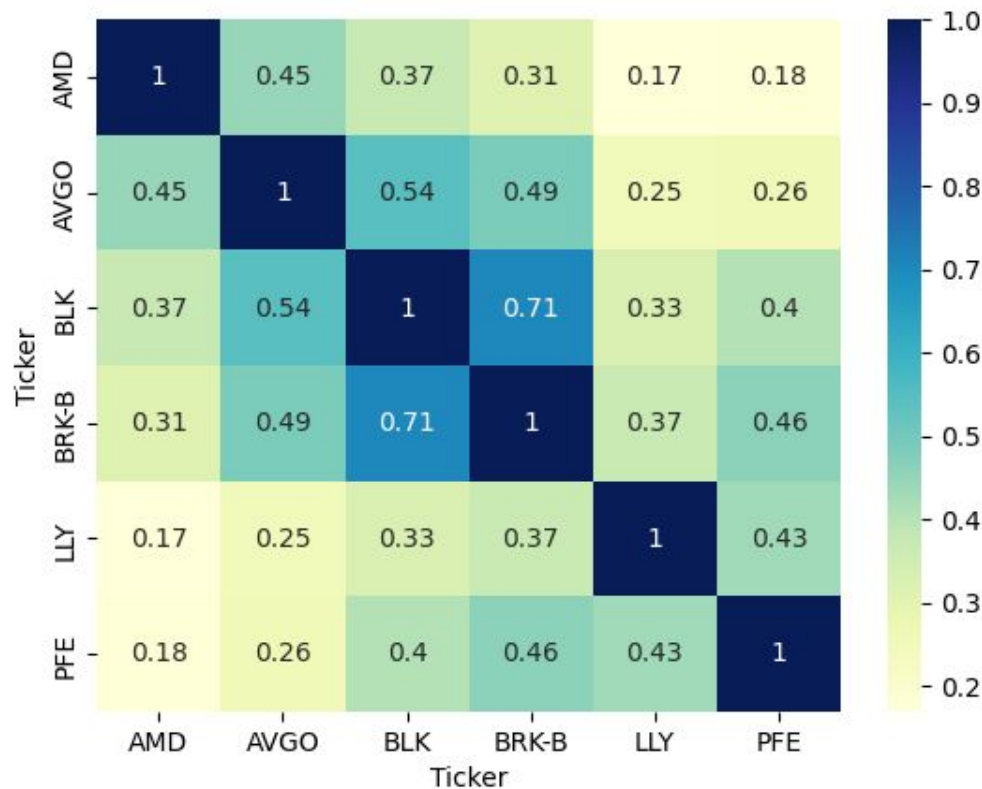


Statistiche descrittive univariate

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	0.7396	0.1503	-0.2482	-0.1606
Curtosi:	9.3448	10.6656	10.7269	5.1215	11.2449	8.7427



Matrice di correlazione

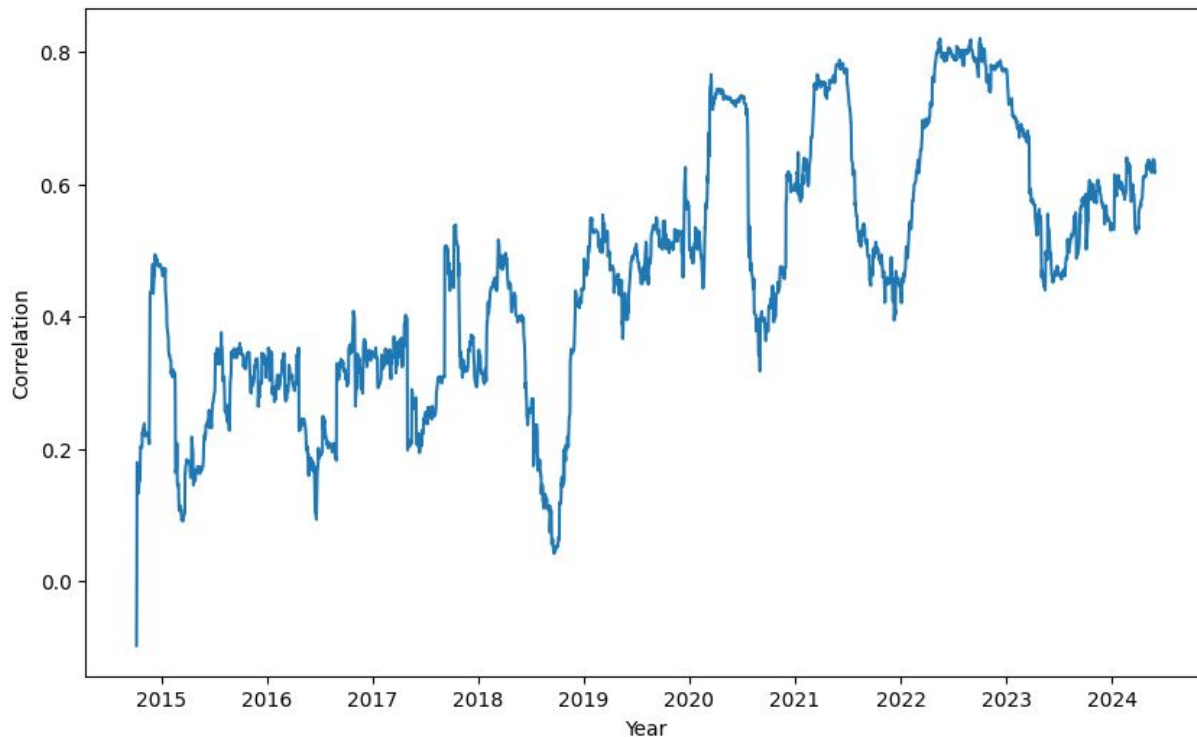


I titoli maggiormente correlati tra loro sono BRK.B e BLK (0.706) che appartengono al settore finance.

I titoli a correlazione minore sono AMD e LLY (0.169) e AMD con PFE (0.178) che appartengono rispettivamente al settore IT e al settore Sanitario.

Correlazione dei titoli nel tempo

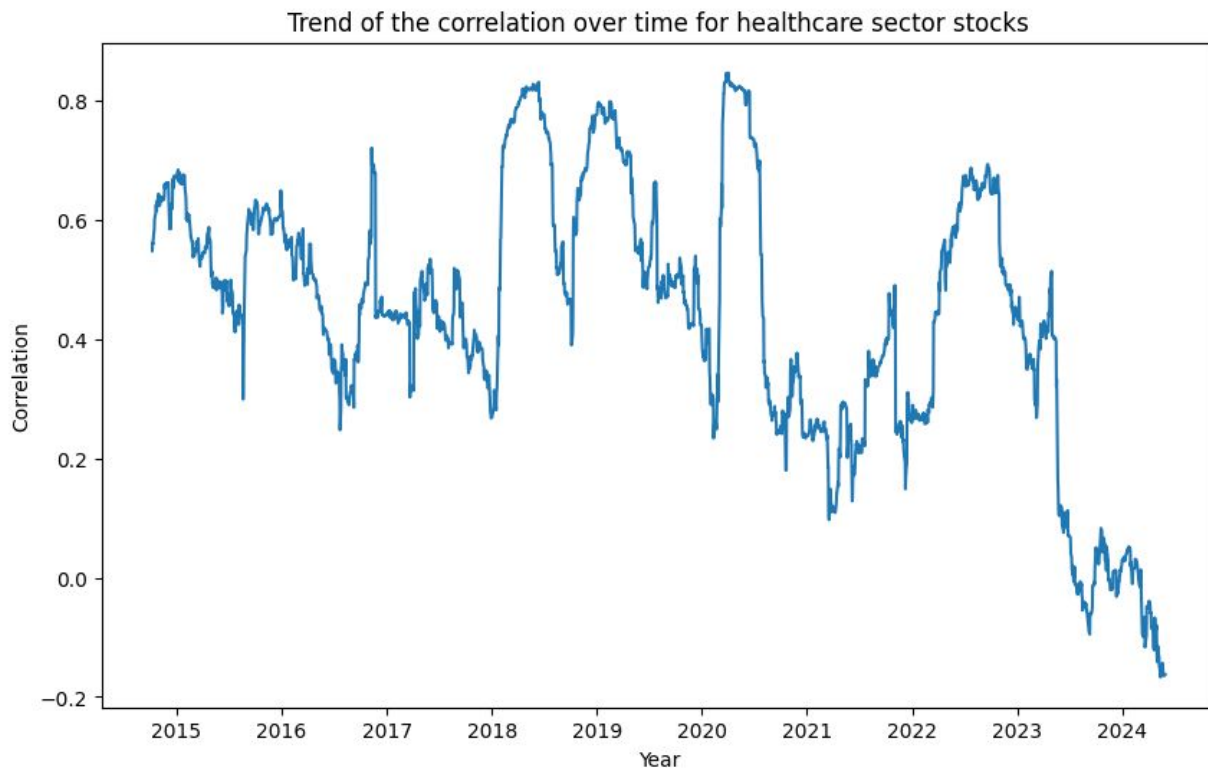
Trend of the correlation over time for IT sector stocks



La correlazione tra i titoli del settore IT ha una crescita graduale.

Dal 2022 la correlazione raggiunge i valori più alti indicando una sincronizzazione nelle performance delle due aziende.

Correlazione dei titoli nel tempo

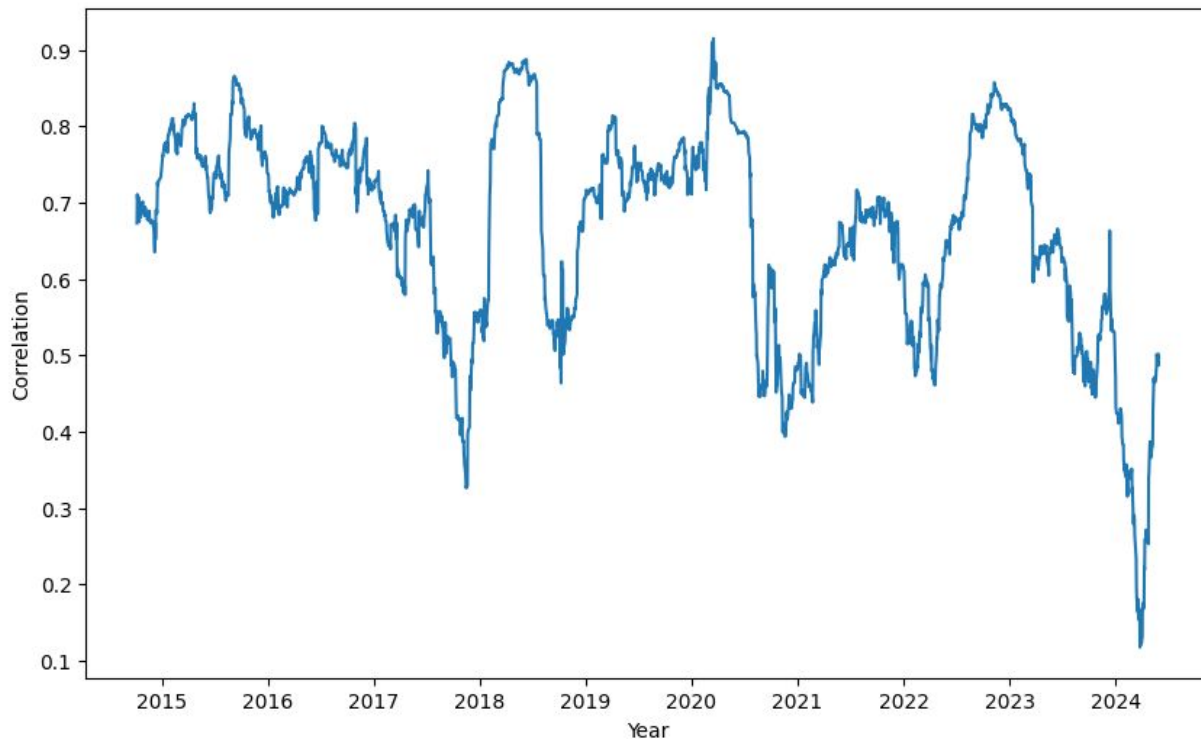


La correlazione tra i titoli del settore sanitario è in decrescita dall'inizio del periodo.

La correlazione ha raggiunto il minimo nel 2024 indicando che i titoli potrebbero muoversi in direzioni opposte.

Correlazione dei titoli nel tempo

Trend of the correlation over time for finance sector stocks



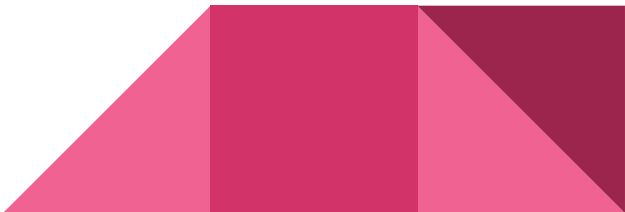
La correlazione tra i titoli del settore finanziario è in lieve decrescita dall'inizio della periodo.

La correlazione ha raggiunto il massimo nel 2020 indicando che i titoli si stavano muovendo in direzioni affini.

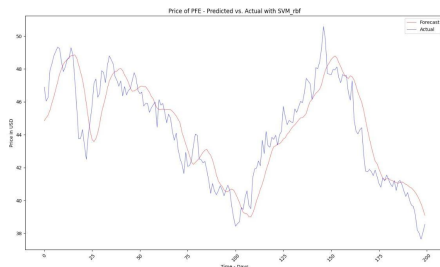
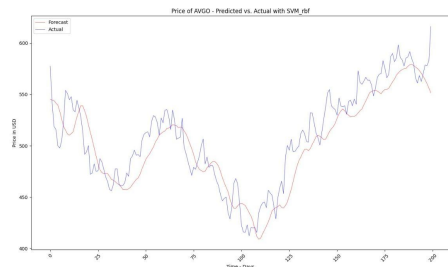
Analisi di previsione

Costruzione modello SVM

Per implementare il modello SVM sono necessari diversi passaggi:

- ❑ Estrazione array feature e target
 - ❑ Divisione dei dati in training, validation e test set
 - ❑ Standardizzazione
 - ❑ Training del modello tramite SVR con kernel RBF
 - ❑ Valutazione del modello
 - ❑ Predizioni e affinamento
 - ❑ Metriche di valutazione (confidence, MAE, MSE, R2)
 - ❑ Renderizzazione
- 

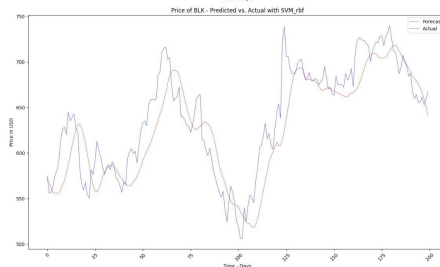
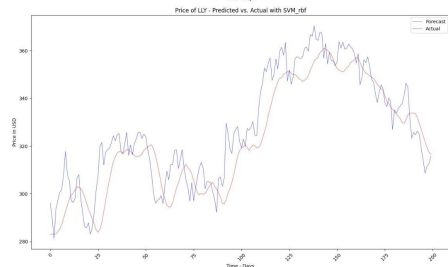
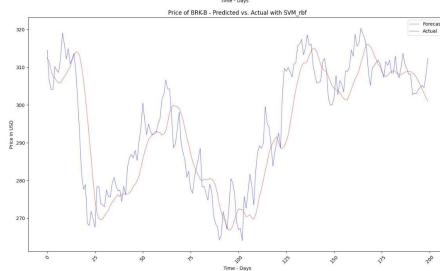
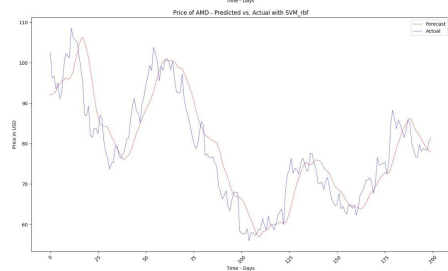
Confronto modello SVM



Il modello con precisione maggiore è quello per l'asset PFE con confidence score pari a 0.945.

L'errore assoluto medio è pari a 1.234 mentre l'errore quadratico medio è 2.752.

Infine l'indice R2 per PFE è pari a 0.933 quindi il modello spiega il 93.3% della variabilità dell'asset.





Strategie di trading e backtesting

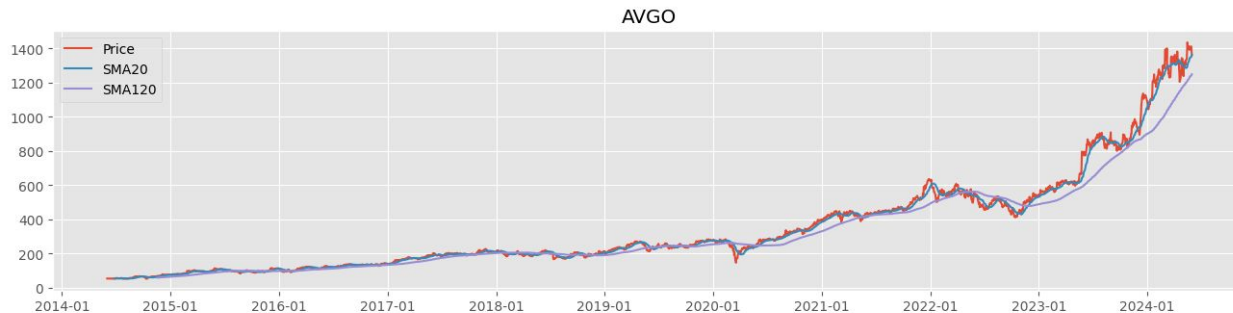
Strategia di trading MACD

Ho utilizzato la strategia di trading Moving Average Converge Divergence come indicatore per identificare i segnali di acquisto o vendita sui mercati finanziari.

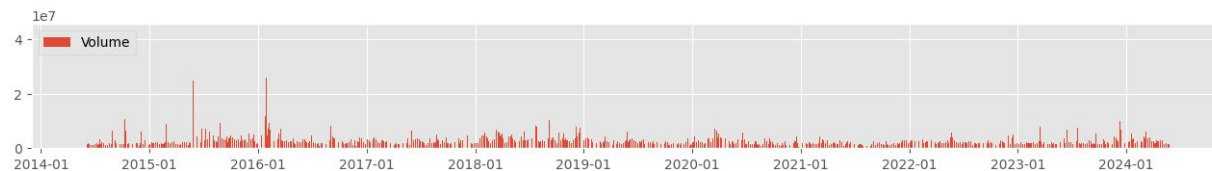
Successivamente ho fatto il confronto con la strategia “Buy and Hold”.



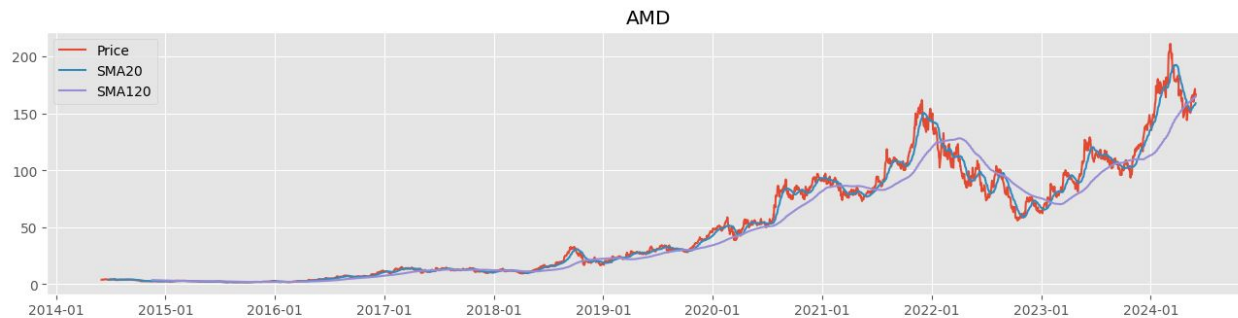
Misura dell'efficienza della strategia



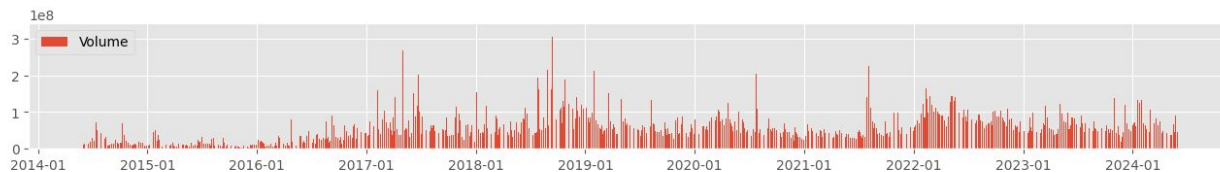
La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.



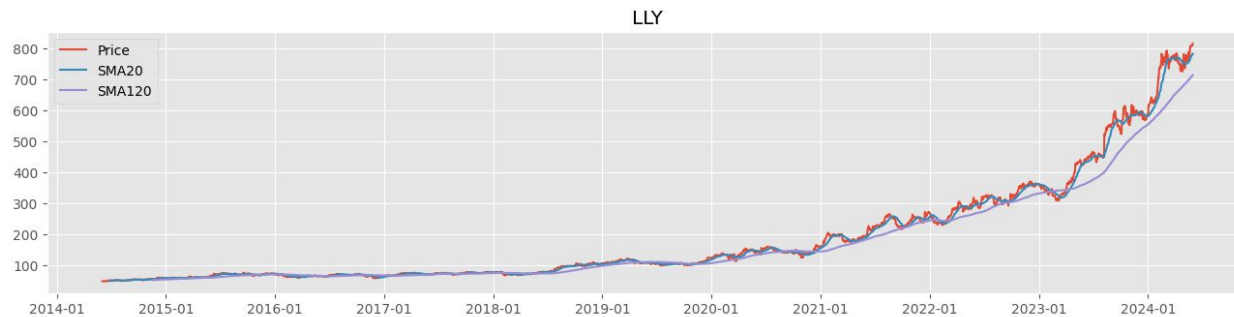
Misura dell'efficienza della strategia



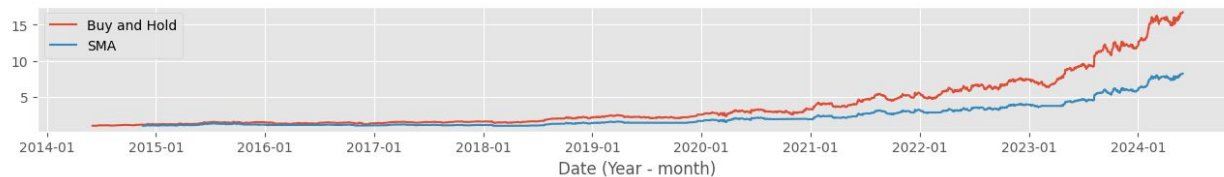
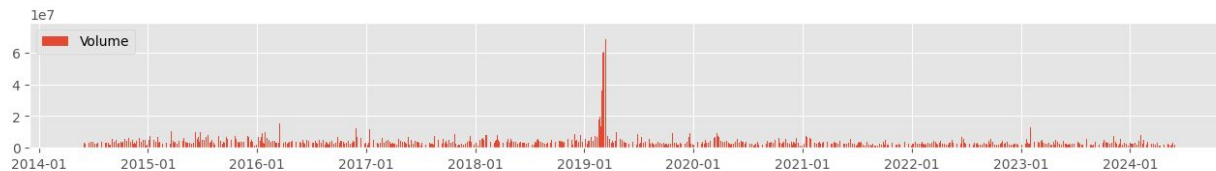
La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.



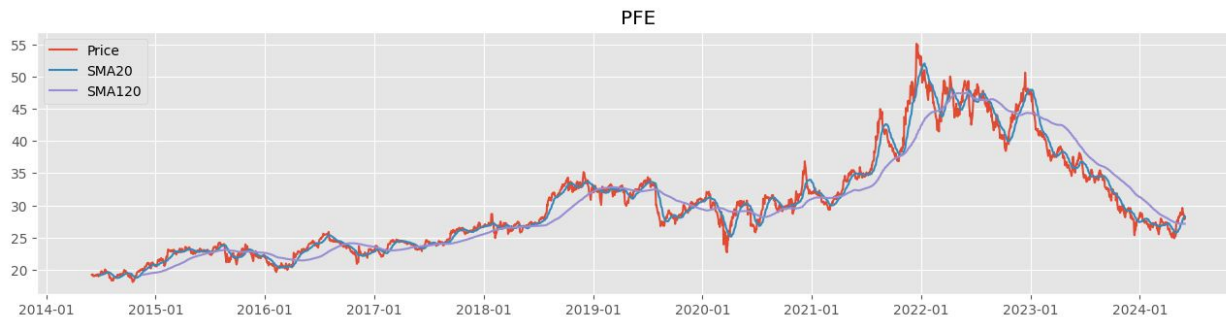
Misura dell'efficienza della strategia



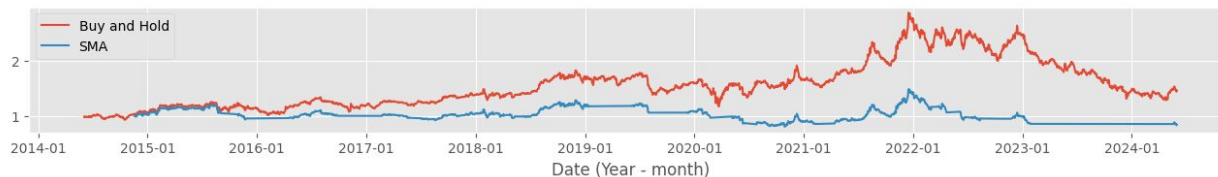
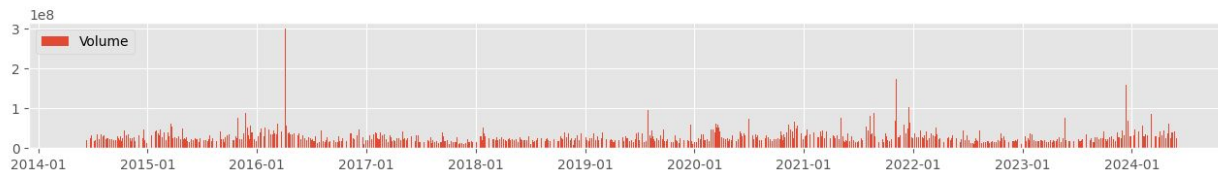
La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.



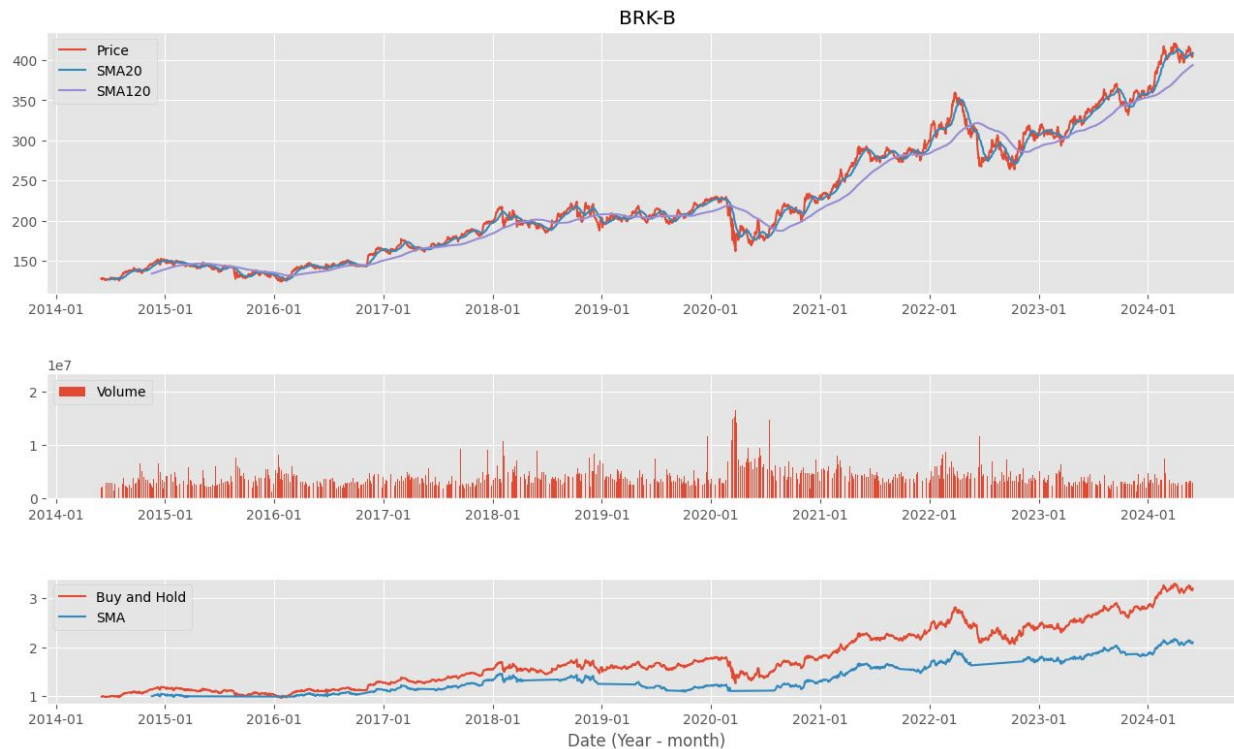
Misura dell'efficienza della strategia



La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.

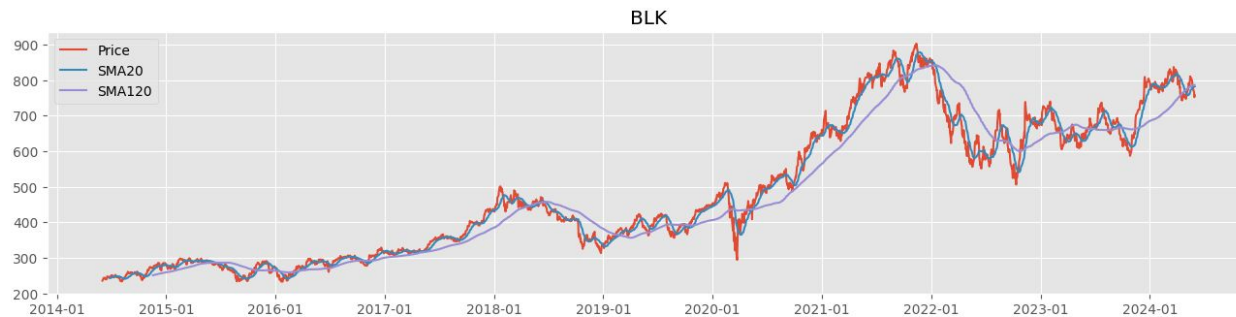


Misura dell'efficienza della strategia

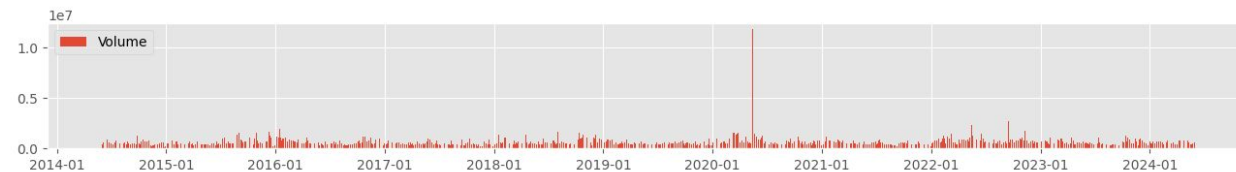


La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.

Misura dell'efficienza della strategia



La strategia Buy and Hold risulta essere superiore alla strategia basata su MACD.





Capital Asset Pricing Model

Beta di ogni titolo

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

Calcolo del rendimento atteso per ogni titolo

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%

Esposizione di rischio ai fattori Fama-French

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
=====
```

Il mercato (mkt) è l'unico fattore significativo ($p < 0.05$) con un coefficiente di 1.0351.

Gli altri due fattori, SMB e HML, non sono significativi.

L'R2 del modello è 0.344, indicando che il 34.4% della variazione nei rendimenti in eccesso di AVGO è spiegato dai fattori del modello.

Esposizione di rischio ai fattori Fama-French

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
=====
```

Il mercato (mkt) è significativo ($p < 0.05$) con un coefficiente di 2.0822.

Anche il fattore HML è significativo con un coefficiente di -0.8083, indicando un'influenza negativa.

Il fattore SMB non è significativo.

L'R2 del modello è 0.363, indicando che il 36.3% della variazione nei rendimenti in eccesso di AMD è spiegato dai fattori del modello.

Esposizione di rischio ai fattori Fama-French

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
=====
```

Il mercato (mkt) con un coefficiente di 0.3181, e il fattore HML con un coefficiente di -0.4501 sono significativi ($p < 0.05$).

Il fattore SMB non è significativo.

L'R2 del modello è 0.104, indicando che solo il 10.4% della variazione nei rendimenti in eccesso di LLY è spiegato dai fattori del modello.

Esposizione di rischio ai fattori Fama-French

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
=====

```

Il mercato (mkt) è significativo ($p < 0.05$) con un coefficiente di 0.6917.

Anche il fattore SMB è significativo con un coefficiente di -0.4709, indicando un'influenza negativa.

Il fattore HML non è significativo.

L'R2 del modello è 0.220, indicando che il 22.0% della variazione nei rendimenti in eccesso di PFE è spiegato dai fattori del modello.

```

=====
              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
=====

```

Esposizione di rischio ai fattori Fama-French

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
=====

```

Tutti e tre i fattori sono significativi ($p < 0.05$).

Il mercato (mkt) ha un coefficiente di 0.9105, SMB ha un coefficiente di -0.5782 e HML ha un coefficiente di 0.3232.

L'R2 del modello è 0.695, indicando che il 69.5% della variazione nei rendimenti in eccesso di BRK.B è spiegato dai fattori del modello.

Esposizione di rischio ai fattori Fama-French

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
=====

```

Il mercato (mkt) è significativo ($p < 0.05$) con un coefficiente di 1.3475.

I fattori SMB e HML non sono significativi.

L'R2 del modello è 0.709, indicando che il 70.9% della variazione nei rendimenti in eccesso di BLK è spiegato dai fattori del modello.

```

=====
              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
=====

```

Costruzione di un portafoglio

Costruzione di un portafoglio

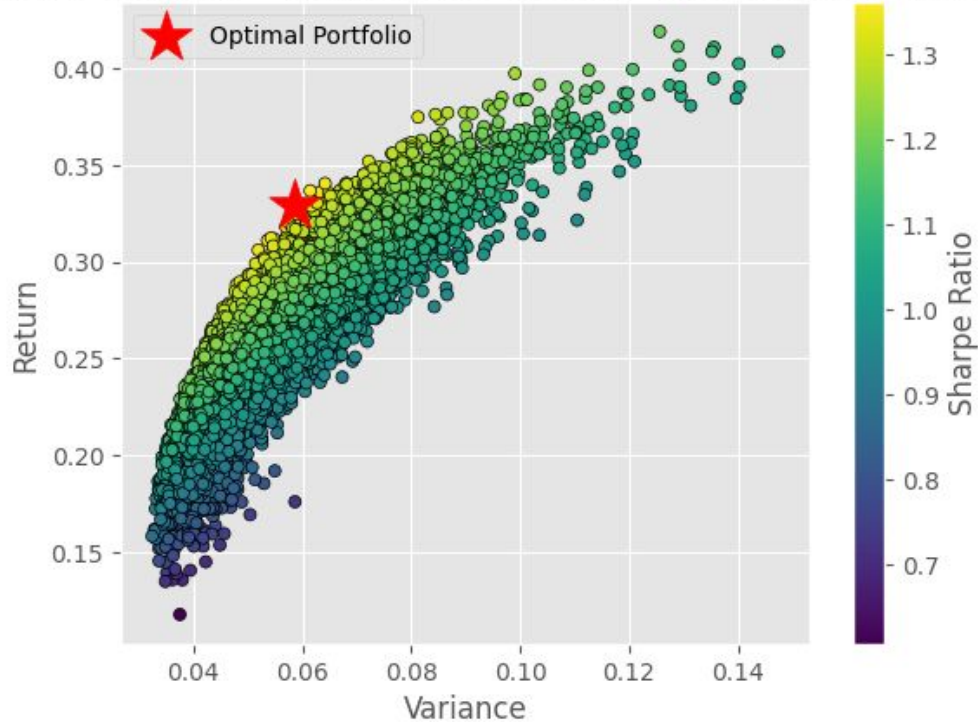
In questa sezione viene costruito un portafoglio in termini di media-varianza sugli ultimi 108 mesi di dati.

Usiamo la frontiera efficiente per visualizzare il portafoglio ottimale basato sui rendimenti passati e sui rendimenti attesi calcolati al punto 5.



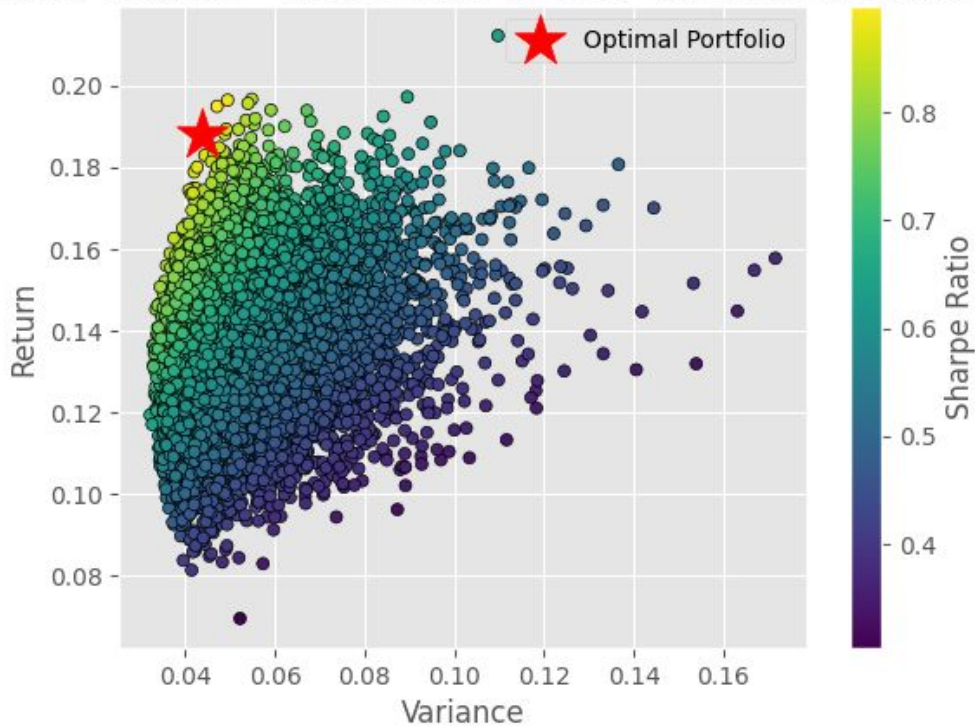
Portafoglio ottimale - rendimenti passati

Optimal Portfolio - Monte Carlo Method - Historical Returns



Portafoglio ottimale - rendimenti attesi

Optimal Portfolio - Monte Carlo Method - Predicted Returns



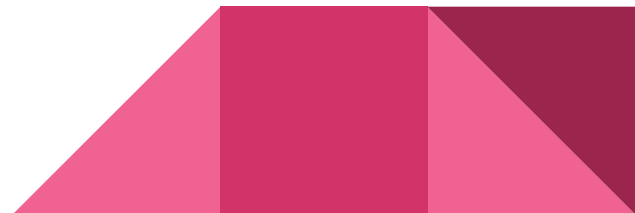
Beta del portafoglio

Il Beta calcolato per il portafoglio ottimale è **1.57**, superiore al Beta standard per SP500 pari a **1.0**.

Di conseguenza il portafoglio è più volatile rispetto al mercato.

Il Beta del portafoglio effettivo (dove tutti gli asset hanno lo stesso peso) è **0.14** che è inferiore rispetto al Beta di SP500 (**1.0**).

Di conseguenza il portafoglio è meno volatile rispetto al mercato.



Conclusioni

Conclusioni

I dati seguono una distribuzione normale e la deviazione standard mostra che la volatilità è contenuta. Gli asset appartenenti allo stesso settore hanno una forte correlazione.

Il modello di previsione ha un livello di approssimazione alto.

La strategia di previsione basata su MACD si è rivelata inferiore rispetto alla classica Buy and Hold.

Infine il portafoglio ottimale è stato identificato.

