



Università degli Studi di Milano Bicocca

Dipartimento di Informatica, Sistemistica e Comunicazione

Corso di laurea in Informatica

# **Progetto di Business Intelligence per i servizi finanziari**

Progetto di:  
Riccardo Mattia  
Matricola 885964

**Anno Accademico 2023-2024**

# 1 Sommario Dei Dati utilizzati

## 1.1 Titoli Utilizzati

per questo progetto di business intelligence sono stati scelti i seguenti titoli per settore:

### 1.1.1 Settore Tecnologico:

#### **NVIDIA (NVDA)**

NVIDIA è una grande azienda che produce schede video, si autodefiniscono “leader mondiali nel computing con intelligenza artificiale”. ho scelto questo titolo dato che recentemente è schizzato a causa della crescente domanda da parte delle grandi aziende tech in velocità di calcolo a causa della corsa alle AI<sup>1</sup>.

#### **INTEL (INTC)**

INTEL è una multinazionale che si occupa di sviluppare processori, ho scelto questo titolo perché a differenza di NVIDIA, INTEL non è schizzato, anzi è sceso e in questo momento potrebbe essere una buona opportunità di investimento dato che è scontato e potrebbe salire per le stesse motivazioni di Nvidia <sup>2</sup>.

---

<sup>1</sup> [Riferimento NVIDIA](#)

<sup>2</sup> [Riferimento INTEL](#)

## 1.1.2 Settore Bancario

### JP Morgan Chase & Co. (JPM)

JP Morgan Chase & Co. è una multinazionale statunitense di servizi finanziari con sede a New York. È una delle banche Big Four statunitensi insieme a Bank of America, Citigroup e Wells Fargo, ed è la più grande banca al mondo con una capitalizzazione di mercato di oltre 420 miliardi di dollari.

La sua posizione offre una panoramica su come le istituzioni finanziarie gestiscono il rischio, l'innovazione tecnologica nei servizi finanziari e la regolamentazione del settore <sup>3</sup>.

### Bank Of America (BAC)

Bank of America Corporation è una banca multinazionale degli Stati Uniti d'America e una società di servizi finanziari. È la seconda più grande istituzione bancaria negli Stati Uniti, dopo JP Morgan Chase. L'ho scelta per le sue possibilità di crescita <sup>4</sup>

## 1.1.3 Settore Bellico

Ho deciso di scegliere due aziende che fanno parte del settore bellico perché a causa delle guerre potrebbero vedere il loro valore salire e potrebbe essere interessante analizzarle, sono state scelte solo per questo motivo, non supporto la guerra in alcuno modo. entrambe le aziende sono state scelte consultando questo sito <sup>5</sup>.

---

<sup>3</sup> [Riferimento JPCM](#)

<sup>4</sup> [Riferimento BAC](#)

<sup>5</sup> [Riferimento settore bellico](#)

## TransDigm Group Inc. (TDG)

TransDigm progetta e produce parti originali di aeromobili per produttori e parti di ricambio per operatori di aerei commerciali e militari. La maggior parte delle sue entrate proviene da fonti dell'aviazione civile. L'azienda sta beneficiando dell'integrazione delle economie globali, che sta stimolando l'aggiunta di flotte di aerei di linea, e del potere di determinazione dei prezzi in quanto unico fornitore di alcuni articoli.

## Huntington Ingalls Industries (HII)

Huntington Ingalls Industries, Inc. (HII) è la più grande azienda di costruzioni navali militari negli Stati Uniti, nonché un fornitore di servizi professionali e partner del governo.

## 1.2 Funzioni Utilizzate

Per scaricare i dati da Yahoo è stata utilizzata la funzione **download()** offerta dalla libreria di python `yfinance`<sup>6</sup>.

A questa funzione viene passata una lista di tickers una data di inizio una data di fine e lui scarica e memorizza tutti i dati relativi ai titoli passati nella lista in un pandas dataframe. Non sono state utilizzate funzioni per fondere i dati dal momento che sono stati scaricati tutti insieme.

## 1.3 Presentazione Dati

si riportano le prime righe del dataframe in Fig. 1.1 e il grafico che raffigura la serie dei prezzi di chiusura di ogni titolo in Fig. 1.2

---

<sup>6</sup> [Documentazione della libreria](#)

Ticker	NVDA	INTC	HII	TDG	JPM	BAC
Date						
2014-06-02	0.450743	20.772072	84.787079	116.768105	41.902775	12.456399
2014-06-03	0.448839	21.076866	84.854218	117.469086	42.092049	12.415581
2014-06-04	0.449314	21.031143	84.887802	117.542862	42.152618	12.415581
2014-06-05	0.451218	21.076866	85.458565	117.426048	42.871807	12.595163
2014-06-06	0.452884	21.465479	86.037781	120.260712	43.129208	12.725771
2014-06-09	0.453360	21.267363	86.658920	121.133850	43.469879	12.929838
2014-06-10	0.455740	21.518820	86.121719	120.162323	43.833263	12.995144
2014-06-11	0.461690	21.282608	84.795471	120.051643	43.356312	12.725771
2014-06-12	0.464546	21.305462	83.217415	118.260452	43.182209	12.587006
2014-06-13	0.465022	22.760883	82.965599	119.036163	43.182209	12.603328

Fig. 1.1 Prime righe del dataframe

dato che TDG ha avuto prezzi di chiusura molto più alti degli altri ho prodotto anche un grafico visibile in Fig. 1.3 senza includere lo stock di TDG per vedere più nel dettaglio gli altri stocks.

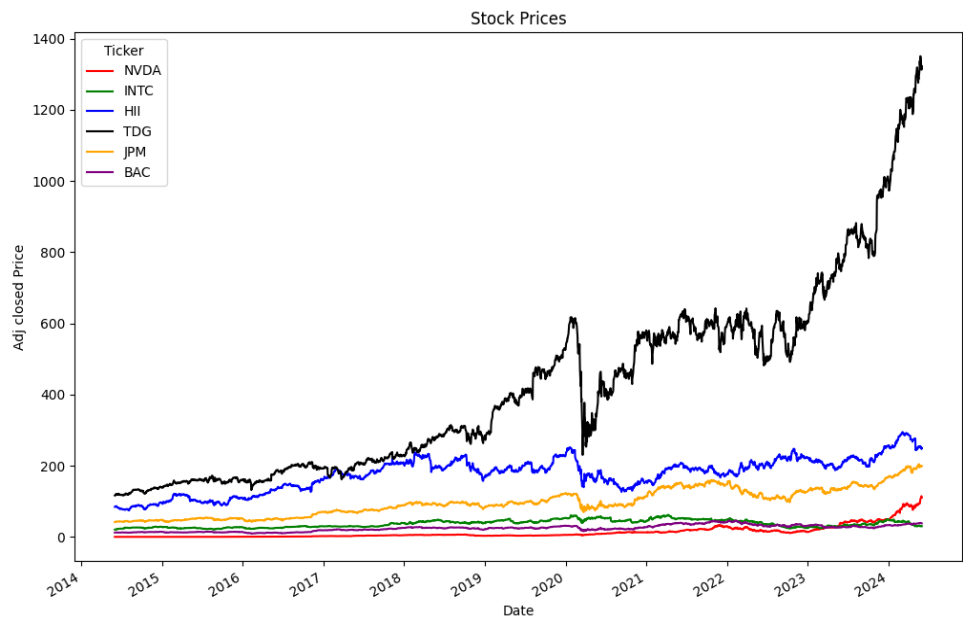


Fig. 1.2 Serie dei prezzi di chiusura

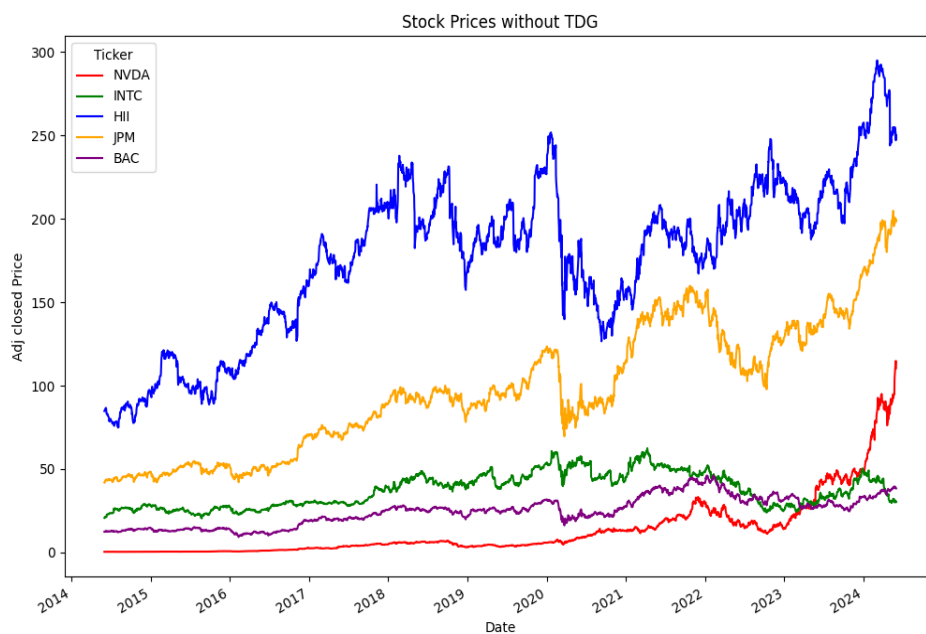
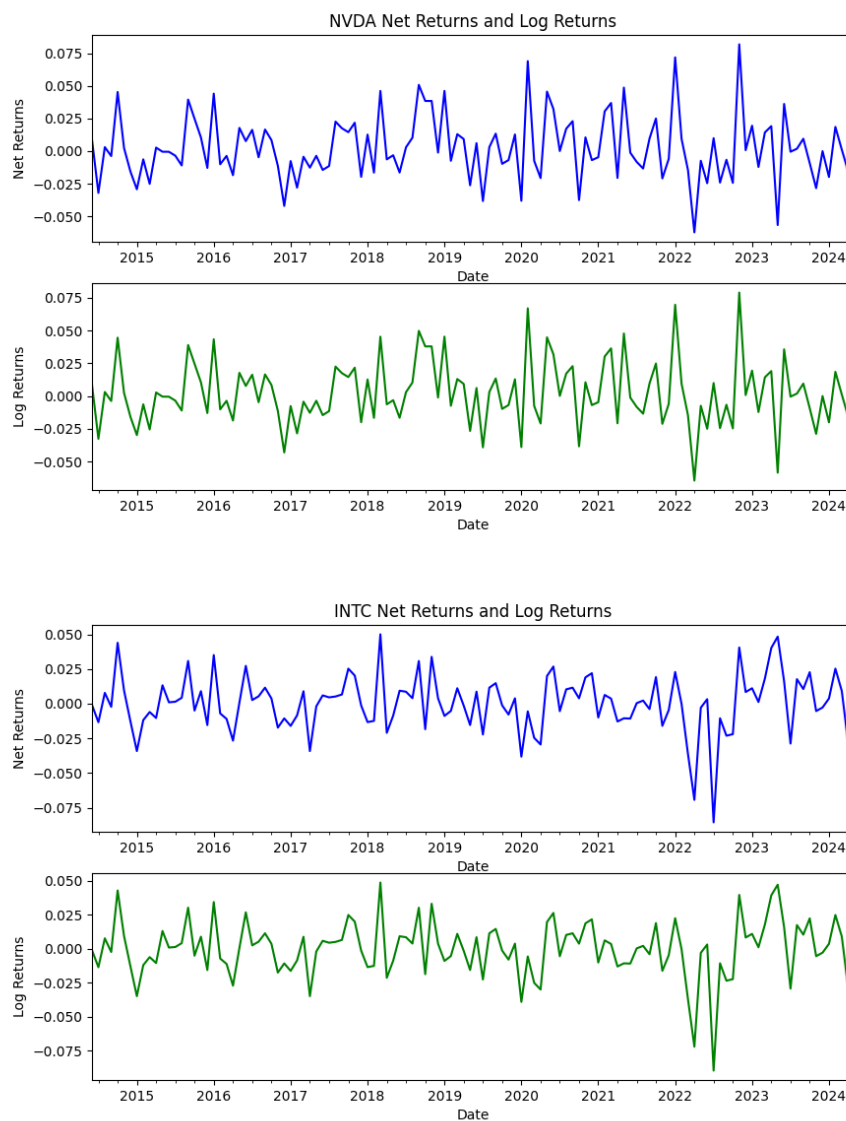


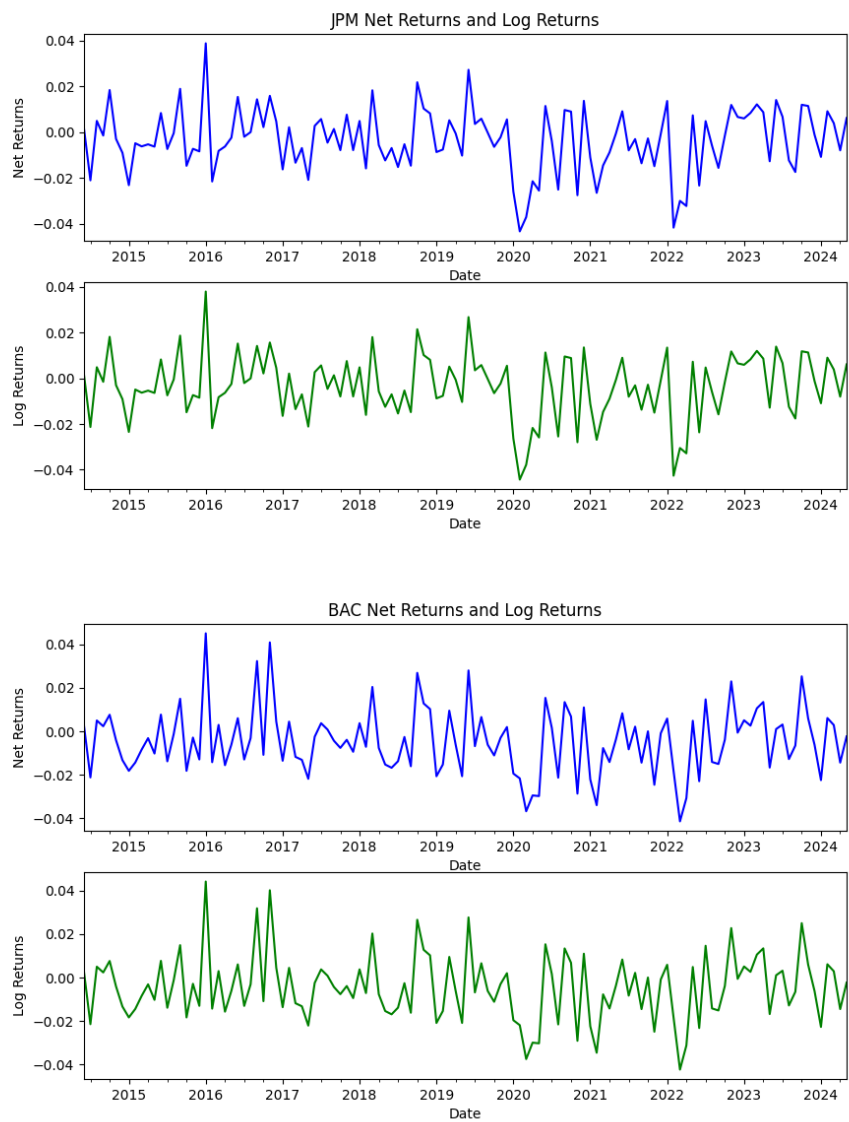
Fig. 1.3 Serie dei prezzi di chiusura senza TDG

## 2 Statistiche Descrittive

### 2.1 Grafici dei Rendimenti e Commenti

#### Ritorni Semplici e Logaritmici



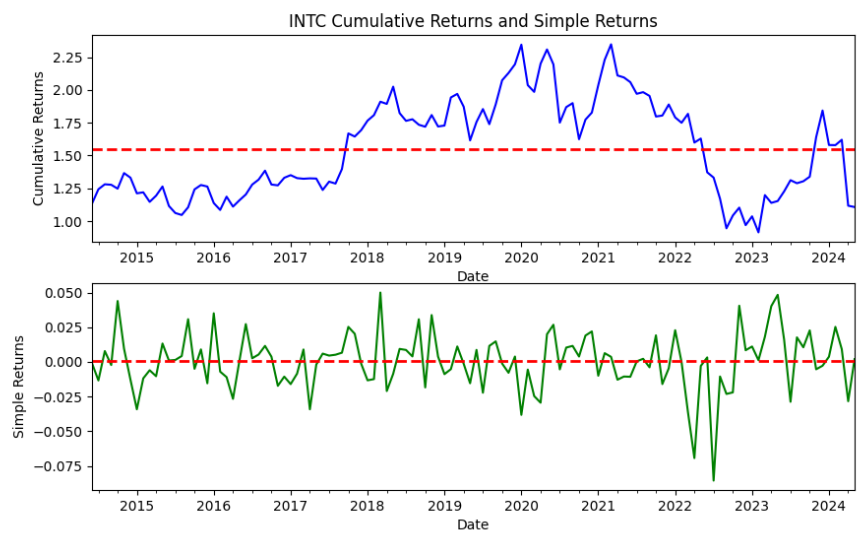
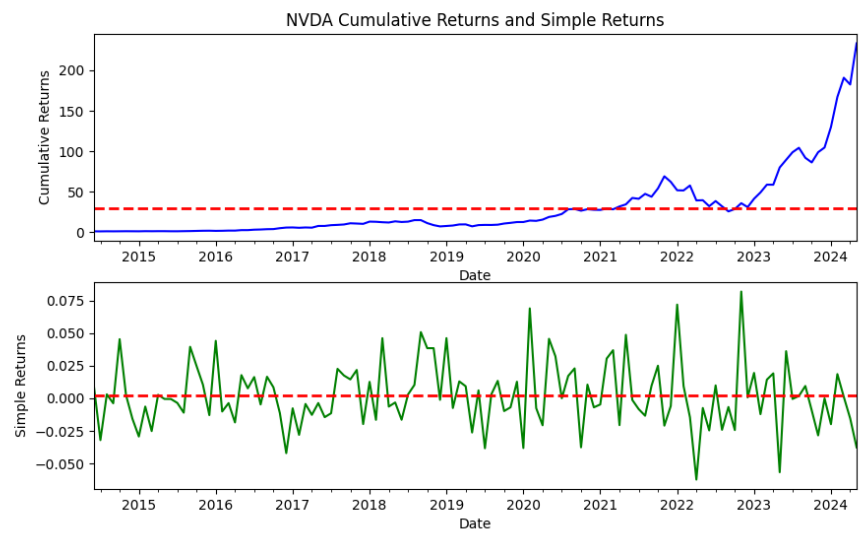


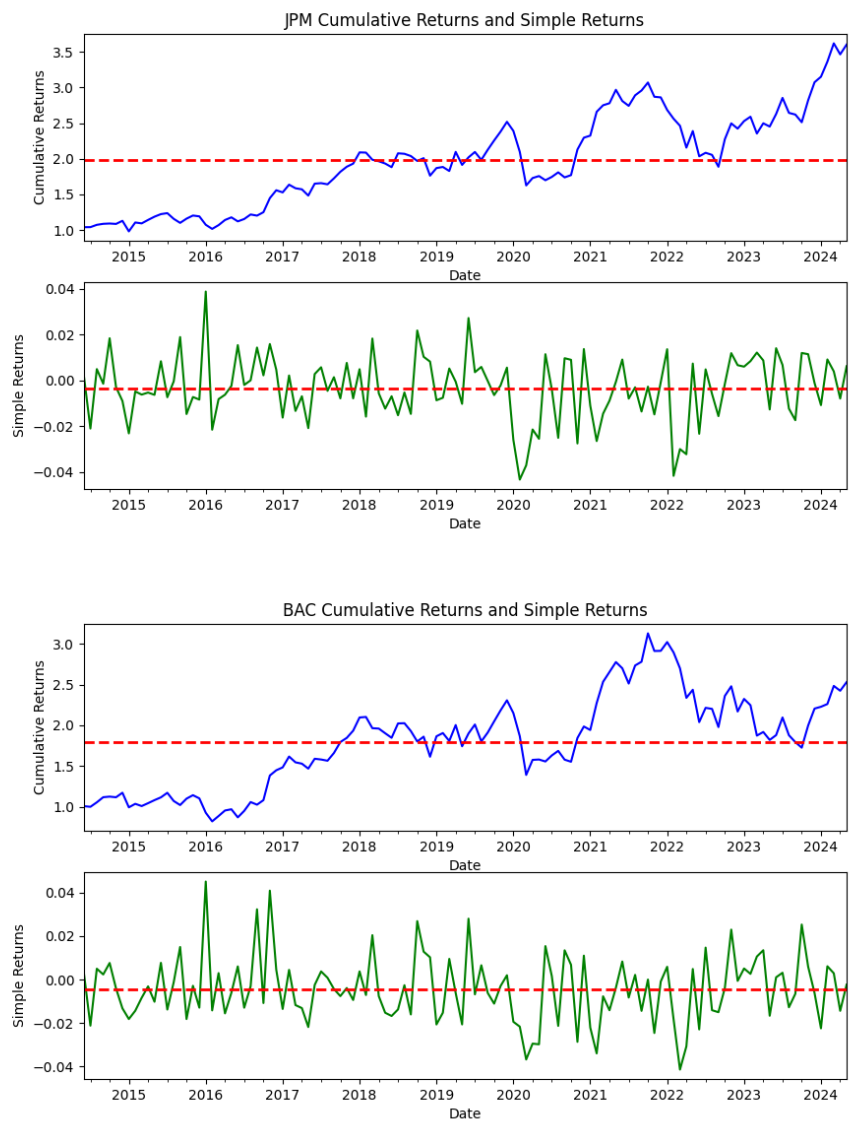


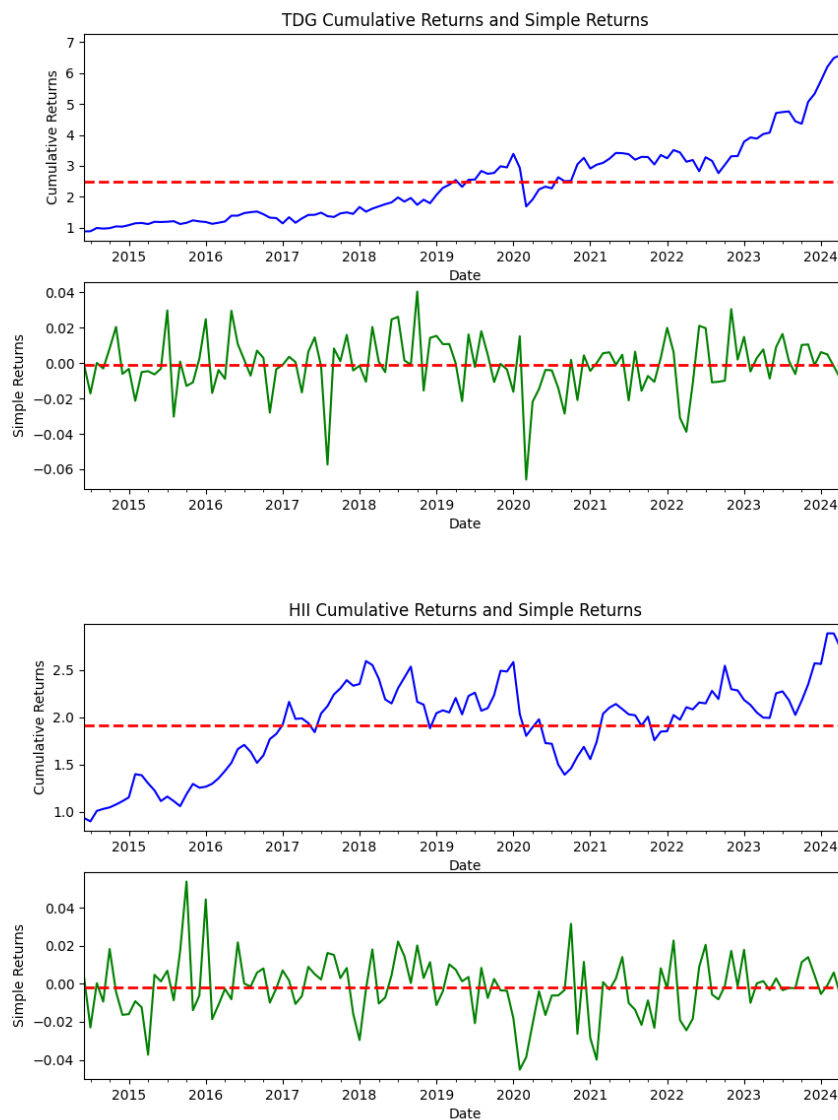


## Rendimenti cumulati e semplici

Nonostante nel testo (sezione 2.b) ci fosse scritto di graficare solo i rendimenti semplici e logaritmici ho voluto inserire anche i grafici a due sezioni dei rendimenti cumulati con i rendimenti semplici, in questi grafici ho anche inserito una linea rossa tratteggiata che indica la media.







## Rendimenti Annualizzati

NVDA\_annualizedReturn 0.732781

INTC\_annualizedReturn 0.038066

HII\_annualizedReturn 0.114060

TDG\_annualizedReturn 0.274381

JPM\_annualizedReturn 0.168631

BAC\_annualizedReturn 0.119051

## Commenti

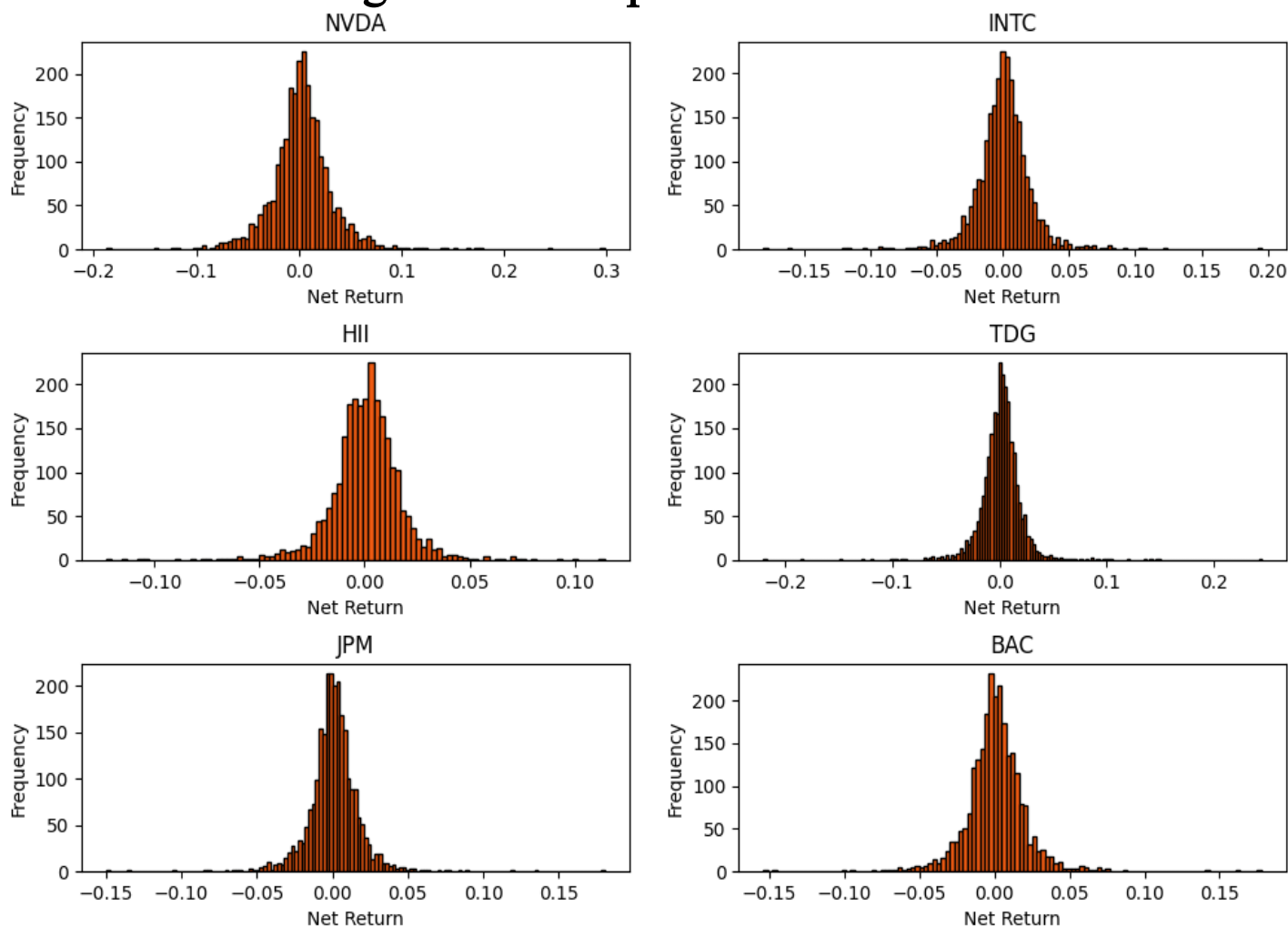
- I. Prima di iniziare con i commenti sui rendimenti, volevo esplicitare che i grafici sono stati prodotti mensilizzando i rendimenti per rendere i grafici più “morbidi”. In generale le serie storiche hanno in comune un **trend crescente**, questo si vede molto bene analizzando i grafici dei ritorni composti e dando un occhio ai rendimenti annualizzati. Un'altra cosa che le serie storiche hanno in comune è una serie di rendimenti negativi in corrispondenza del 2020 questo potrebbe essere stato causato dalla pandemia di **covid19** che ha **impattato negativamente i mercati finanziari**. Oltre all'impatto negativo che la pandemia ha avuto si può notare che dopo la decrescita c'è stata una rapida ricrescita verso l'alto (fine 2021) su tutti gli stock. Inoltre verso fine 2021, inizio 2022 si può notare un'altra decrescita, questo può essere dovuto alla guerra in ucraina che ha portato un periodo di crisi in europa<sup>7</sup>.
- II. C'è sicuramente una correlazione positiva tra società dello stesso settore, a occhio mi pare che i due stocks del settore bancario siano maggiormente correlati rispetto agli altri. Le due società del settore tecnologico non sembrano essere fortemente correlate, soprattutto nell'ultimo periodo. E nemmeno per quanto riguarda il settore bellico noto correlazioni significative.
- III. Di momenti lontani dalla media, oltre al periodo di pandemia possiamo osservare: NVIDIA nell'anno 2023 è salita notevolmente, la possibile causa è che c'è stato il boom delle AI che richiedono molta capacità di calcolo che NVIDIA è in grado

---

<sup>7</sup> [Articolo 2022](#)

di offrire<sup>8</sup>. INTEL nell'anno 2024 è scesa vertiginosamente al contrario di NVIDIA. Ho trovato un articolo<sup>9</sup> che elenca una serie di motivi per cui intel sta rimanendo indietro nel mercato dei semiconduttori

## 2.2 Istogrammi e Dispersione



---

<sup>8</sup> [Articolo NVIDIA](#)

<sup>9</sup> [Articolo INTEL](#)

## Confronto tra i titoli:

### 1) NVDA (NVIDIA Corporation):

- **Osservazione:** La distribuzione dei rendimenti di NVDA è relativamente ampia, con una gamma significativa di rendimenti sia positivi che negativi. Ci sono valori estremi che si estendono oltre  $\pm 0.2$ .
- **Conclusione:** NVDA ha una dispersione alta, indicando una maggiore volatilità.

### 2) INTC (Intel Corporation):

- **Osservazione:** La distribuzione dei rendimenti di INTC è più stretta rispetto a NVDA, con la maggior parte dei rendimenti compresi tra -0.1 e 0.1.
- **Conclusione:** INTC ha una dispersione moderata, indicando una volatilità inferiore rispetto a NVDA.

### 3) JPM (JPMorgan Chase & Co.):

- **Osservazione:** La distribuzione dei rendimenti di JPM è abbastanza stretta, con pochi rendimenti che superano  $\pm 0.1$ .
- **Conclusione:** JPM mostra una dispersione relativamente bassa, indicando una volatilità bassa.

### 4) BAC (Bank of America Corporation):

- **Osservazione:** La distribuzione dei rendimenti di BAC è simile a quella di JPM, con una gamma di rendimenti principalmente tra -0.1 e 0.1.
- **Conclusione:** BAC ha una dispersione bassa, indicando una volatilità ridotta.

### 5) TDG (TransDigm Group Incorporated):

- **Osservazione:** La distribuzione dei rendimenti di TDG è relativamente stretta, con pochi rendimenti che superano  $\pm 0.1$  ma ha outliers che superano lo 0.2.
- **Conclusione:** TDG ha una dispersione moderata, potrebbe avere una volatilità bassa/intermedia.

6) HII (Huntington Ingalls Industries, Inc.):

- **Osservazione:** La distribuzione dei rendimenti è concentrata tra -0.10 e +0.10 con pochi outliers significativi.
- **Conclusione:** HII ha una bassa, questo indica anche una bassa volatilità

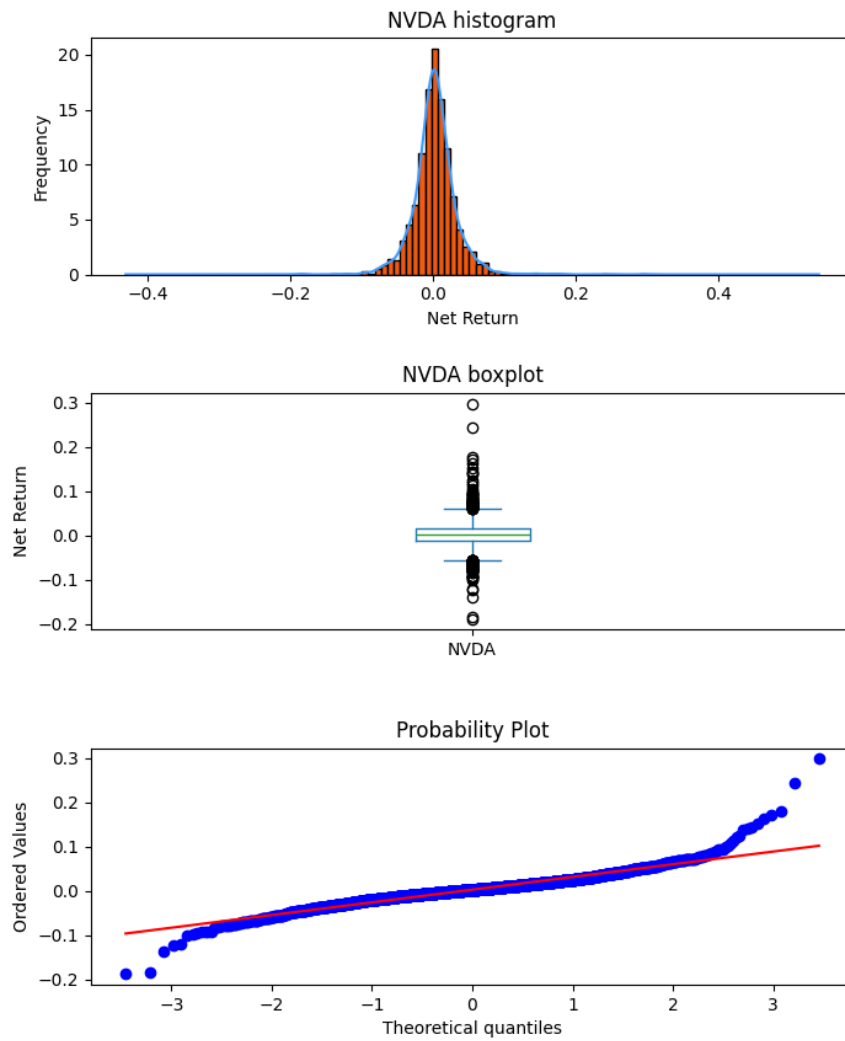
### **Deviazioni Standard dei titoli**

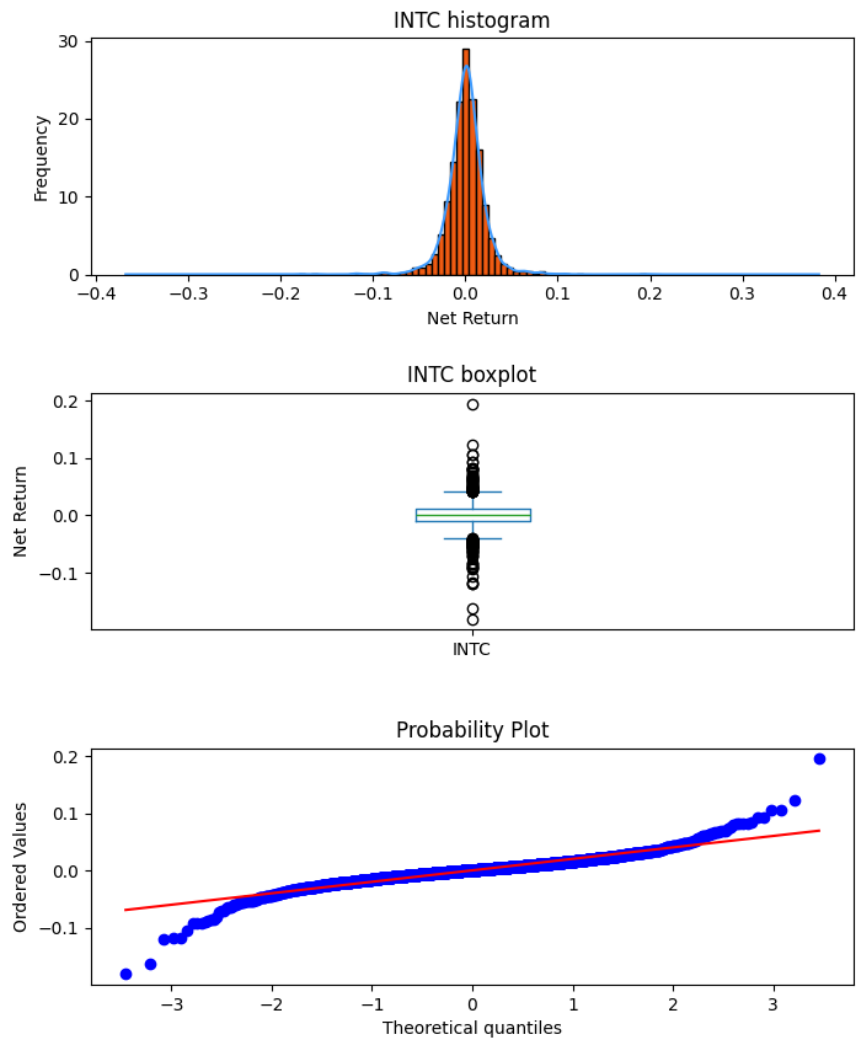
- NVDA (NVIDIA Corporation): 0.029824
- INTC (Intel Corporation): 0.021047
- HII (Huntington Ingalls Industries, Inc.): 0.017221
- TDG (TransDigm Group Incorporated): 0.020850
- JPM (JPMorgan Chase & Co.): 0.016972
- BAC (Bank of America Corporation): 0.019474

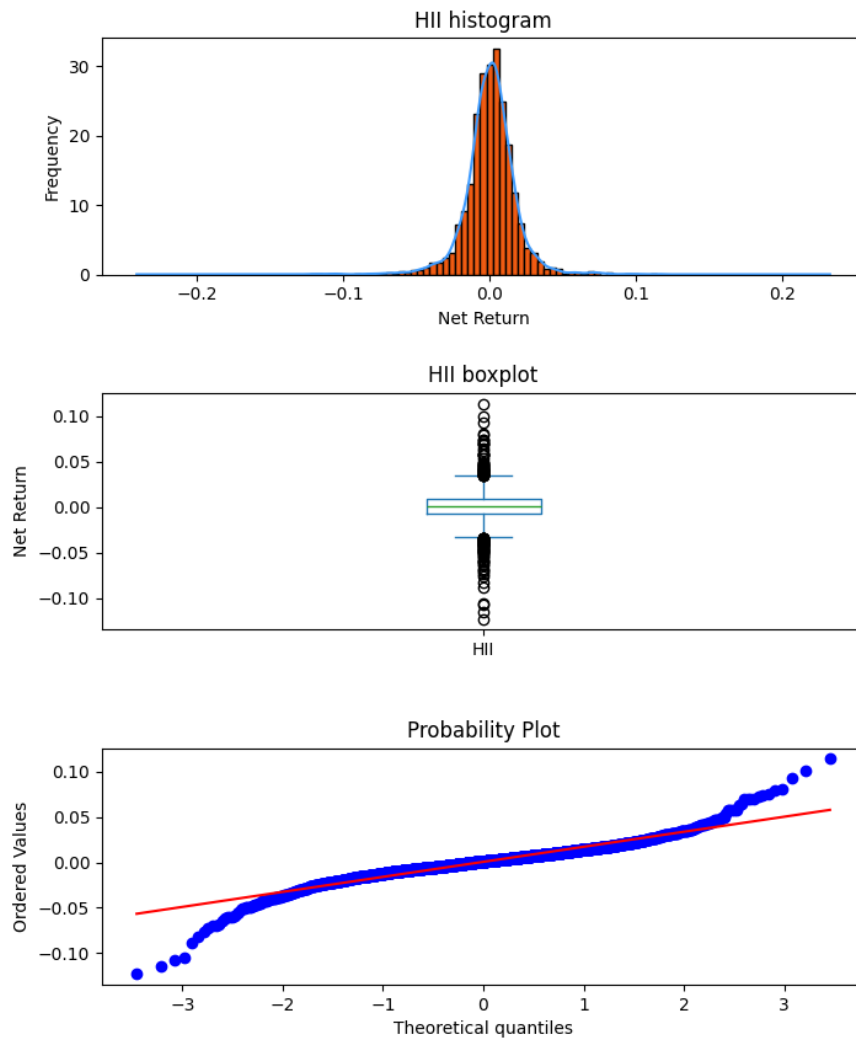
Questi valori confermano l'analisi effettuata sugli istogrammi. In conclusione: **NVDA** è il titolo più volatile. **INTC** e **TDG** hanno una volatilità moderata, riflettendo una stabilità relativa pur essendo soggetti a variazioni significative. **HII**, **BAC** e **JPM** sono i titoli meno volatili, con **JPM** che è il più stabile.

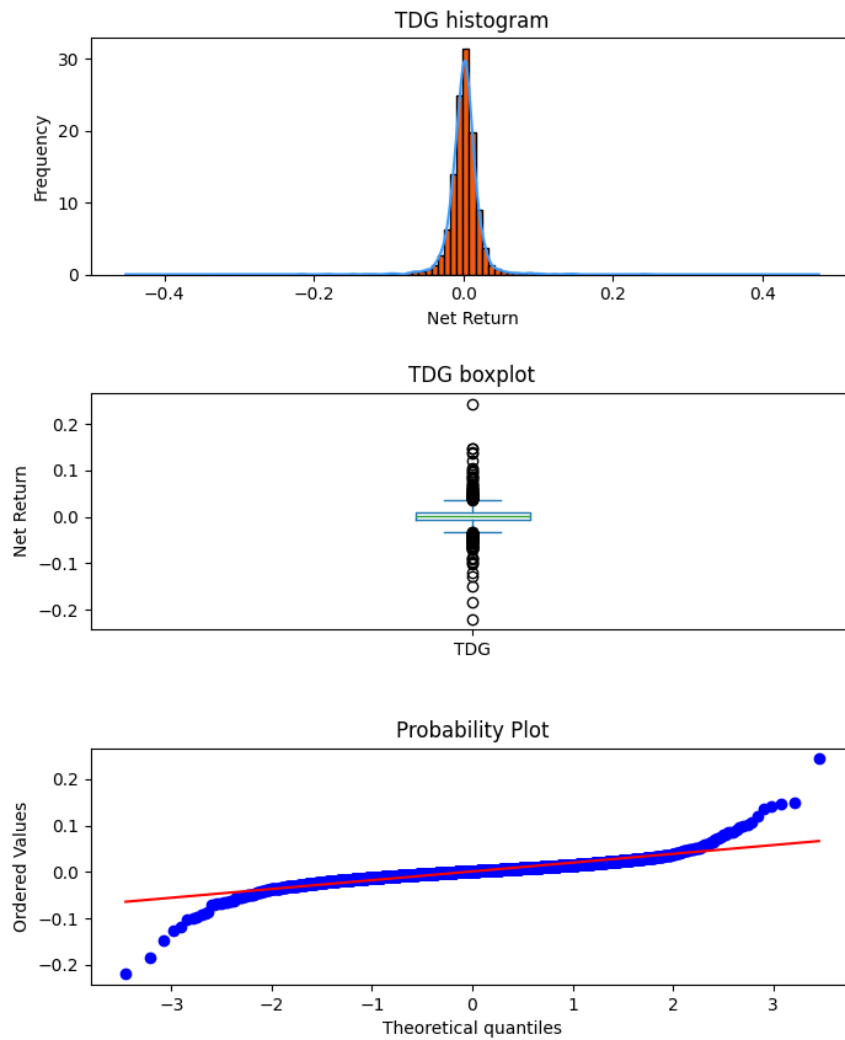


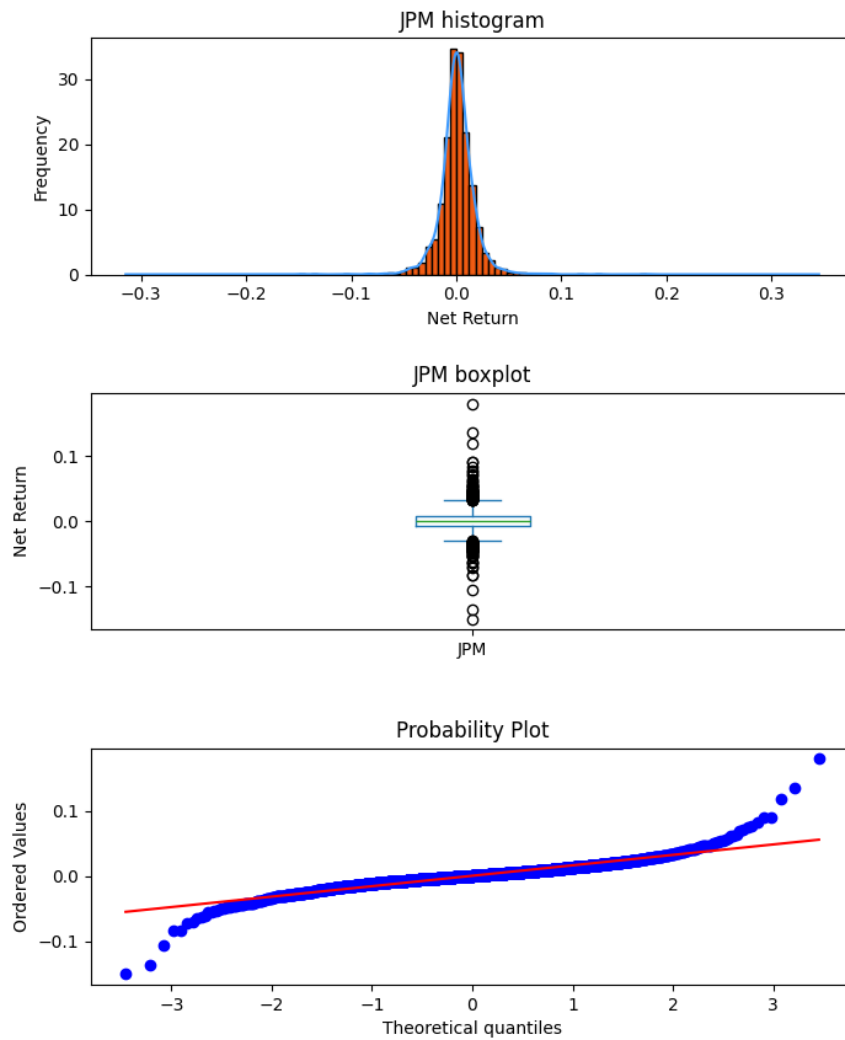
## 2.3 Grafici Diagnostici a Tre Sezioni

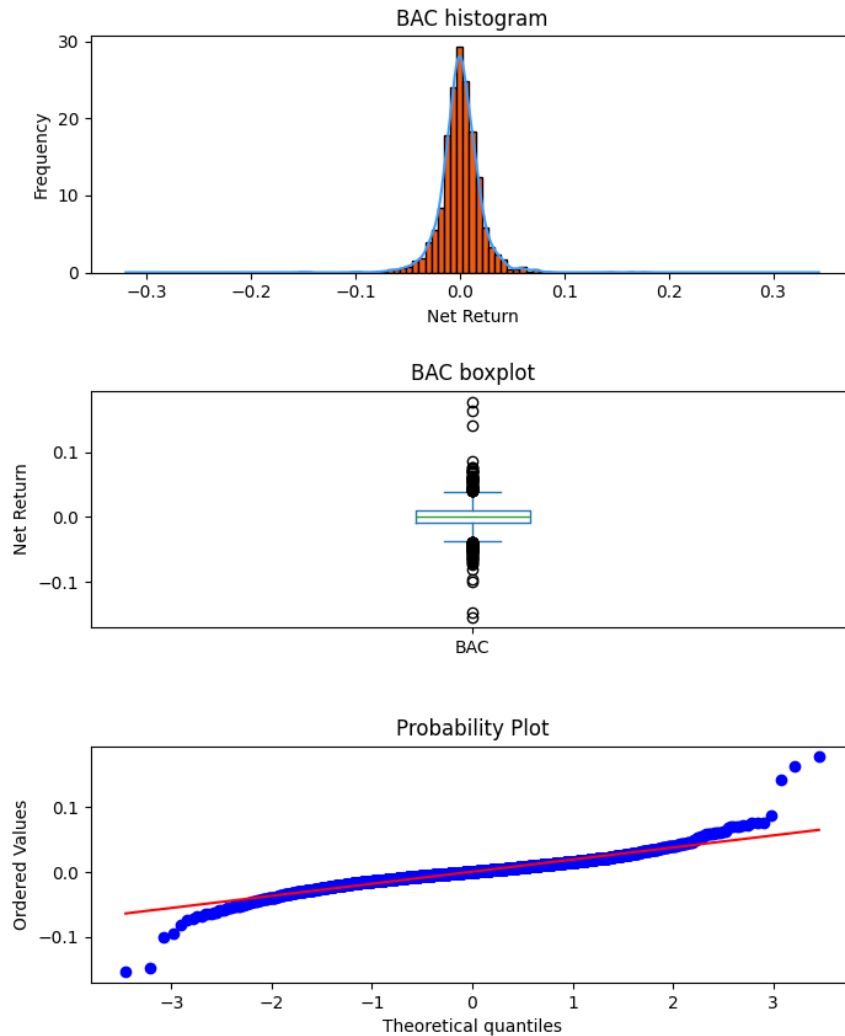












## Commenti

Per tutti i titoli i rendimenti hanno una distribuzione pressoché normale, tuttavia, la distribuzione devia nelle code, questo significa che ci sono rendimenti estremi più frequenti di quanto ci si aspetterebbe in una distribuzione normale.

Per quanto riguarda i boxplot possiamo osservare che ci sono molti **outliers** nei rendimenti dei vari titoli, questo indica che i titoli hanno avuto giornalmente rendimenti più estremi (in positivo o negativo) di quanto ci si aspetterebbe se fossero distribuiti normalmente.

## 2.4 Statistiche Descrittive Univariate

	Media	varianza	Deviazione standard	Asimmetria	Curtosi
NVDA	0.002629	0.000889	0.029824	0.690360	9.271242
INTC	0.000371	0.000443	0.021047	-0.231853	10.310511
HII	0.000579	0.000297	0.017221	-0.281539	6.791653
TDG	0.001183	0.000435	0.020850	0.125962	22.424329
JPM	0.000764	0.000288	0.016972	0.297650	14.113210
BAC	0.000637	0.000379	0.019474	0.282791	10.097505

### Commenti

- i. L'azione con il rendimento più alto è NVIDIA, mentre quella a rendimento più basso è INTEL e HII. Questo dato è osservabile dalla media dei rendimenti.
- ii. Le azioni con la deviazione standard più alta sono NVIDIA e INTEL, confermando la volatilità del settore tecnologico e le analisi effettuate fino ad ora. Le azioni a deviazione standard più basse sono JPM e HII.
- iii. Senza una rappresentazione temporale è difficile dire come si evolvono nel tempo rendimento e volatilità, possiamo aiutarci con i grafici dei rendimenti mostrati prima: La volatilità è sempre aumentata dopo i periodi di crisi spiegati nella sezione Commenti a pag. 13. Per il settore bancario volatilità e rendimenti sono sempre rimasti contenuti confermando la stabilità del settore in

questione. Per il settore bellico abbiamo una volatilità intermedia che lascia comunque spazio a rendimenti positivi. Per ultimo, il settore tecnologico presenta le volatilità più alte rispetto agli altri settori, solo che NVIDIA presenta rendimenti nettamente positivi anche con un'elevata volatilità mentre INTEL no.

- iv. L'azione che ha distribuzione dei rendimenti più vicina alla normale è HII mentre quella più lontana è TDG. Questo dato è osservabile dalla curtosi

## 2.5 Matrici di varianze e Covarianze e Correlazioni

```
NVDA    0.000889
INTC     0.000443
HII      0.000297
TDG      0.000435
JPM      0.000288
BAC      0.000379
dtype: float64
```

Fig. 2.1 Matrice delle varianze

	NVDA	INTC	HII	TDG	JPM	BAC
NVDA	0.000889	0.000306	0.000100	0.000229	0.000175	0.000197
INTC	0.000306	0.000443	0.000108	0.000153	0.000156	0.000172
HII	0.000100	0.000108	0.000297	0.000152	0.000134	0.000151
TDG	0.000229	0.000153	0.000152	0.000435	0.000190	0.000202
JPM	0.000175	0.000156	0.000134	0.000190	0.000288	0.000296
BAC	0.000197	0.000172	0.000151	0.000202	0.000296	0.000379

Fig. 2.2 Matrice delle Covarianze



	NVDA	INTC	HII	TDG	JPM	BAC
NVDA	1.000000	0.488112	0.195491	0.368466	0.345753	0.339757
INTC	0.488112	1.000000	0.298821	0.347971	0.436530	0.420698
HII	0.195491	0.298821	1.000000	0.423178	0.459849	0.449738
TDG	0.368466	0.347971	0.423178	1.000000	0.538235	0.497808
JPM	0.345753	0.436530	0.459849	0.538235	1.000000	0.895274
BAC	0.339757	0.420698	0.449738	0.497808	0.895274	1.000000

Fig. 2.3 Matrice delle Correlazioni

## Commenti

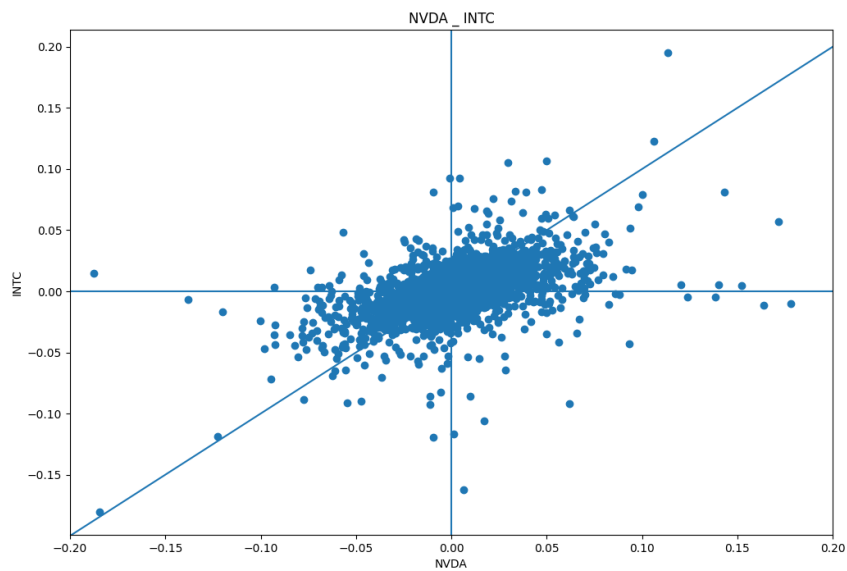
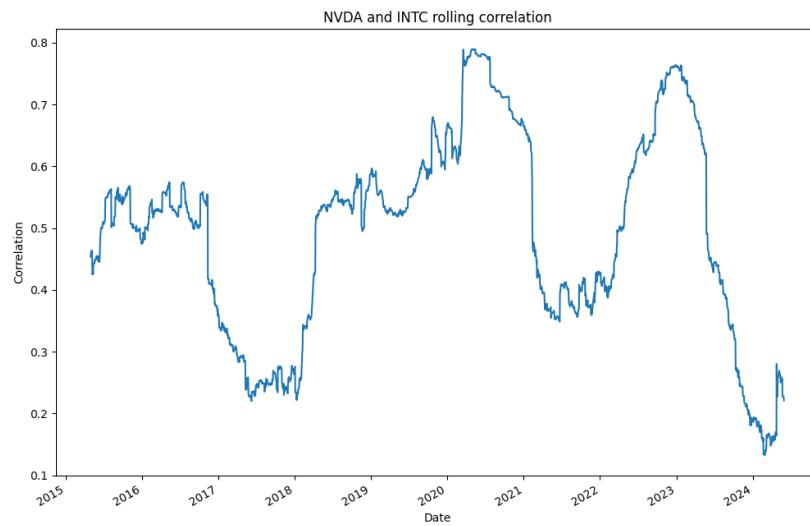
I titoli a **correlazione più alta** sono JPM e BAC, questo risultato è atteso perché entrambe le società operano nel **settore bancario**.

I titoli a **correlazione più bassa** sono NVDA e HII, anche in questo caso non è un risultato imprevedibile dal momento che NVDA opera nel settore tecnologico, mentre HII opera nel settore bellico e della difesa, l'andamento dei rendimenti di questi due settori è influenzato da fattori diversi.

Altre correlazioni degne di nota sono tra TDG e JPM con una correlazione 0,538 e tra TDG e BAC con indice di correlazione a 0,497 che pur non essendo forti correlazioni sono comunque alte tenendo in considerazione che si tratta di due settori diversi.

## 2.6 Correlazioni nel tempo e Scatter plots

### 2.6.1 Settore Tecnologico



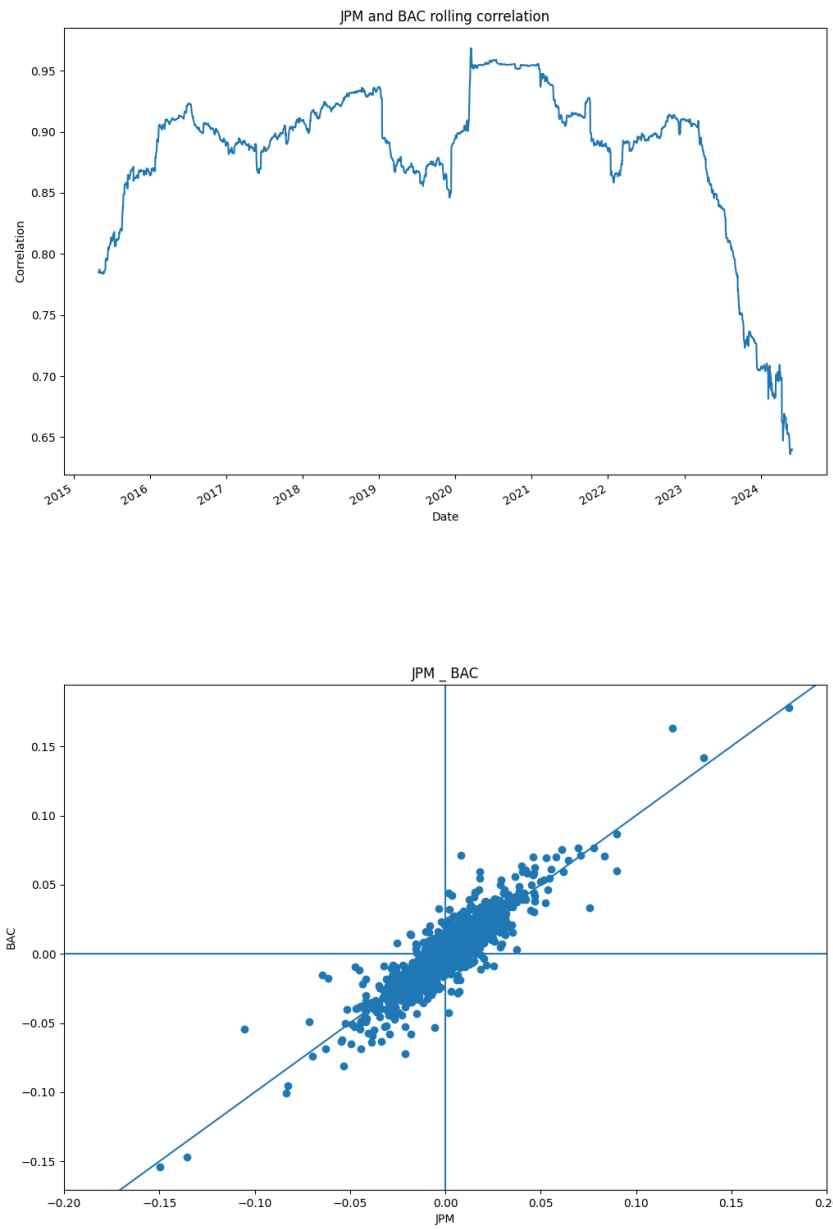
## **Relazioni e Andamento nel tempo**

Il grafico della correlazione nel tempo tra NVDA e INTC mostra variazioni significative nella correlazione tra le due azioni. Inizialmente, la correlazione è relativamente alta, con valori intorno a 0.5 e 0.6, ma scende a valori inferiori a 0.3 nel 2017. Successivamente, la correlazione risale, raggiungendo picchi intorno a 0.77 nel 2020, prima di scendere nuovamente. Questo indica che la relazione tra i rendimenti di NVDA e INTC è variabile e soggetta a cambiamenti nel tempo.

Per quanto riguarda i rendimenti di NVDA e INTC, possiamo osservare dallo scatter plot che la maggior parte dei punti si trova vicino all'origine, indicando che piccoli rendimenti positivi o negativi sono comuni per entrambe le azioni. Tuttavia, ci sono alcuni punti di dispersione maggiore che indicano rendimenti estremi. La correlazione tra i due titoli sembra essere più forte per i rendimenti più piccoli, mentre i rendimenti estremi mostrano una maggiore variabilità.

La dispersione dei punti nello scatter plot conferma una relazione lineare positiva tra i rendimenti di NVDA e INTC, ma non è perfetta. La maggior parte dei punti si distribuisce lungo una linea diagonale positiva, suggerendo che quando i rendimenti di NVDA sono positivi, anche quelli di INTC tendono a essere positivi, e viceversa. Tuttavia, ci sono punti che si discostano da questa linea, indicando che la relazione lineare non è perfetta e ci sono periodi in cui i rendimenti delle due azioni non si muovono insieme.

## 2.6.2 Settore Bancario



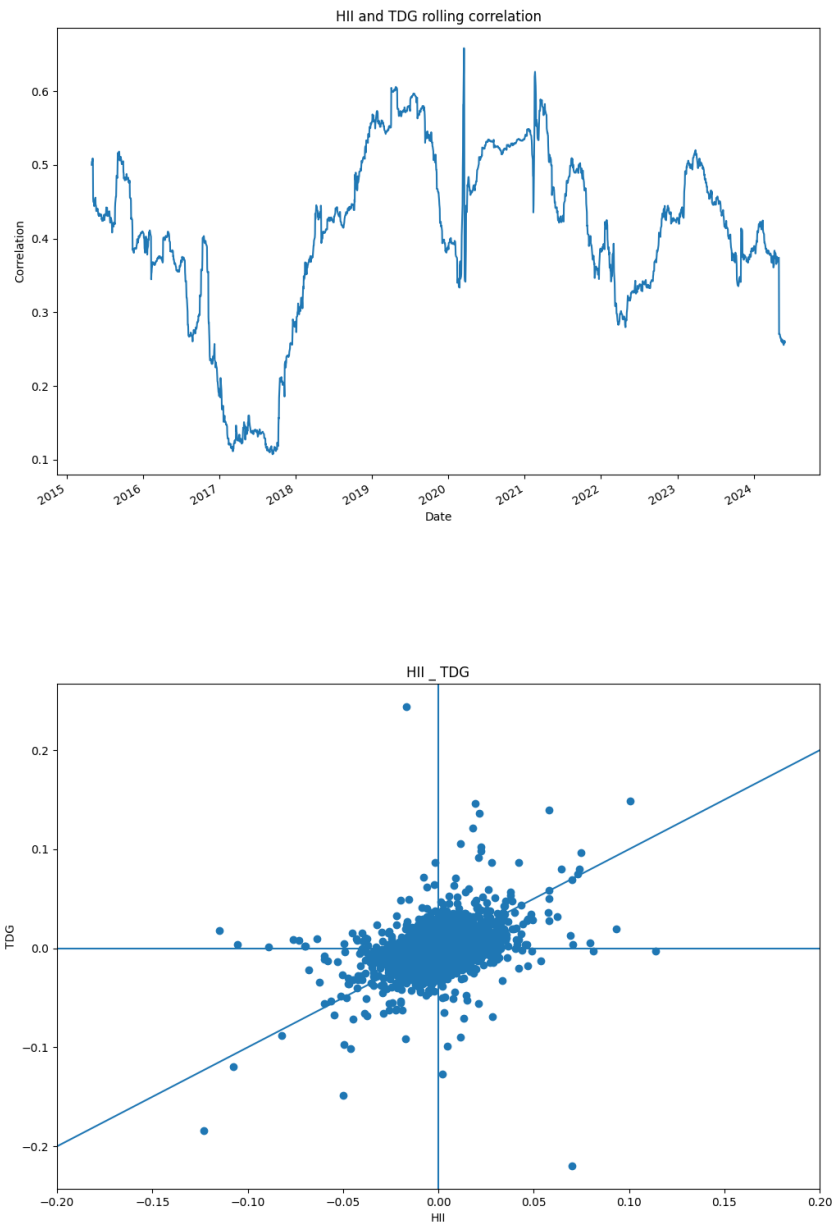
## **Relazioni e Andamento nel tempo**

Dal grafico che ci mostra la correlazione nel tempo di JPM e BAC possiamo notare che verso la fine del 2016 la correlazione supera il valore di 0.85 non andando mai al di sotto di questo valore fino a metà 2023. In questo arco di tempo la correlazione è sempre rimasta tendenzialmente alta arrivando a toccare il picco a 0.96 nel 2002. nell'ultimo periodo la correlazione è scesa parecchio ma rimane comunque superiore allo 0.60.

uno scatter plot dei rendimenti giornalieri di JPM e BAC, mostra che c'è una chiara relazione lineare positiva tra i rendimenti delle due azioni. Questo è indicato dalla distribuzione stretta dei punti intorno alla linea di regressione (diagonale), suggerendo che quando il rendimento di JPM aumenta, anche il rendimento di BAC tende ad aumentare, e viceversa.

Dunque, la dispersione dei punti nello scatterplot conferma una relazione lineare tra i due rendimenti.

## 2.6.3 Settore Bellico



## Relazioni e Andamento nel tempo

La correlazione tra HII e TDG cambia nel tempo mostrando una serie di alti e bassi, finora è il grafico più ondeggiante. In sintesi, la correlazione si è mossa in una banda molto larga di valori tra 0.1 e 0.6. nei 10 anni analizzati il valore medio di correlazione è 0.42 ma sempre avendo valori o superiori o inferiori.

Lo scatter plot dei rendimenti semplici di HII e TDG ci fornisce ulteriori dettagli sulle correlazioni in funzione dei rendimenti: La densità dei punti attorno all'origine suggerisce che la maggior parte dei rendimenti sono vicini allo zero come per il settore tecnologico. La nuvola di punti mostra una distribuzione che suggerisce una correlazione positiva, ma non perfetta.

In conclusione, lo scatter conferma la correlazione positiva tra i rendimenti dei due titoli ma questa relazione non è perfettamente lineare per via della presenza di molti punti sparsi lontani dalla linea diagonale.

## 3 Analisi di previsione

### 3.1.1 Metodologia

Ho deciso di costruire un modello di previsione dei **prezzi** utilizzando delle **SVMs**, la previsione viene effettuata giorno per giorno: dato in input il prezzo di chiusura del giorno di oggi, come output si avrà la previsione del prezzo di chiusura di domani. per ogni titolo sono stati eseguiti diversi passi:

1. Creazione dei dataset di **TRAIN**, **VALIDAZIONE** e **TEST**, rispettivamente di dimensioni di: 80 mesi, 30 mesi, 10 mesi.
2. Scalare le variabili indipendenti per rendere le computazioni più leggere e avere risultati migliori.
3. Creazione di un modello allenato con i dati di **TRAIN** e iperparametri ottimizzati sulla base dello score sul **VALIDATION** set

4. Refit del miglior modello con i dati di **TRAIN** e **VALIDATION**
5. Collezione delle predizioni: si parte dal primo prezzo del test set e la predizione viene effettuata giorno per giorno. Dopo che viene effettuata la predizione su un giorno il prezzo e la previsione reale che si è verificata vengono aggiunti al train set dopodiché refit del modello su tutto il train e si prosegue al giorno seguente.
6. Valutazione delle previsioni e visualizzazione dei dati.

### 3.1.2 Ottimizzazione Iperparametri

Gli iperparametri che ho fatto variare per ottenere i modelli sono:

Nome Iperparametro	Possibili Valori
C	[0.005, 0.01, 0.1, 1, 10, 100]
gamma	['scale', 'auto']
kernel	['rbf', 'linear']
epsilon	[0.01, 0.1, 0.5, 1]

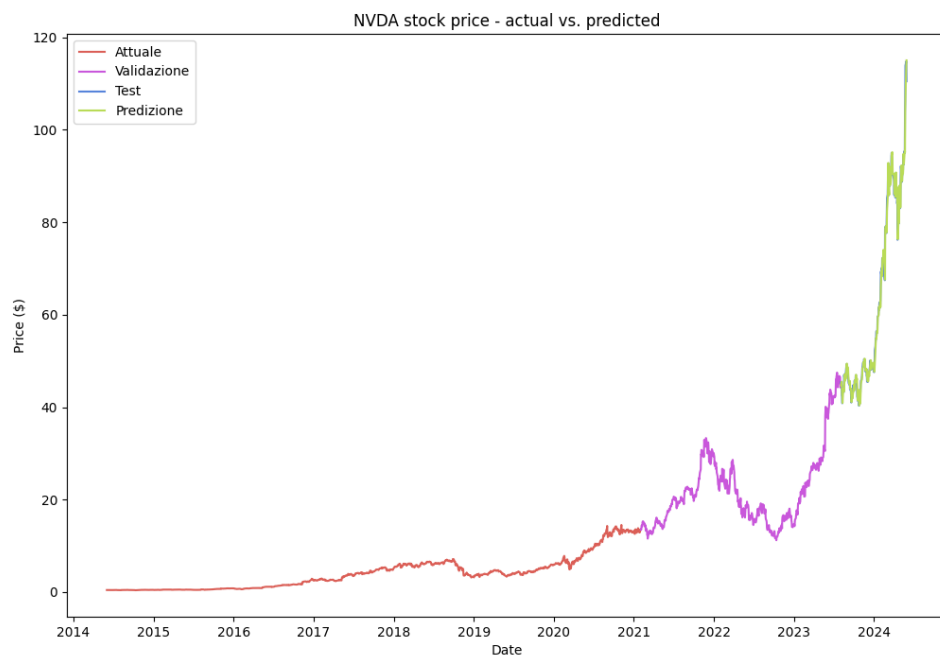
Ho scelto questi valori per la ricerca a griglia per esplorare una gamma ampia e diversificata di iperparametri. I valori di 'C' e 'epsilon' coprono diversi ordini di grandezza per garantire che il modello possa essere testato per varie complessità e tolleranze di errore, mentre il 'kernel' è stato fatto variare per trovare quello che meglio performa in base alla disposizione dei dati

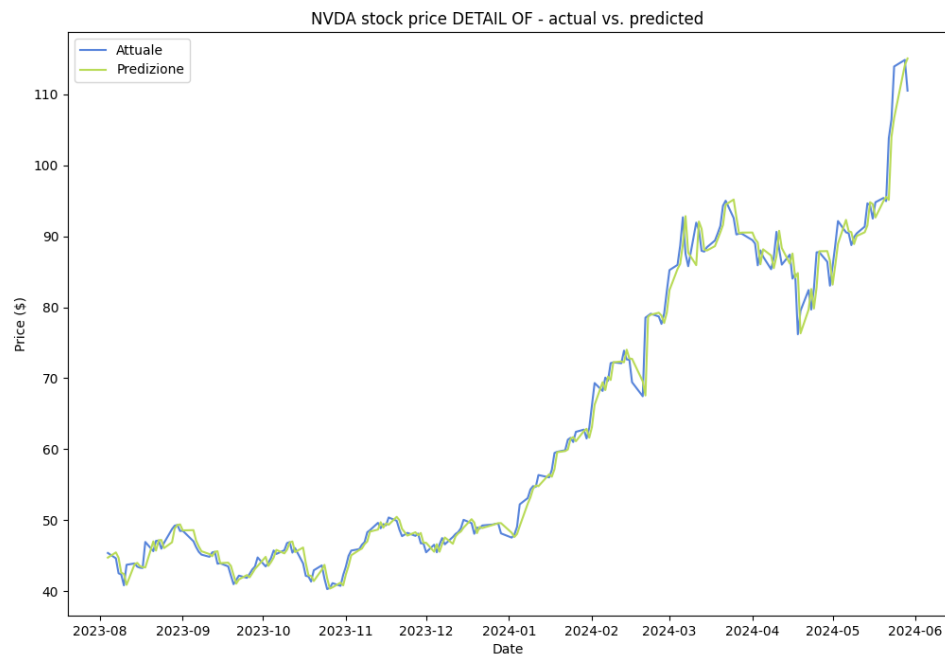


## 3.2 Risultati

### 3.2.1 NVIDIA

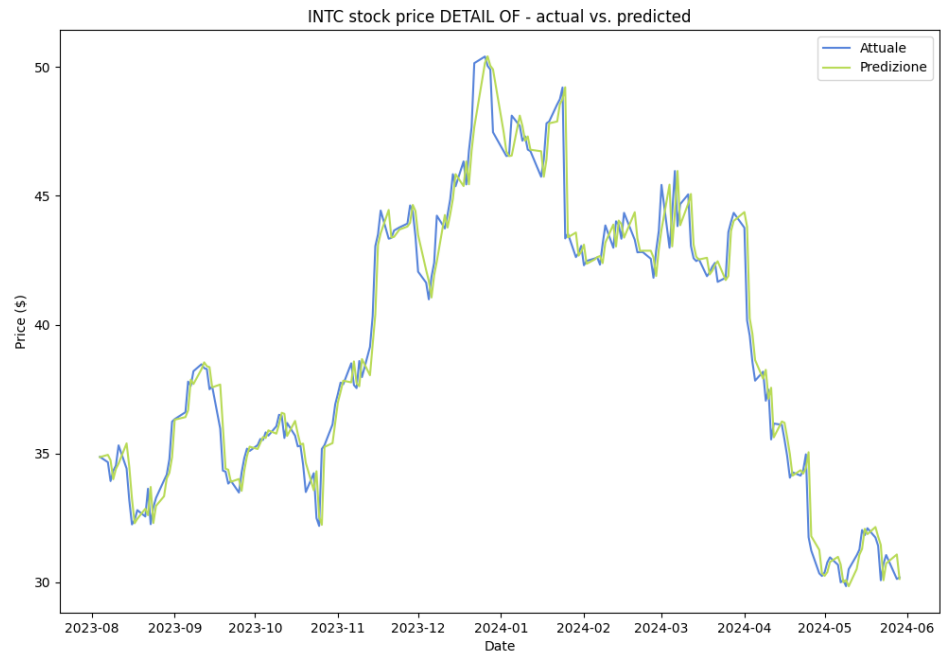
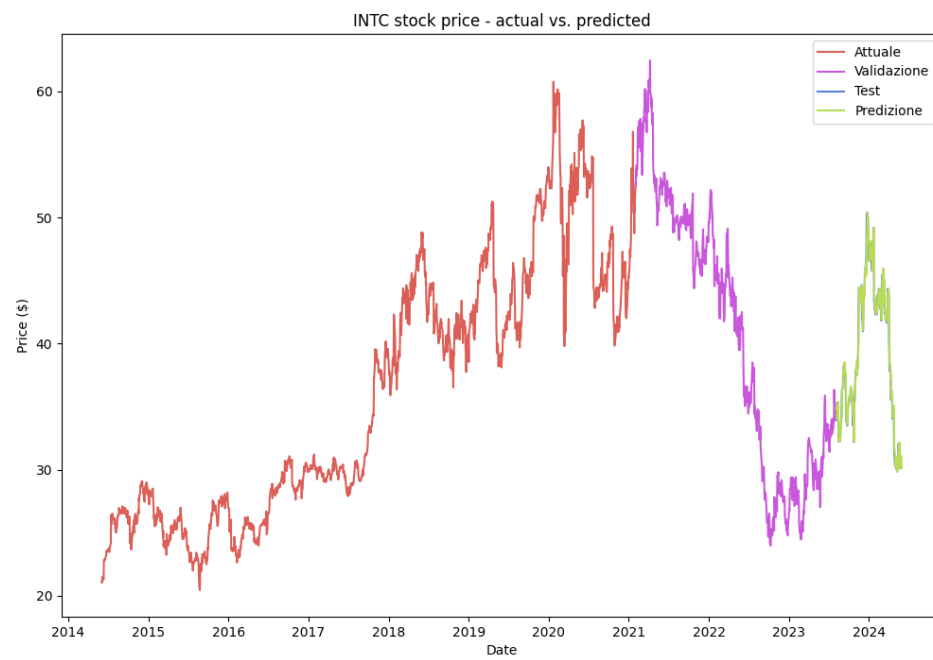
- Migliore configurazione:
  - Kernel: linear
  - C: 10
  - Epsilon: 0.01
  - Gamma: scale
- Risultati:
  - MSE: 4.2237848788378995
  - RMSE: 2.055184877045834
  - R2 SCORE: 0.9897057454210





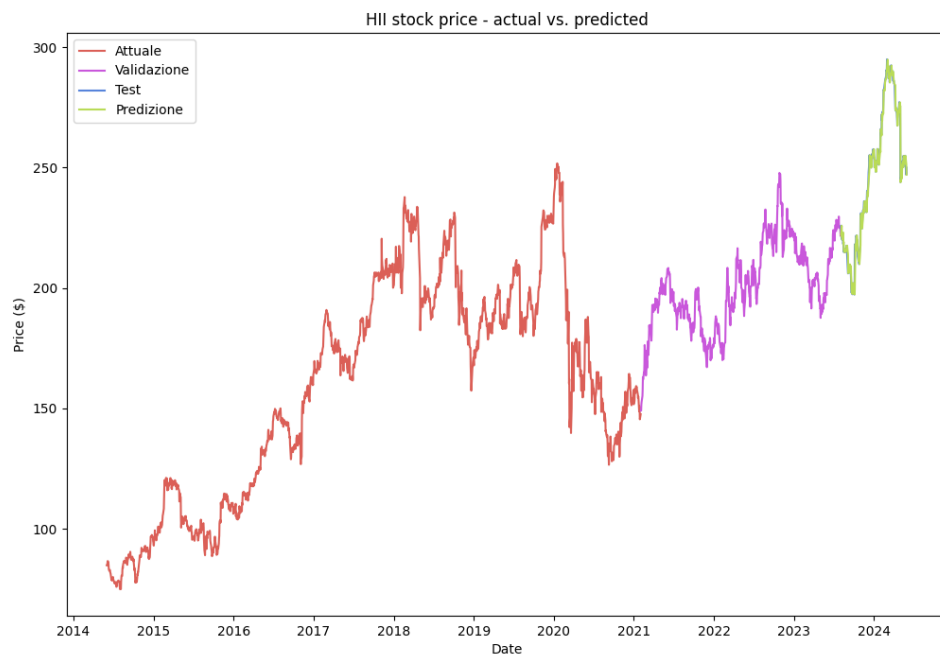
### 3.2.2 INTC

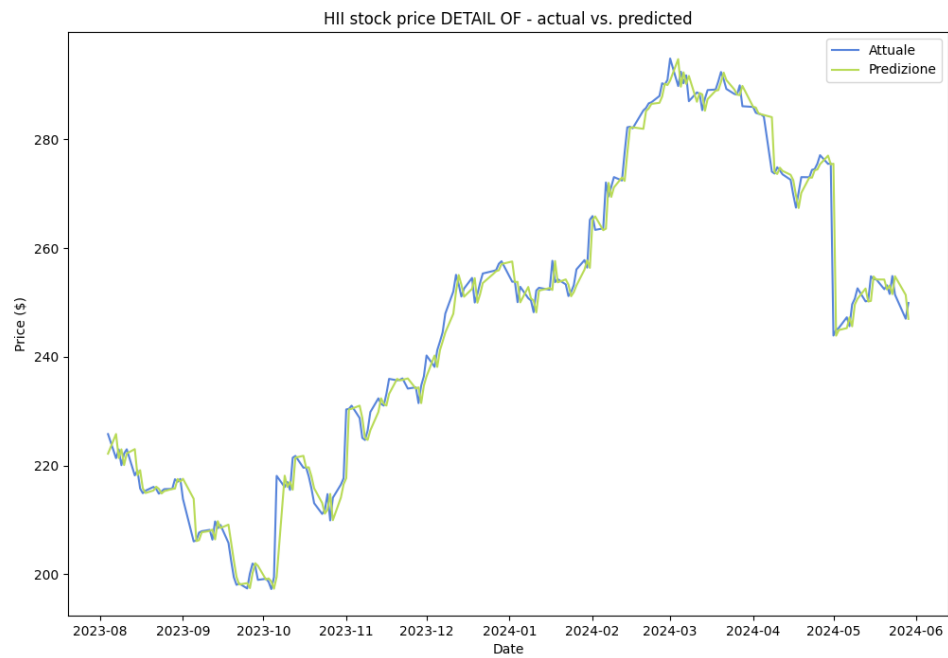
- Migliore configurazione:
  - Kernel: rbf
  - C: 100
  - Epsilon: 0.1
  - Gamma: scale
- Risultati:
  - MSE: 0.9776679279601863
  - RMSE: 0.9887709178369812
  - R2 SCORE: 0.9674698230824967



### 3.2.3 HII

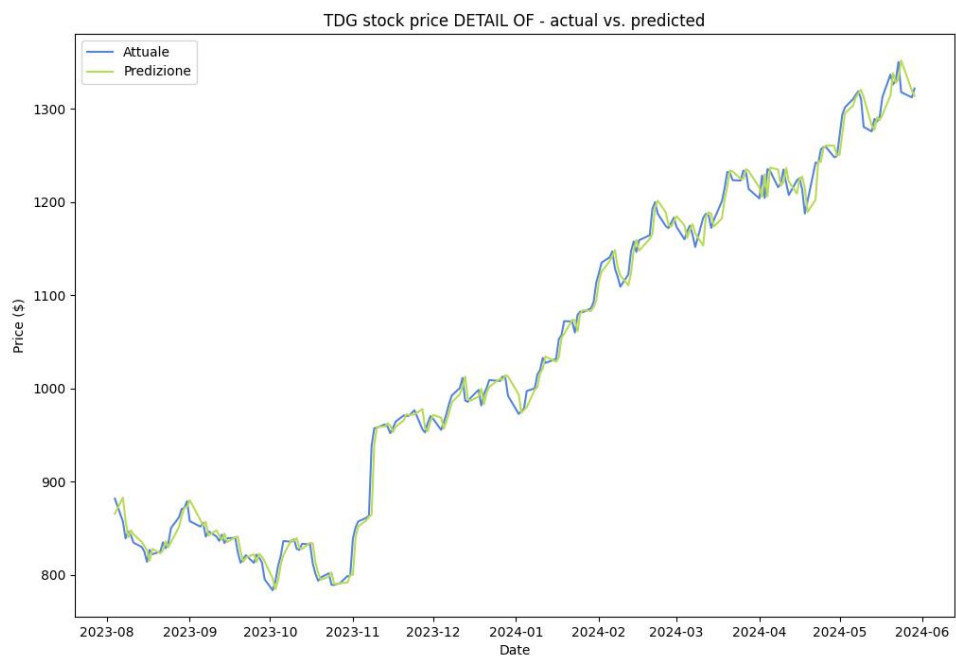
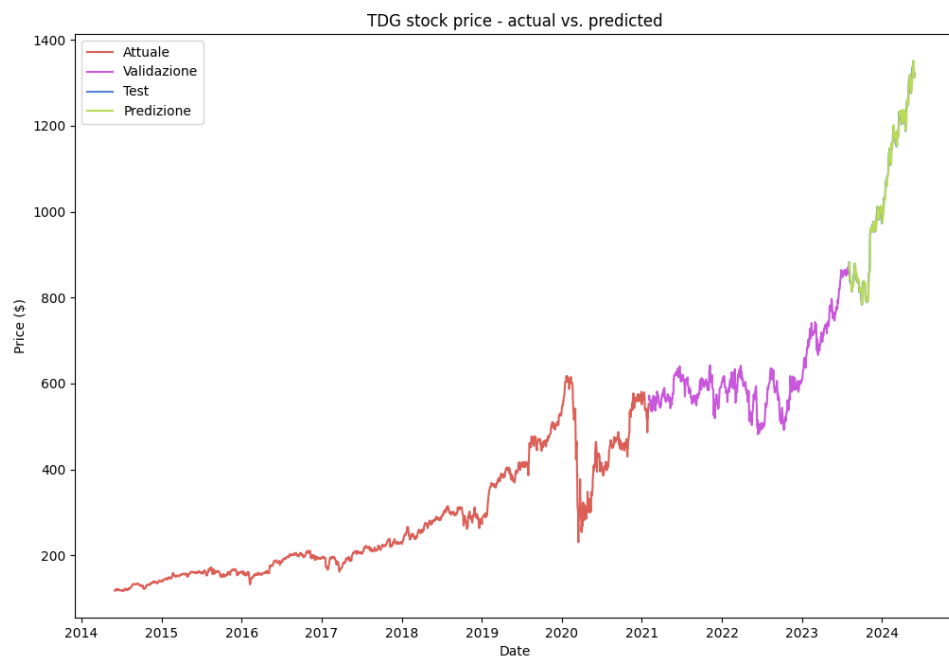
- Migliore configurazione:
  - Kernel: linear
  - C: 1
  - Epsilon: 1
  - Gamma: scale
- Risultati:
  - MSE: 13.671180834715445
  - RMSE: 3.6974559949667345
  - R2 SCORE: 0.9829669542947151





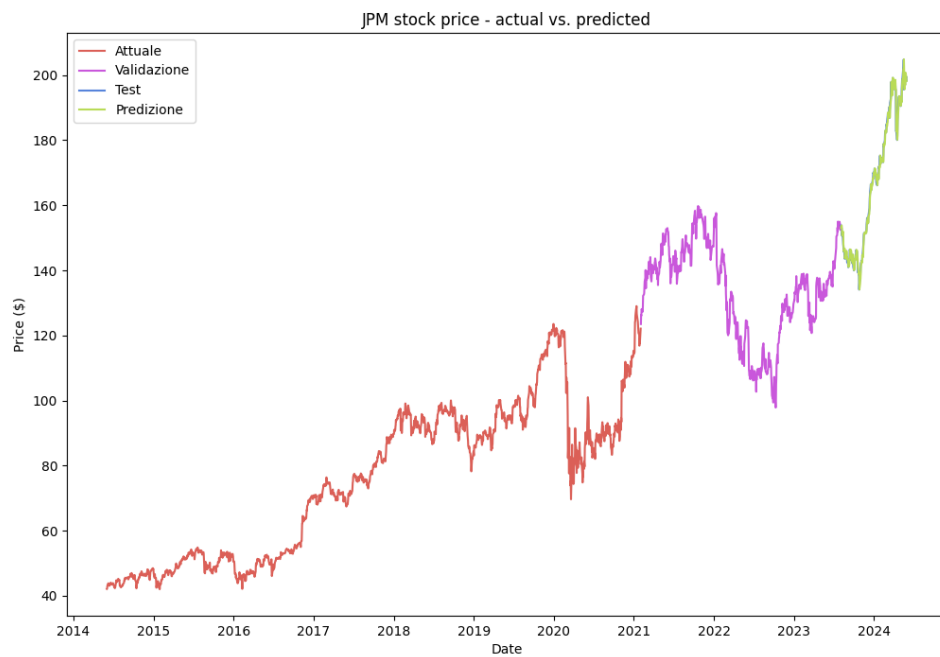
### 3.2.4 TDG

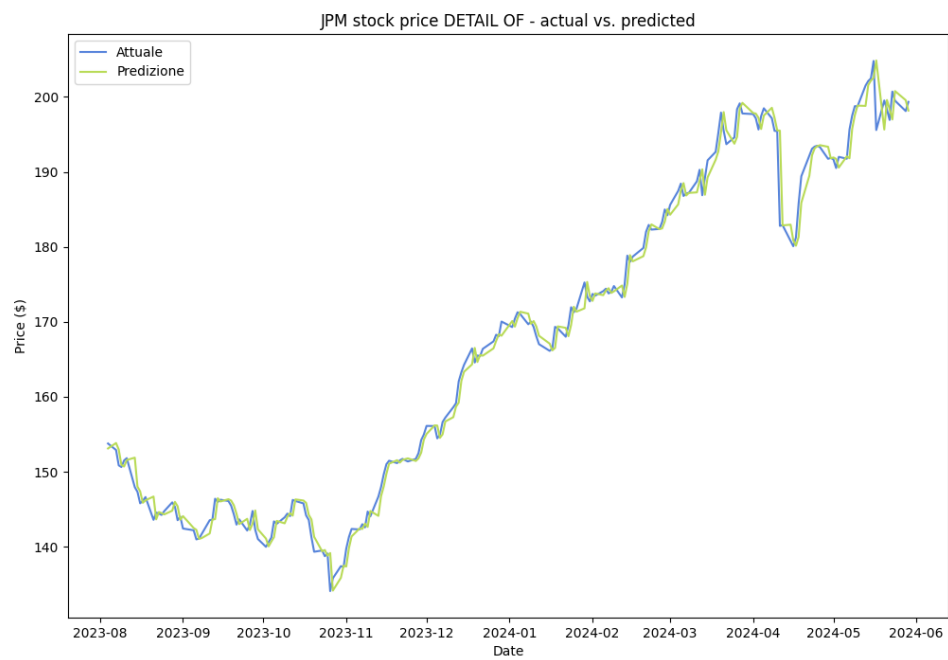
- Migliore configurazione:
  - Kernel: linear
  - C: 10
  - Epsilon: 0.5
  - Gamma: scale
- Risultati:
  - MSE: 181.47463122669566
  - RMSE: 13.471252028920537
  - R2 SCORE: 0.9939597686940409



### 3.2.5 JPM

- Migliore configurazione:
  - Kernel: linear
  - C: 100
  - Epsilon: 0.5
  - Gamma: scale
- Risultati:
  - MSE: 3.5600340128534733
  - RMSE: 1.8868052397779356
  - R2 SCORE: 0.9919917648062034

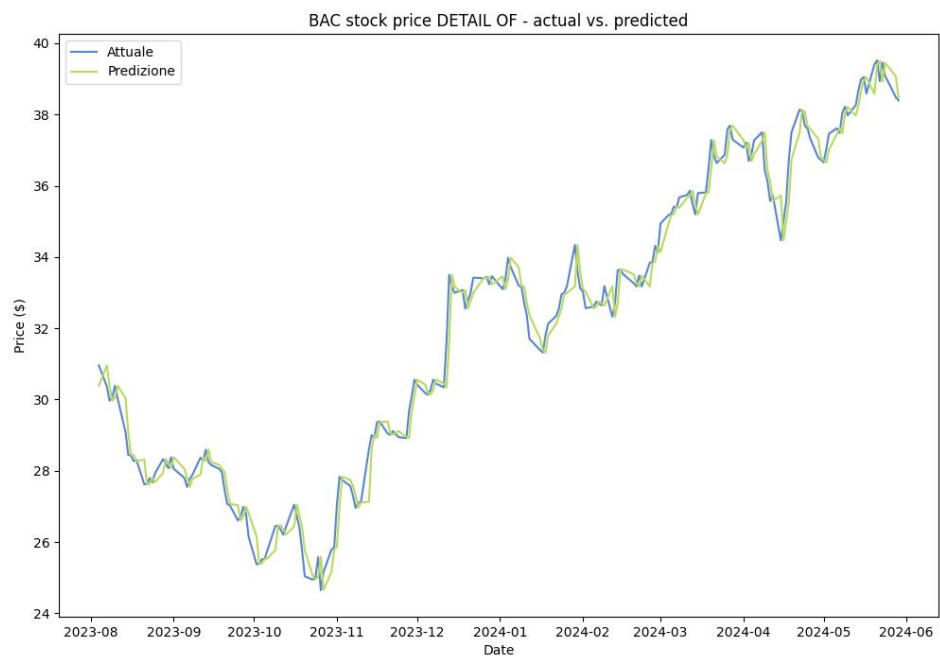
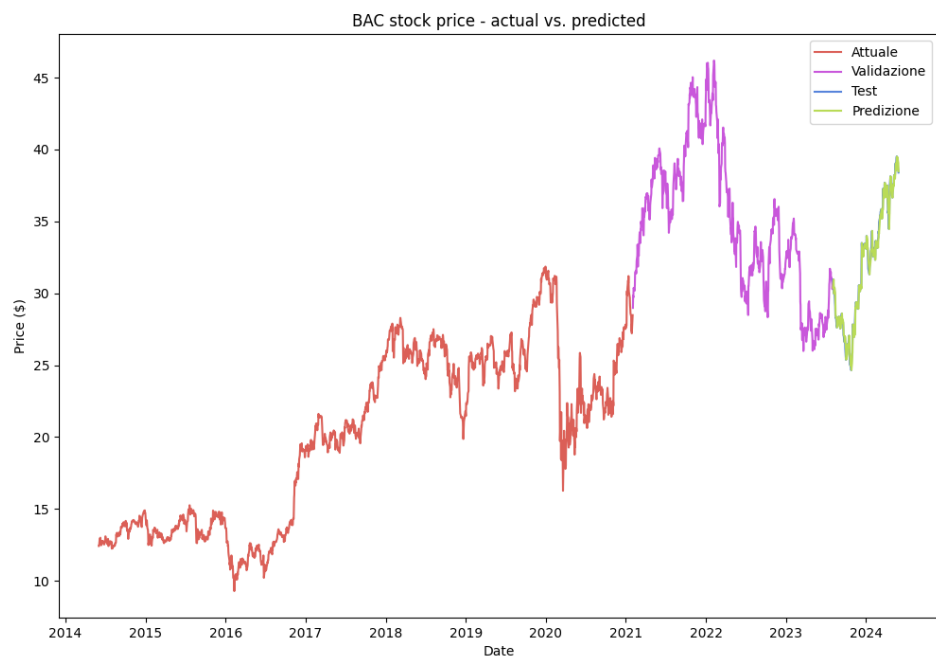




### 3.2.6 BAC

- Migliore configurazione:
  - Kernel: linear
  - C: 1
  - Epsilon: 0.1
  - Gamma: scale
- Risultati:
  - MSE: 0.21229063258266936
  - RMSE: 0.4607500760528091
  - R2 SCORE: 0.9876807217814838





### 3.2.7 Conclusioni

Tutti i modelli non hanno errori eccessivamente alti e intervalli di confidenza sempre superiori al 96%, i modelli prodotti sembrano essere accurati. Tuttavia, questi modelli hanno il difetto di prevedere soltanto il giorno successivo e quindi aumentando l'orizzonte temporale su cui fare le previsioni potrebbero essere meno performanti.

## 4 Strategie di Trading e Backtesting

Come strategia di Trading ho scelto di utilizzare la strategia che si basa sulle medie mobili semplici. Questa strategia consiste in:

1. Definire i periodi su cui calcolare le medie mobili semplici del titolo, vengono calcolate due medie mobili semplici una “corta” e una “lunga” (ad esempio: 50 giorni e 200 giorni)
2. Si applica la strategia sull'andamento del titolo:
  - Quando la SMA<sup>10</sup> più corta è sopra la SMA più lunga si assume una posizione long sul titolo.
  - Quando la SMA più corta è sotto la SMA più lunga si assume una posizione short sul titolo

### 4.1 Metodologia

Per la costruzione della strategia è stata utilizzata la seguente metodologia:

- **Scelta del titolo:** Ho scelto NVIDIA per applicare la mia strategia di trading perché ritengo sia interessante verificare se una strategia di trading possa superare i significativi rendimenti

---

<sup>10</sup> “simple moving average” o “media mobile semplice”

del metodo buy and hold, che NVIDIA ha ottenuto senza l'uso di alcuna strategia specifica.

- **Suddivisione Dataset:** Ho deciso di dividere il dataset in training (primi 90 mesi) e testing (ultimi 30 mesi) in modo da evitare l'overfitting dei parametri
- **Semplificazioni:** Non ho considerato i costi per le transazioni dato che si effettuano pochi trades nel corso di molti anni.
- **Ottimizzazione parametri:** per forza bruta ho trovato i due parametri cioè i due periodi su cui calcolare le medie mobili semplici che massimizzano i rendimenti sui primi 90 mesi.
- **BackTesting:** Ho applicato la strategia con i valori dei parametri trovati al punto prima e verificato i rendimenti con la strategia sugli ultimi 30 mesi.
- **Confronto con la buy and hold:** ho verificato la differenza di rendimenti tra la strategia e la buy and hold e ho verificato anche il livello di rischio tra le due.

## 4.2 Costruzione della strategia

Per la parte di Ottimizzazione ho trovato che i parametri migliori sui primi 90 mesi sono:

- **SMA1:** 32
- **SMA2:** 200
- **Market cum returns:** 58.23
- **Strategy cum returns:** 103.30
- **Out Performance (market returns – strategy returns):** 45.06

Applicando questi valori di SMA1 e SMA2 agli ultimi 30 mesi ho ottenuto la strategia visibile in Fig. 4.1. Fig. 4.1 Visualizzazione della strategia basata su medie mobili

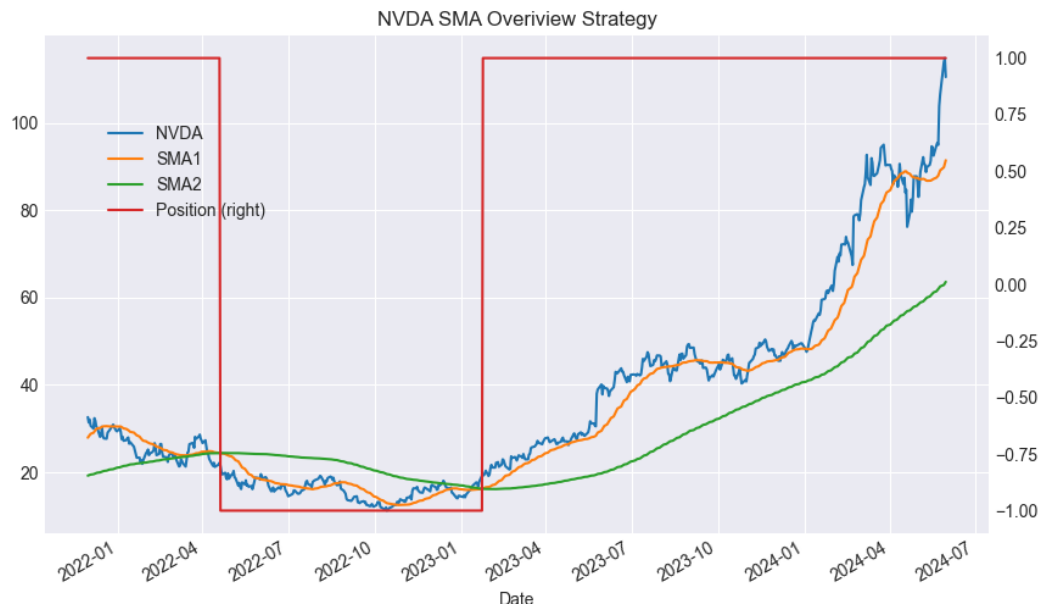


Fig. 4.1 Visualizzazione della strategia basata su medie mobili

Nel grafico in figura è possibile osservare:

1. **Medie Mobili:**

- **SMA1 (Linea Arancione):** Questa è la media mobile semplice a breve termine.
- **SMA2 (Linea Verde):** Questa è la media mobile semplice a lungo termine.

2. **Segnali di Trading:**

- **Segnale di Acquisto (Long):** Quando SMA1 (la media mobile a breve termine) incrocia verso l'alto SMA2 (la media mobile a lungo termine), la strategia indica di prendere una posizione lunga (long). Questo è rappresentato da un valore di posizione pari a 1.
- **Segnale di Vendita (Short):** Quando SMA1 incrocia verso il basso SMA2, la strategia indica di prendere una posizione

corta (short). Questo è rappresentato da un valore di posizione pari a 0.

## 4.3 Risultati e Confronto con la Buy and Hold

- **Ritorno cumulato Della strategia:** 4.370
- **Ritorno cumulato Buy and Hold:** 3.520
- **Out Performance (market returns – strategy returns):** 0.850

Rispetto ai ritorni è possibile osservare che la strategia ha performato meglio di 81 punti percentuali rispetto alla Buy and Hold.

In Fig. 4.2 è possibile osservare l'andamento delle due strategie, inoltre è possibile osservare le posizioni assunte, come possiamo vedere la strategia è riuscita a capitalizzare meglio il momento in cui il titolo è sceso, andando in short selling, ottenendo così rendimenti migliori.

Confrontiamo ora la volatilità annualizzata delle strategie:

- **Strategia:** 0.555129
- **Buy and Hold:** 0.554786

La piccola differenza tra 0.555129 e 0.554786 suggerisce che la strategia di trading applicata ha una volatilità molto simile a quella del titolo stesso. In altre parole, la strategia di trading non introduce un livello significativo di volatilità aggiuntiva e quindi abbiamo una performance migliore con lo stesso livello di rischio.

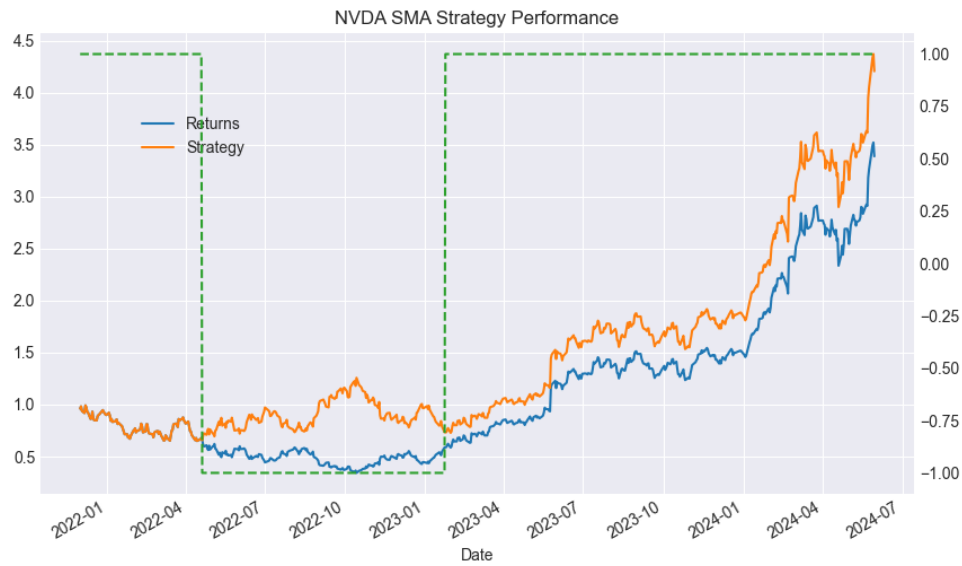


Fig. 4.2 Visualizzazione dei ritorni cumulati rispetto alle due strategie

## 5 Capital Asset Pricing Model - CAPM

Il CAPM (Capital Asset Pricing Model) `e un modello che rappresenta la relazione tra i rendimenti attesi di un indice rischioso ed il rischio di mercato (chiamato anche rischio sistematico). Per calcolare il rendimento atteso di un titolo nel CAPM si usa la seguente equazione:

$$E(r_i) = r_f + \beta_i(E(r_m) - r_f)$$

Dove:

$E(r_i)$  è il rendimento atteso dell'asset i

$r_f$  è il rate risk free

$E(r_m)$  è il rendimento atteso del mercato

$\beta_i$  è il coefficiente beta dell'asset i

## 5.1 Calcolo del beta

Per calcolare il beta ho utilizzato la formula:

$$\beta = \frac{cov(R_i, R_m)}{var(R_m)}$$

Per il calcolo del beta ho considerato l'indice S&P 500 (^GSPC)

I beta dei vari titoli sono:

- Indice beta per NVDA: 1.8058
- Indice beta per INTC: 0.9808
- Indice beta per HII: 0.7361
- Indice beta per TDG: 1.2942
- Indice beta per JPM: 1.1381
- Indice beta per BAC: 1.4049

### Valore del beta

### Interpretazione

$\beta < 0$	L'asset si muove nella direzione opposta al movimento del mercato
$\beta = 0$	I movimenti dell'asset e del mercato sono non correlati
$0 < \beta \leq 1$	L'asset si muove nella stessa direzione del mercato; la volatilità dell'asset può essere < o > della volatilità del mercato
$\beta > 1$	L'asset si muove nella stessa direzione del mercato ma con maggiore volatilità

## 5.2 Calcolo dei rendimenti attesi

Per il calcolo dei rendimenti attesi sono stati considerati i seguenti tassi:

- Tasso di interesse risk free: 0.0484 (tasso del tbill a un anno 2023)
- Tasso di interesse di mercato: 0.12299429408195972 (rendimento medio annualizzato di S&P500)

### Rendimenti attesi

Rendimento atteso per NVDA: 0.1831

Rendimento atteso per INTC: 0.1216

Rendimento atteso per HII: 0.1033

Rendimento atteso per TDG: 0.1449

Rendimento atteso per JPM: 0.1333

Rendimento atteso per BAC: 0.1532

## 5.3 Esposizione dei titoli ai fattori di rischio

### Fama e French

Il modello di Fama e French a tre fattori è un'estensione del Capital Asset Pricing Model (CAPM) ed è utilizzato per spiegare i rendimenti dei titoli attraverso tre fattori distinti. I tre fattori sono:

1. **Rendimento di Mercato (MKT)**: ha lo stesso significato del beta nel CAPM
2. **Dimensione (SMB - Small Minus Big)**: È calcolato come la differenza tra i rendimenti di un portafoglio di piccole imprese e un portafoglio di grandi imprese. Rappresenta il premio per il rischio legato alla dimensione delle aziende.



3. **Valore (HML - High Minus Low):** Rappresenta il premio per il rischio legato al rapporto valore/mercato (Value vs. Growth). È calcolato come la differenza tra i rendimenti di un portafoglio di azioni value e un portafoglio di azioni growth.

### 5.3.1 Formulazione matematica

Il modello di Fama e French a tre fattori può essere espresso matematicamente come:

$$R_i - R_f = \alpha + \beta(R_m - R_f) + s \cdot SMB + h \cdot HML$$

Dove:

- $R_i$  è il rendimento del titolo  $i$
- $R_f$  è il rendimento privo di rischio
- $\alpha$  è l'intercetta del modello, rappresenta il rendimento anomalo non spiegato dai fattori.
- $\beta$  è la sensibilità del rendimento del titolo al fattore di mercato (premio per il rischio di mercato).
- $s$  è la sensibilità del rendimento del titolo al fattore dimensione (premio per le piccole imprese).
- $h$  è la sensibilità del rendimento del titolo al fattore valore (premio per le azioni value).

## 5.3.2 Esposizione dei titoli

### NVDA

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn		R-squared:	0.419		
Model:	OLS		Adj. R-squared:	0.404		
Method:	Least Squares		F-statistic:	27.69		
Date:	Thu, 04 Jul 2024		Prob (F-statistic):	1.49e-13		
Time:	12:02:02		Log-Likelihood:	101.79		
No. Observations:	119		AIC:	-195.6		
Df Residuals:	115		BIC:	-184.5		
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0364	0.010	3.686	0.000	0.017	0.056
mkt	1.7879	0.220	8.124	0.000	1.352	2.224
smb	-0.0962	0.367	-0.262	0.794	-0.823	0.631
hml	-0.9307	0.258	-3.609	0.000	-1.441	-0.420
=====						
Omnibus:	6.778		Durbin-Watson:	1.837		
Prob(Omnibus):	0.034		Jarque-Bera (JB):	6.412		
Skew:	0.474		Prob(JB):	0.0405		
Kurtosis:	3.629		Cond. No.	39.3		

NVIDIA è un titolo altamente volatile, più simile a una grande azienda e fortemente orientato alla crescita. Questi fattori lo rendono attraente per gli investitori che cercano potenziali rendimenti elevati e sono disposti ad accettare una maggiore volatilità e rischio.

## INTC

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn		R-squared:		0.250	
Model:	OLS		Adj. R-squared:		0.231	
Method:	Least Squares		F-statistic:		12.80	
Date:	Thu, 04 Jul 2024		Prob (F-statistic):		2.82e-07	
Time:	12:04:09		Log-Likelihood:		140.97	
No. Observations:	119		AIC:		-273.9	
Df Residuals:	115		BIC:		-262.8	
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	-0.0043	0.007	-0.605	0.546	-0.018	0.010
mkt	0.9465	0.158	5.978	0.000	0.633	1.260
smb	-0.0806	0.264	-0.305	0.761	-0.604	0.443
hml	0.0488	0.186	0.263	0.793	-0.319	0.416
=====						
Omnibus:	14.277		Durbin-Watson:		1.903	
Prob(Omnibus):	0.001		Jarque-Bera (JB):		42.917	
Skew:	-0.238		Prob(JB):		4.79e-10	
Kurtosis:	5.903		Cond. No.		39.3	
=====						

Intel è un titolo che si comporta in linea con il mercato complessivo, con una leggera inclinazione verso le caratteristiche delle grandi aziende e delle azioni di valore.

## TDG

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn		R-squared:	0.448		
Model:	OLS		Adj. R-squared:	0.434		
Method:	Least Squares		F-statistic:	31.13		
Date:	Thu, 04 Jul 2024		Prob (F-statistic):	8.30e-15		
Time:	12:04:49		Log-Likelihood:	158.79		
No. Observations:	119		AIC:	-309.6		
Df Residuals:	115		BIC:	-298.5		
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0115	0.006	1.874	0.063	-0.001	0.024
mkt	1.2641	0.136	9.273	0.000	0.994	1.534
smb	-0.0972	0.227	-0.427	0.670	-0.548	0.353
hml	0.1339	0.160	0.838	0.404	-0.182	0.450
=====						
Omnibus:	14.339		Durbin-Watson:	2.385		
Prob(Omnibus):	0.001		Jarque-Bera (JB):	26.410		
Skew:	-0.498		Prob(JB):	1.84e-06		
Kurtosis:	5.082		Cond. No.	39.3		

TDG è un titolo che è più volatile rispetto al mercato complessivo, con una leggera inclinazione verso le caratteristiche delle grandi aziende e delle azioni di valore. Adatto ad investitori che cercano potenziali rendimenti elevati pur accentuando più rischio.

### HII

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn	R-squared:	0.257			
Model:	OLS	Adj. R-squared:	0.237			
Method:	Least Squares	F-statistic:	13.25			
Date:	Thu, 04 Jul 2024	Prob (F-statistic):	1.73e-07			
Time:	12:05:17	Log-Likelihood:	160.39			
No. Observations:	119	AIC:	-312.8			
Df Residuals:	115	BIC:	-301.7			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0053	0.006	0.879	0.381	-0.007	0.017
mkt	0.7091	0.134	5.273	0.000	0.443	0.976
smb	-0.0459	0.224	-0.205	0.838	-0.490	0.399
hml	0.4721	0.158	2.996	0.003	0.160	0.784
=====						
Omnibus:	0.228	Durbin-Watson:	2.044			
Prob(Omnibus):	0.892	Jarque-Bera (JB):	0.114			
Skew:	-0.076	Prob(JB):	0.945			
Kurtosis:	3.008	Cond. No.	39.3			
=====						

HII è un titolo che è meno volatile rispetto al mercato complessivo, con una leggera inclinazione verso le caratteristiche delle grandi aziende e delle azioni di valore. Questi fattori lo rendono attraente per gli investitori che cercano stabilità e minori rischi di mercato.

## JPM

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn		R-squared:	0.749		
Model:	OLS		Adj. R-squared:	0.742		
Method:	Least Squares		F-statistic:	114.2		
Date:	Thu, 04 Jul 2024		Prob (F-statistic):	2.45e-34		
Time:	12:05:53		Log-Likelihood:	230.60		
No. Observations:	119		AIC:	-453.2		
Df Residuals:	115		BIC:	-442.1		
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0056	0.003	1.669	0.098	-0.001	0.012
mkt	1.0861	0.075	14.568	0.000	0.938	1.234
smb	0.0598	0.124	0.481	0.631	-0.187	0.306
hml	0.8499	0.087	9.731	0.000	0.677	1.023
=====						
Omnibus:	0.362		Durbin-Watson:	1.836		
Prob(Omnibus):	0.834		Jarque-Bera (JB):	0.371		
Skew:	-0.128		Prob(JB):	0.830		
Kurtosis:	2.903		Cond. No.	39.3		

JPMorgan è un titolo che è leggermente più volatile rispetto al mercato complessivo, con una leggera inclinazione verso le caratteristiche delle aziende di medie dimensioni e una forte caratterizzazione di titolo di valore. Questi fattori lo rendono attraente per gli investitori che cercano una combinazione di crescita moderata e stabilità, con una solida base nei fondamentali aziendali.

## BAC

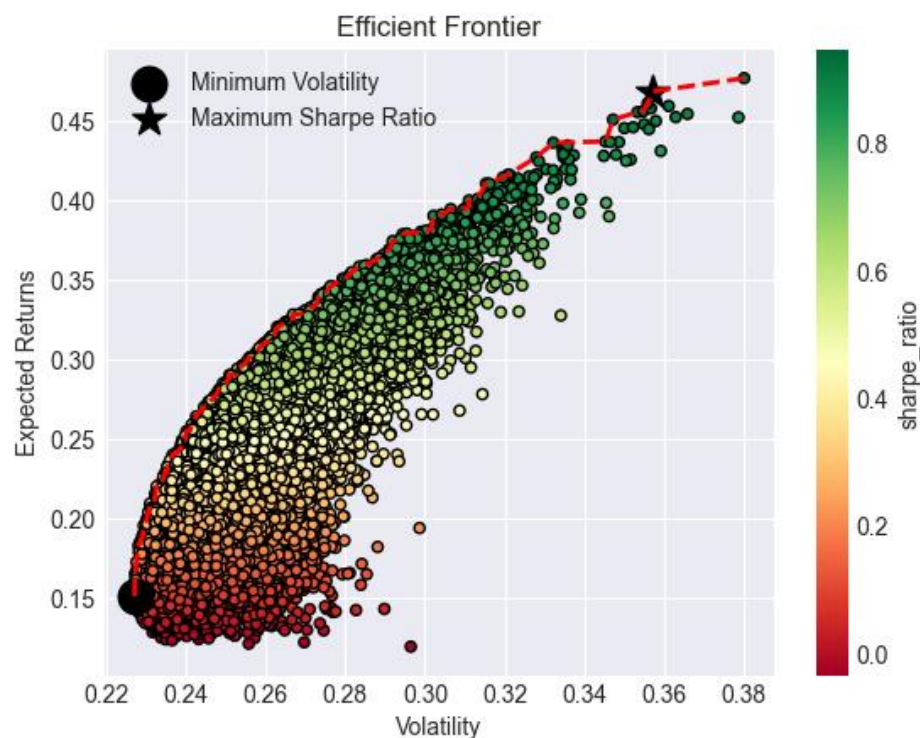
OLS Regression Results						
=====						
Dep. Variable:	excess_rtn	R-squared:	0.762			
Model:	OLS	Adj. R-squared:	0.756			
Method:	Least Squares	F-statistic:	123.1			
Date:	Thu, 04 Jul 2024	Prob (F-statistic):	9.56e-36			
Time:	12:06:18	Log-Likelihood:	210.08			
No. Observations:	119	AIC:	-412.2			
Df Residuals:	115	BIC:	-401.0			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
Intercept	0.0014	0.004	0.351	0.726	-0.006	0.009
mkt	1.3412	0.089	15.141	0.000	1.166	1.517
smb	0.1093	0.148	0.740	0.461	-0.183	0.402
hml	1.0299	0.104	9.924	0.000	0.824	1.235
=====						
Omnibus:	1.977	Durbin-Watson:	1.896			
Prob(Omnibus):	0.372	Jarque-Bera (JB):	1.582			
Skew:	-0.077	Prob(JB):	0.453			
Kurtosis:	3.544	Cond. No.	39.3			

Bank of America è un titolo che è più volatile rispetto al mercato complessivo, con una leggera inclinazione verso le caratteristiche delle aziende di medie dimensioni e una forte caratterizzazione di titolo di valore. Questi fattori lo rendono attraente per gli investitori che cercano potenziali rendimenti elevati avendo comunque una certa stabilità derivante dalle azioni di valore.

## 6 Creazione di un portafoglio

### 6.1 Metodo di simulazione

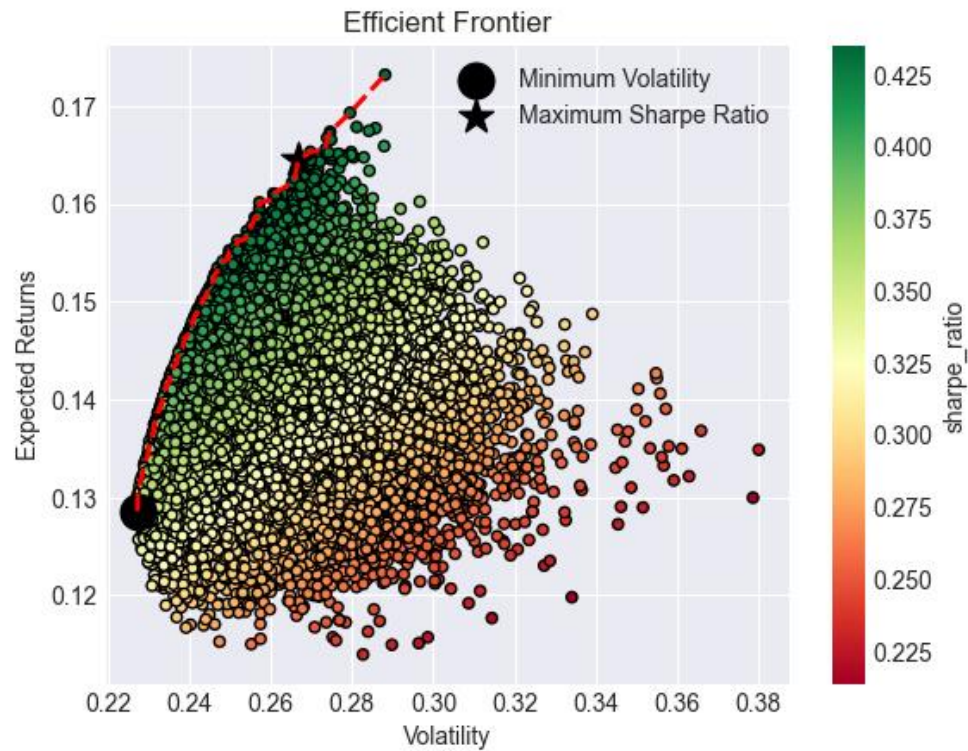
#### 6.1.1 Primi 108 mesi



```
Minimum Volatility Portfolio:
Returns: 0.15157344723892582
Volatility: 0.2270294018252584
Sharpe Ratio: 0.09546537613488333
Weights:
{'BAC': 0.025156280824076796, 'HII': 0.3993715053451653, 'INTC': 0.20082718127963528,
'JPM': 0.2910684366005092, 'NVDA': 0.014945837426698935, 'TDG': 0.06863075852391458}

Maximum Sharpe Ratio Portfolio:
Returns: 0.4682283660777049
Volatility: 0.35728427942816954
Sharpe Ratio: 0.9469444516819957
Weights:
[{'BAC': 0.01023381710950099, 'HII': 0.020691553652888082, 'INTC': 0.02675238836466937,
'JPM': 0.024898914800193305, 'NVDA': 0.6161067486778329, 'TDG': 0.3013165773949153}]
```

## 6.1.2 Rendimenti attesi



```
Minimum Volatility Portfolio:
Returns: 0.1285954464346385
Volatility: 0.2270294018252584
Sharpe Ratio: 0.35323815237095985
Weights:
{'BAC': 0.025156280824076796, 'HII': 0.3993715053451653, 'INTC': 0.20082718127963528,
'JPM': 0.2910684366005092, 'NVDA': 0.014945837426698935, 'TDG': 0.06863075852391458}

Maximum Sharpe Ratio Portfolio:
Returns: 0.1644385320779197
Volatility: 0.2667215819666686
Sharpe Ratio: 0.43505490340268255
Weights:
{'BAC': 0.5583383552170579, 'HII': 0.11337943072094825, 'INTC': 0.02223450280008831,
'JPM': 0.04814473014216739, 'NVDA': 0.01814098759203266, 'TDG': 0.2397619935277055}
```



## 6.2 Metodo Analitico

### 6.2.1 Primi 108 mesi

```
Minimum Volatility Portfolio:
Returns: 0.1541131810783119
Volatility: 0.22653142811544136
Sharpe Ratio: 0.46666010962699633
Weights:
{'BAC': 2.770824240557247e-17, 'HII': 0.3951068519722865, 'INTC': 0.20765890415551475,
'JPM': 0.2960163514878126, 'NVDA': 0.016075944437095722, 'TDG': 0.08514194794729042}

Maximum Sharpe Ratio Portfolio:
Returns: 0.6013045625109045
Volatility: 0.4655001636259854
Sharpe Ratio: 1.1877644858469818
Weights:
{'BAC': 7.586569934065849e-17, 'HII': 0.0, 'INTC': 2.313575392157747e-17,
'JPM': 0.0, 'NVDA': 0.9684588048640864, 'TDG': 0.03154119513591375}
```

### 6.2.2 Rendimenti Attesi

```
Minimum Volatility Portfolio:
Returns: 0.12757386805269877
Volatility: 0.22653142811544136
Sharpe Ratio: 0.34950500560280506
Weights:
{'BAC': 2.770824240557247e-17, 'HII': 0.3951068519722865, 'INTC': 0.20765890415551475,
'JPM': 0.2960163514878126, 'NVDA': 0.016075944437095722, 'TDG': 0.08514194794729042}

Maximum Sharpe Ratio Portfolio:
Returns: 0.16929563049704896
Volatility: 0.2742714933904925
Sharpe Ratio: 0.44078817307099605
Weights:
{'BAC': 0.6799060741938546, 'HII': 0.13329106377154534, 'INTC': 0.0,
'JPM': 4.483175983227561e-17, 'NVDA': 0.000817946979823517, 'TDG': 0.18598491505477646}
```

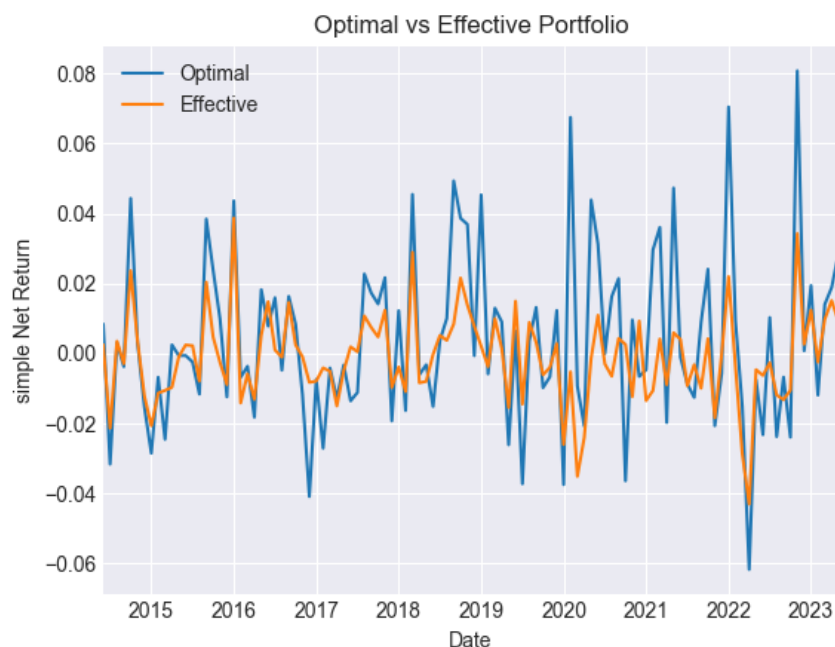
## 6.3 Beta dei portafogli Ottimali

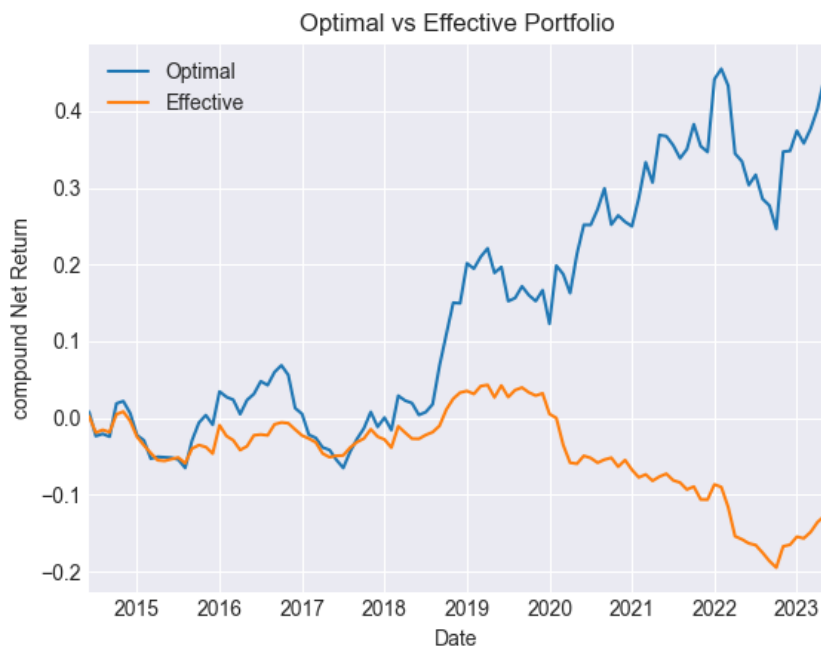
- Beta of Maximum Sharpe Ratio Portfolio (historical data):  
1.7896519268658122
- Beta of Maximum Sharpe Ratio Portfolio (forecasted data):  
1.2954675919537502

## 6.4 Confronto con il Portafogli Effettivo

### 6.4.1 Rispetto ai primi 108 mesi

```
Effective Portfolio:  
Returns: 0.2349050220130427  
Volatility: 0.24640334858600232  
Sharpe Ratio: 0.7569094457656963  
Weights:  
{'BAC': 0.16666666666666666, 'HII': 0.16666666666666666, 'INTC': 0.16666666666666666,  
'JPM': 0.16666666666666666, 'NVDA': 0.16666666666666666, 'TDG': 0.16666666666666666}
```





Dai grafici e dai numeri è possibile osservare che il portafogli ottimale presenta dei rendimenti nettamente maggiori rispetto a quello effettivo.

## 6.4.2 Confronto rispetto ai rendimenti attesi

```
Effective Portfolio:
Returns: 0.13990011520771153
Volatility: 0.24640334714202924
Sharpe Ratio: 0.3713428257732635
Weights:
{'BAC': 0.16666666666666666, 'HII': 0.16666666666666666, 'INTC': 0.16666666666666666,
'JPM': 0.16666666666666666, 'NVDA': 0.16666666666666666, 'TDG': 0.16666666666666666}
```

In questo caso non ho l'opportunità di creare un grafico, in quanto questo portafoglio è stato calcolato usando i ritorni attesi tra un anno. tuttavia è comunque possibile confrontarlo rispetto a quello ottimale trovato col metodo analitico sempre rispetto ai dati di previsione.

In questo caso, a differenza del confronto sui dati reali in cui il portafoglio ottimale rende meglio rispetto a quello effettivo, possiamo notare che i portafogli sono molto simili sia in termini di ritorni che di volatilità

### **6.4.3 Conclusioni**

Abbiamo osservato che il portafoglio ottimizzato sui primi 108 mesi ha una performance significativamente migliore rispetto al portafoglio effettivo. Tuttavia, questo portafoglio include NVIDIA al 96%, risultando poco diversificato e potenzialmente soggetto a ingenti perdite.

D'altro canto, il portafoglio ottimizzato basato sui rendimenti attesi è leggermente più diversificato, includendo principalmente tre titoli (BAC, TDG, HII). Nonostante ciò, presenta uno Sharpe Ratio relativamente basso (0.44), rendendolo un'opzione di investimento meno attraente.

Per migliorare questo portafoglio, sarebbe opportuno aggiungere titoli meno correlati tra loro con rendimenti alti e includere un asset risk-free.