

REPORT PROGETTO BUSINESS INTELLIGENCE PER I SERVIZI FINANZIARI
Agnese Belloni matricola 880126

SOMMARIO DEI DATI UTILIZZATI

Titoli S&P500 utilizzati per questo progetto:

1) Settore tecnologia dell'informazione

- a) **Alphabet Inc. (GOOGL)**: è una holding multinazionale americana, società madre di Google e di diverse altre consociate precedentemente controllate da Google. Alphabet Inc. è organizzata come un conglomerato, ovvero possiede una serie di aziende diverse, anche in settori differenti. Alphabet Inc. rappresenta un modello di business innovativo che permette a Google di mantenere la sua leadership nel settore digitale, promuovendo al tempo stesso la ricerca e lo sviluppo di diversi settori emergenti.
- b) **Intel Corporation (INTC)**: si tratta di uno dei principali produttori di semiconduttori e processori al mondo. È nota in particolare per i suoi microprocessori, che alimentano una vasta gamma di dispositivi informatici. Tra i suoi prodotti e tecnologie principali troviamo i microprocessori, i chipset e le schede madri, le memorie e gli storage, le tecnologie di connettività e le piattaforme per i data center e AI. Intel Corporation è una colonna portante dell'industria tecnologica, conosciuta per la sua capacità di innovazione e per il suo impatto significativo sul processo tecnologico globale.

2) Settore salute

- a) **Johnson & Johnson (JNJ)**: è una delle più grandi multinazionali nel settore della salute e del benessere, rinomata per la sua vasta gamma di prodotti farmaceutici, dispositivi medici e di consumo. Inoltre, investe significativamente nella ricerca e sviluppo, possedendo svariati laboratori e centri di ricerca a livello globale.
- b) **Pfizer Inc. (PFE)**: si tratta di una delle più grandi aziende farmaceutiche, nota per lo sviluppo e la produzione di vaccini, farmaci e altre terapie. Pfizer si concentra in svariate aree terapeutiche: l'area farmaceutica, che comprende i vaccini, l'oncologia, l'immunologia e le malattie rare, e i prodotti da banco e di consumo, cioè prodotti da banco, integratori e prodotti per la cura personale.

3) Settore energetico

- a) **Exxon Mobil Corporation (XOM)**: è una multinazionale leader nel settore dell'energia e dei prodotti chimici. Exxon Mobil opera in tre principali settori: l'esplorazione e la produzione (*Upstream*), la raffinazione e la distribuzione (*Downstream*) e i prodotti chimici. Inoltre, la multinazionale investe considerevolmente in ricerca e sviluppo per migliorare l'efficienza di nuove operazioni e sviluppare nuove tecnologie energetiche.
- b) **Chevron Corporation (CVX)**: è un attore chiave nell'industria energetica globale, attiva in tutti i settori principali dell'industria energetica: l'esplorazione e la produzione (*Upstream*), la raffinazione e la distribuzione (*Downstream*) e i prodotti chimici e lubrificanti. La compagnia investe in innovazione e tecnologia per affrontare sfide future, come quella della transizione energetica.

Per scaricare i dati relativi ai titoli unirli in un unico dataset ho usato il seguente codice:

```
[236]: start=datetime.date(2014,5,31)
end=datetime.date(2024,3,31)

[238]: def get(tickers, start, end):
    def data(ticker):
        return yf.download(ticker,start,end)
    datas=map(data,tickers)
    return pd.concat(datas,keys=tickers,names=['Ticker','Date'])

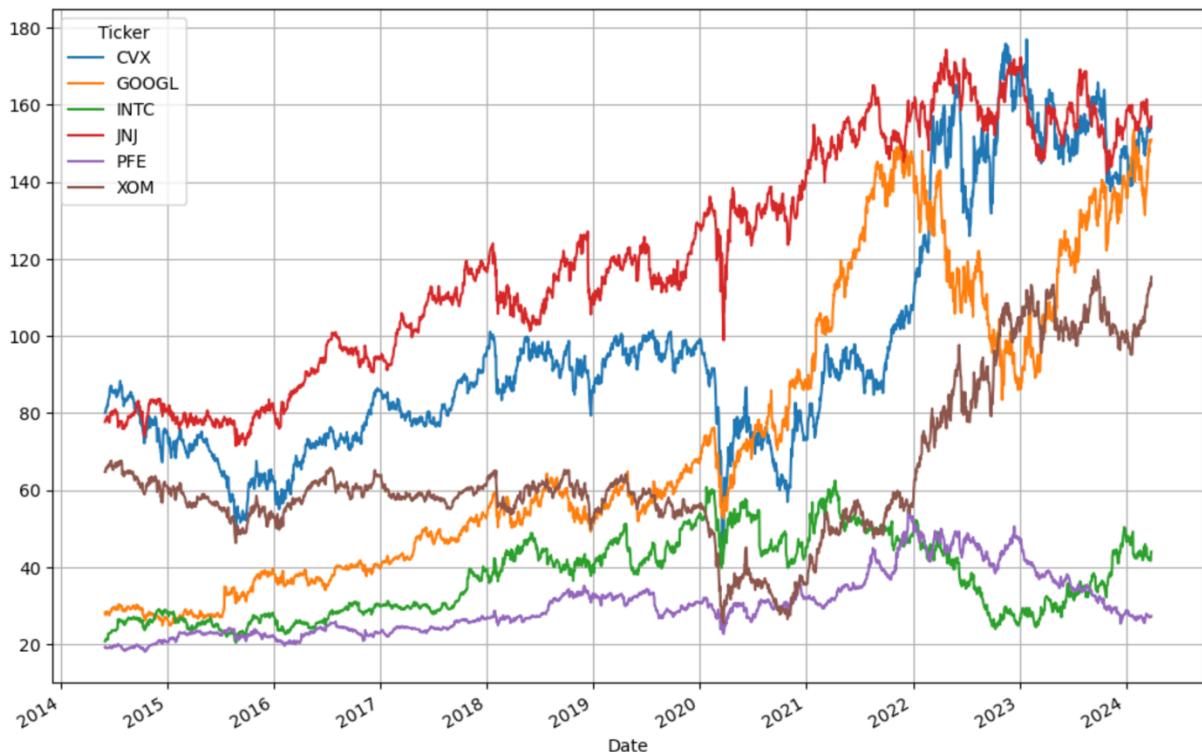
[240]: tickers=['GOOGL','INTC','JNJ','PFE','XOM','CVX']
all_data=get(tickers,start,end)
```

Il codice definisce l'intervallo di tempo tra il 31 maggio 2014 e il 31 marzo 2024 e una lista contenente i sei titoli azionari scelti. Per scaricare i dati storici dei titoli azionari per l'intervallo di tempo specificato ho utilizzato la libreria 'yfinance'. La funzione 'get' scarica i dati e li concatena in un unico DataFrame con multi-indice formato da 'ticker', che contiene tutti i titoli utilizzati, e 'date', che contiene le date.

		Open	High	Low	Close	Adj Close	Volume
Ticker	Date						
GOOGL	2014-06-02	28.487499	28.520500	27.834999	28.216999	28.216999	33210000
	2014-06-03	28.045000	28.120001	27.626499	27.725500	27.725500	40626000
	2014-06-04	27.576000	27.929001	27.427999	27.688000	27.688000	34582000
	2014-06-05	27.855499	28.250000	27.752501	28.246500	28.246500	35868000
	2014-06-06	28.408001	28.440001	27.979000	28.301500	28.301500	34804000

CVX	2024-03-22	155.080002	155.279999	154.160004	154.660004	153.113876	5973700
	2024-03-25	155.479996	157.309998	155.270004	156.470001	154.905777	6058700
	2024-03-26	156.520004	156.860001	154.649994	155.270004	153.717773	6677900
	2024-03-27	154.869995	156.440002	154.850006	156.350006	154.786987	7416900
	2024-03-28	157.149994	158.190002	156.369995	157.740005	156.163086	8331600

14844 rows × 6 columns



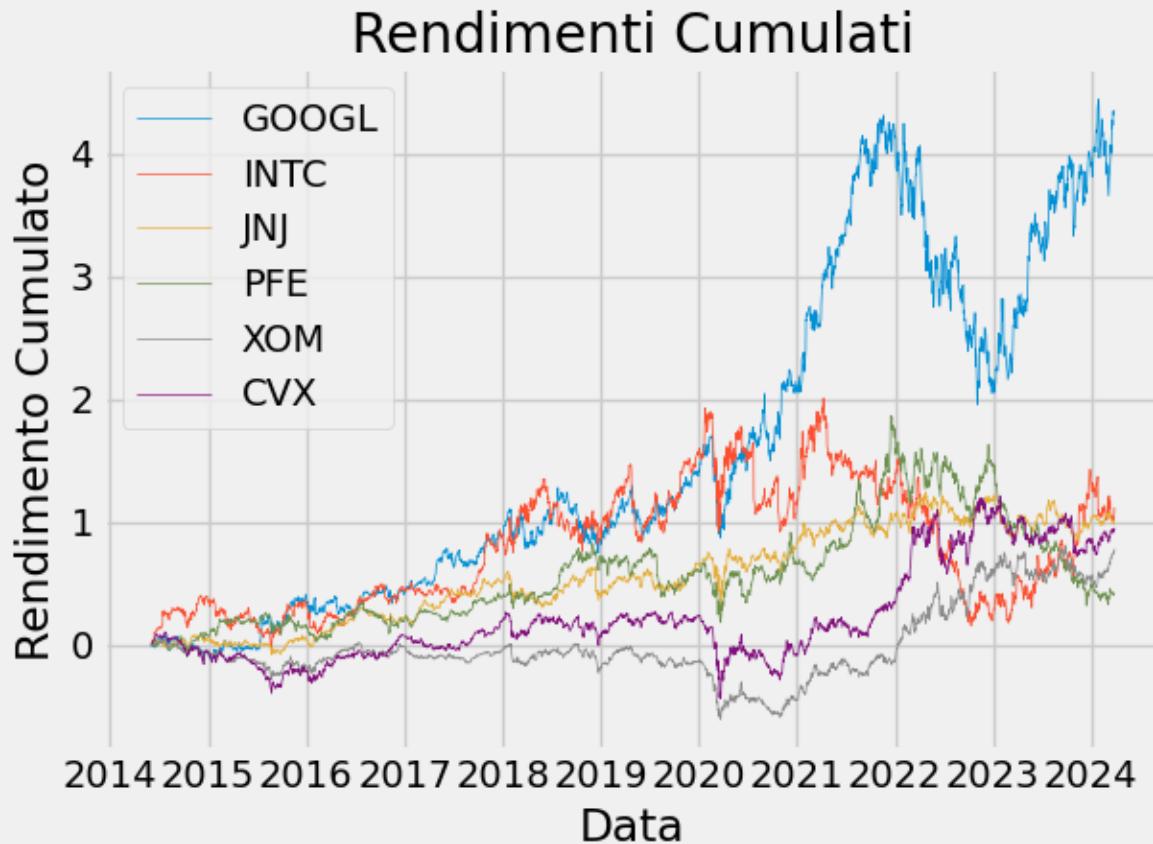
Il grafico mostra l'andamento dei prezzi di chiusura giornalieri aggiustati per i sei titoli presi in considerazione. Sull'asse delle ascisse troviamo gli anni presi in considerazione nella nostra analisi, mentre sull'asse delle ordinate troviamo i prezzi di chiusura aggiustati. Questa rappresentazione permette all'investitore di confrontare la variazione dei prezzi di chiusura nel tempo per ciascun titolo, facilitando l'analisi delle prestazioni relative delle azioni nel portafoglio o in uno specifico settore. Ora analizziamo le prestazioni dei singoli titoli:

- 1) CVX: possiamo notare un generale aumento dei prezzi di chiusura, spezzata però da una decrescita nel 2020, che rende il titolo abbastanza volatile negli anni.
- 2) GOOGL: notiamo che i prezzi di chiusura aumentano con il passare degli anni, con un particolare incremento dopo il 2020.
- 3) INTC: in questo caso i prezzi di chiusura mostrano una crescita moderata e abbastanza costante, con un leggero calo nel 2023 e una ripresa nel 2024.
- 4) JNJ: possiamo notare come i prezzi crescano in modo costante, con una discreta volatilità.
- 5) PFE: i prezzi di chiusura in questo caso crescono in modo molto moderato, con un leggero picco attorno al 2022, e con una volatilità molto bassa rispetto agli altri titoli.
- 6) XOM: in questo caso i prezzi di chiusura in generale decrescono fino al primo trimestre del 2020, e successivamente crescono rapidamente, anche se la crescita è caratterizzata da una notevole volatilità.

STATISTICHE DESCRIPTTIVE

Per prima cosa ho calcolato il rendimento cumulato e composto annuo per ogni titolo nel periodo. Per calcolare i rendimenti cumulati abbiamo bisogno dei rendimenti giornalieri, ovvero il rendimento percentuale di un titolo in un singolo giorno. Ora calcoliamo i rendimenti cumulati, che rappresentano la performance complessiva di ciascun titolo nel tempo, partendo dal presupposto che l'investimento inizia con un valore iniziale di 1. Ogni

giorno, il rendimento giornaliero viene aggiunto a 1, quindi il prodotto cumulativo di questi fattori moltiplicativi per tutti i giorni precedenti da il rendimento cumulativo fino a quel punto. Sottraendo 1 alla fine del calcolo assicuriamo che il risultato finale rappresenti il rendimento totale aggiustato per il periodo considerato.

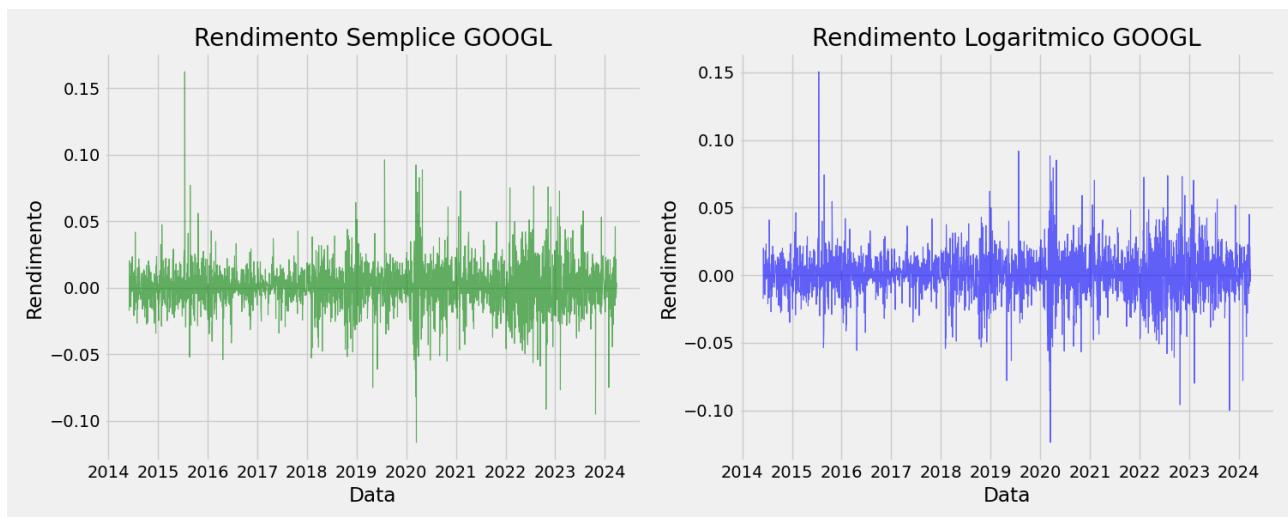


Dal grafico, vediamo come quasi tutti i titoli sono caratterizzati da una crescita o una decrescita dei rendimenti cumulativi moderata, mentre l'unico titolo con una crescita notevole è GOOGL, seppur contraddistinto da una alta volatilità.

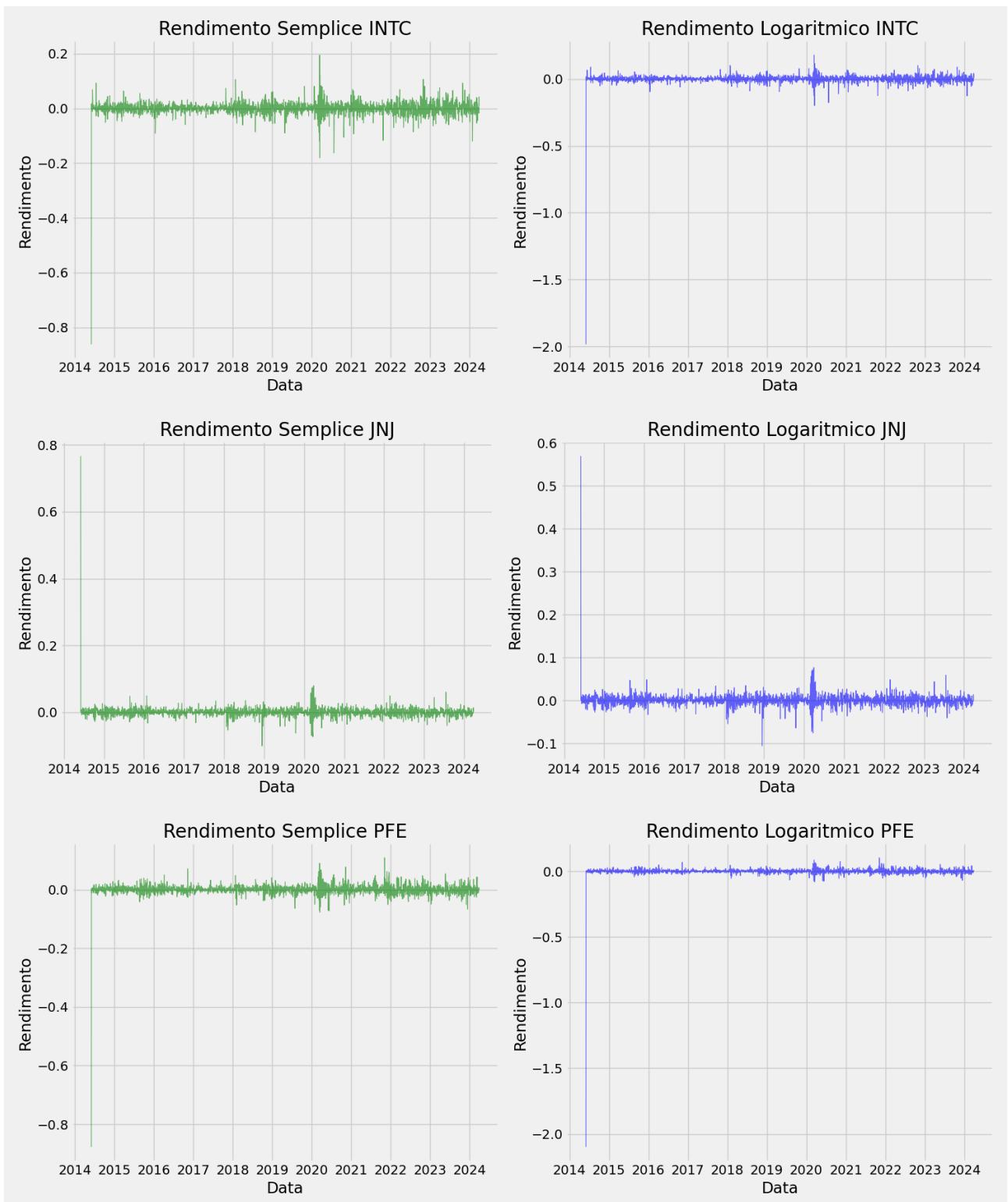
Ora calcoliamo i rendimenti composti annui. Per fare ciò utilizziamo il CARG (Compound Annual Growth Rate), una misura che esprime il tasso di crescita medio annuo di un investimento nel tempo, considerando che il tasso di ritorno è composto annualmente.

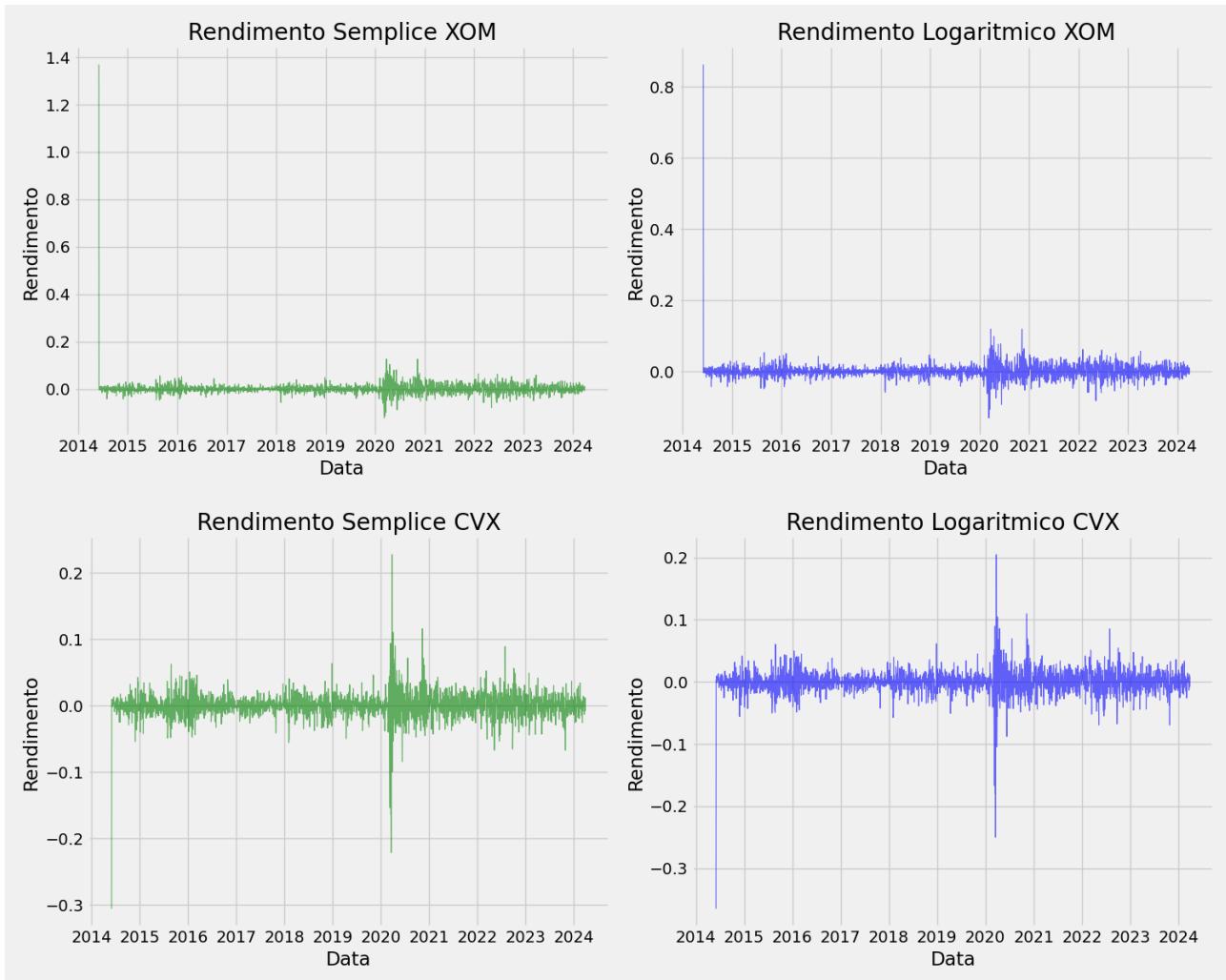
Ticker	Date	Adj Close	Cumulative_Return
GOOGL	2014-06-02	28.216999	NaN
	2014-06-03	27.725500	-0.017419
	2014-06-04	27.688000	-0.018748
	2014-06-05	28.246500	0.001046
	2014-06-06	28.301500	0.002995
Ticker			
CVX		0.070333	
GOOGL		0.185914	
INTC		0.079289	
JNJ		0.074078	
PFE		0.036394	
XOM		0.060389	
Name:	CAGR	dtype: float64	

Tutti i titoli sono caratterizzati da un CAGR positivo, il che vuol dire che in tutti i casi il valore dell'investimento è cresciuto nel tempo a un tasso medio composto annuo positivo. Possiamo vedere come la crescita maggiore è quella di GOOGL, con un tasso del 18,59%, mentre quello minore è quella di PFE, con un tasso del 3,63%. Proseguiamo ora con i rendimenti semplici e logaritmici. I rendimenti semplici contengono le variazioni percentuali giornaliere del prezzo di chiusura aggiustato per ciascun titolo.



Nel caso di GOOGL vediamo come i rendimenti si distribuiscono attorno alla media, con un picco negativo nel 2020 e uno positivo tra il 2015 e il 2016. Nel 2016 abbiamo un aumento dei rendimenti dovuto all'aumento di ricerche tramite device mobile e le pianificazioni video su YouTube.

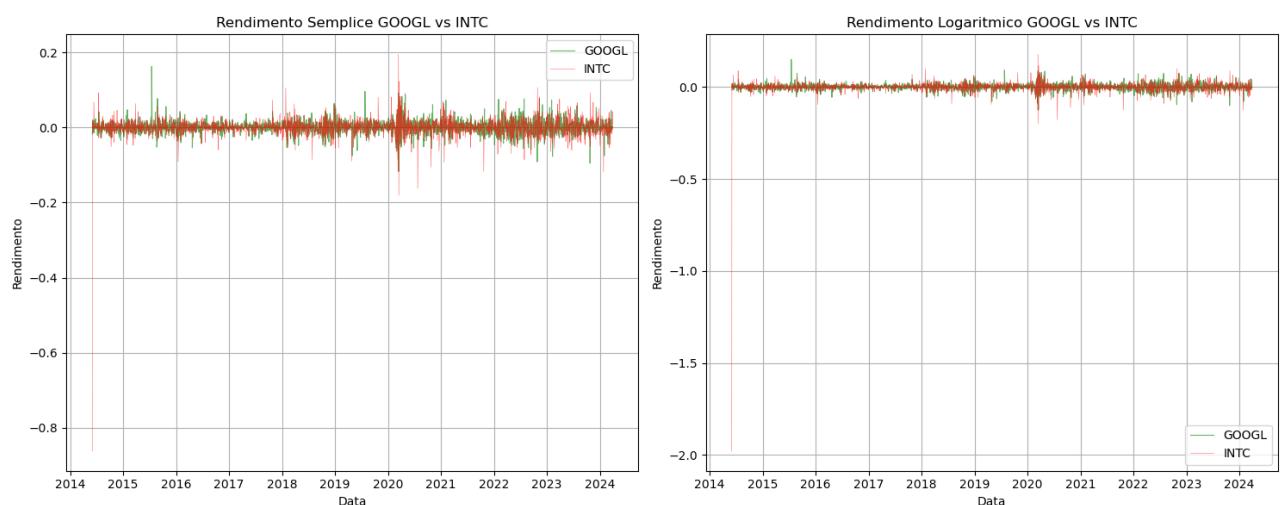




In generale, vediamo come per tutti i titoli restanti il 2014 sia un anno particolare. Infatti per INTL, PFE e CVX vediamo un grande picco negativo iniziale, mentre per XOM e JNJ vediamo un notevole picco positivo.

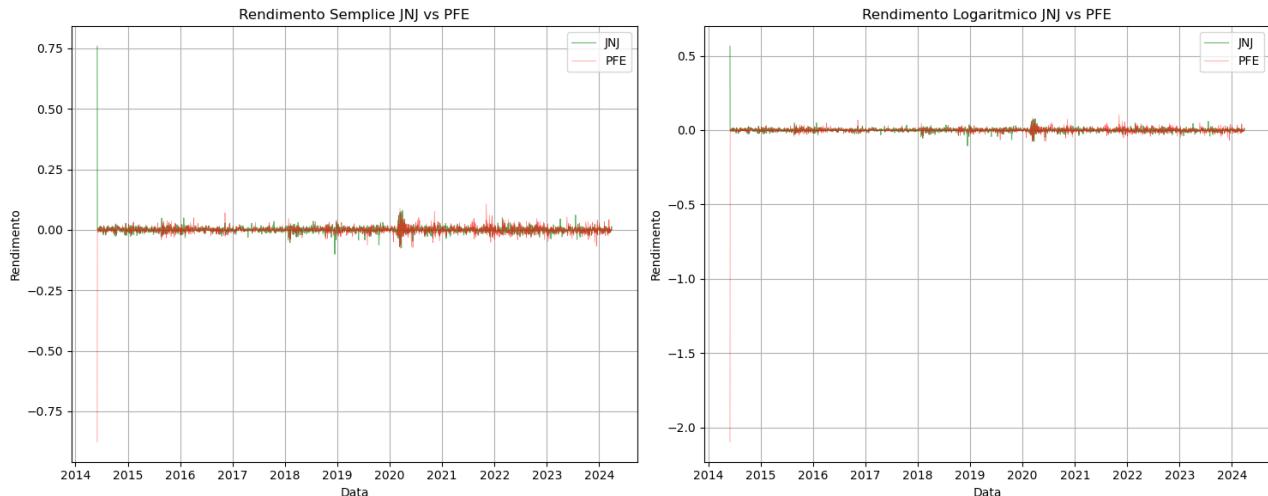
Vediamo ora questi rendimenti confrontando i titoli dei settori analizzati.

Settore tecnologia dell'informazione:



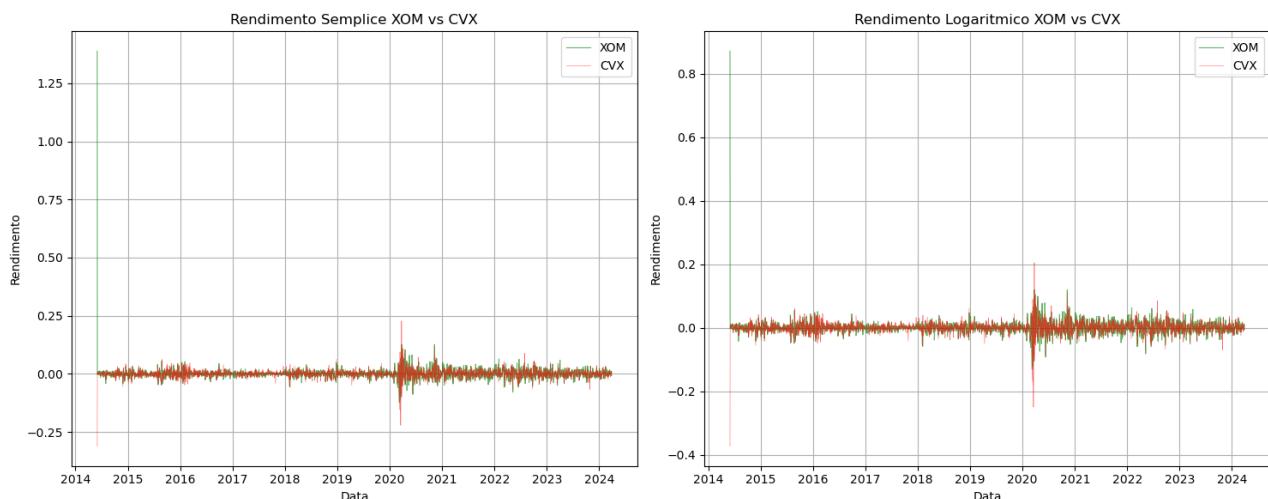
In questo caso i suoi titoli sembrano abbastanza correlati nel loro andamento, evidenziamo però il picco negativo che caratterizza il 2014 per INTC e un picco positivo di GOOGL tra il 2015 e il 2016 non condiviso da INTC.

Settore salute:



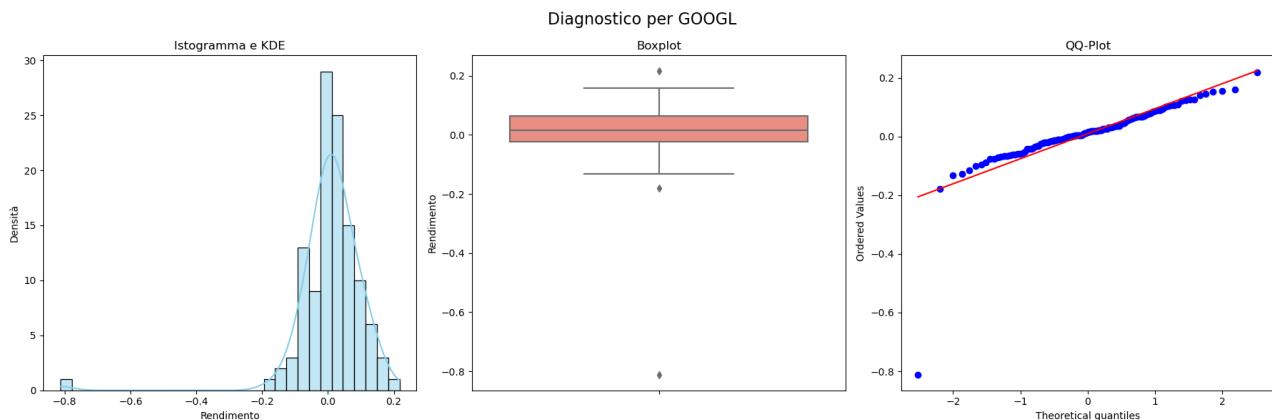
Anche qui troviamo i due titoli abbastanza coerenti nei rendimenti, seppur sia costante quest'anomalia nel 2014 che vede un picco positivo per JNJ e uno positivo per PFE. Inoltre, evidenziamo come in entrambi i rendimenti delle case farmaceutiche dopo il 2020 aumenti la volatilità, probabilmente per via dell'emergenza COVID e del ruolo che queste due multinazionali hanno giocato in questo periodo.

Settore energetico:

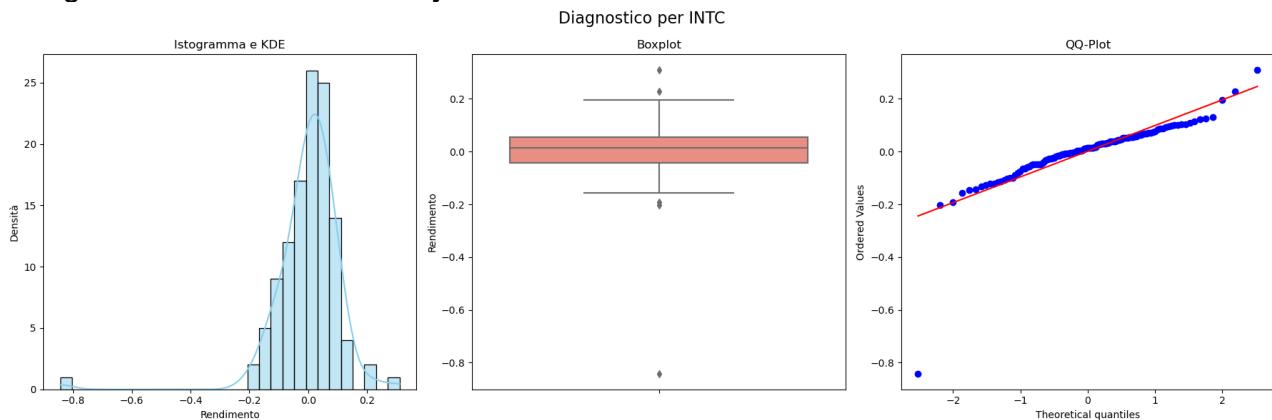


Infine, qui vediamo sempre una certa correlazione tra i rendimenti dei titoli, nonostante l'anomalo picco iniziale in direzioni opposte. Vediamo come subito dopo l'inizio del 2020 vi sia un picco, sia negativo che positivo, per tutte e due le multinazionali, ma in particolare per la CVX, e successivamente un aumento di volatilità per entrambi i titoli.

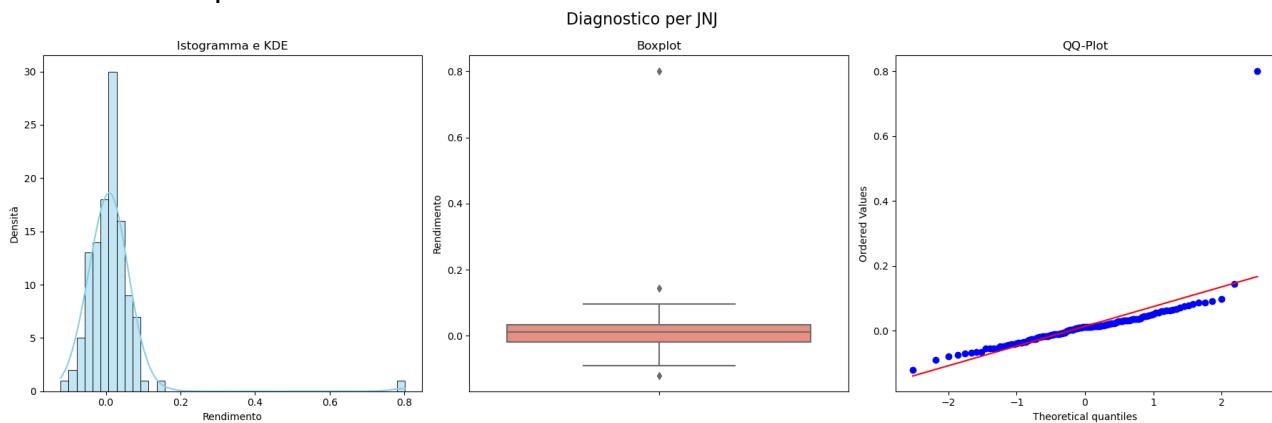
Ora creo dei grafici diagnostici: un istogramma con kernel density, un boxplot e un qq-plot.



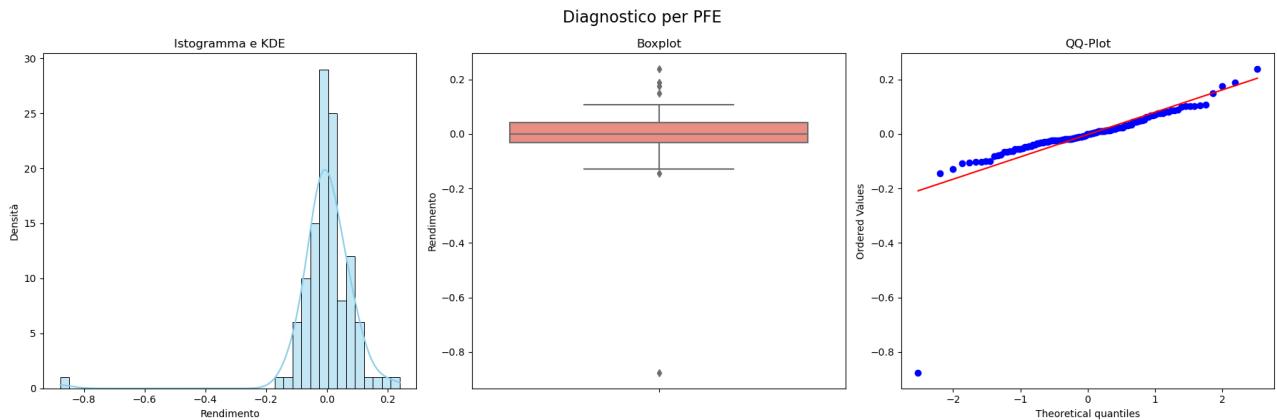
Dall'istogramma possiamo vedere come i rendimenti di GOOGL si distribuiscono bene attorno alla media, con una leggera asimmetria riguardante i rendimenti negativi, mentre dal boxplot vediamo come siano presenti outlier, in particolare uno negativo molto lontano dalla media, e in fine dal qq-plot vediamo come i dati si distribuiscono normalmente, essendo posti all'incirca in linea retta, cosa confermata anche dal confronto tra l'istogramma e la kernel density.



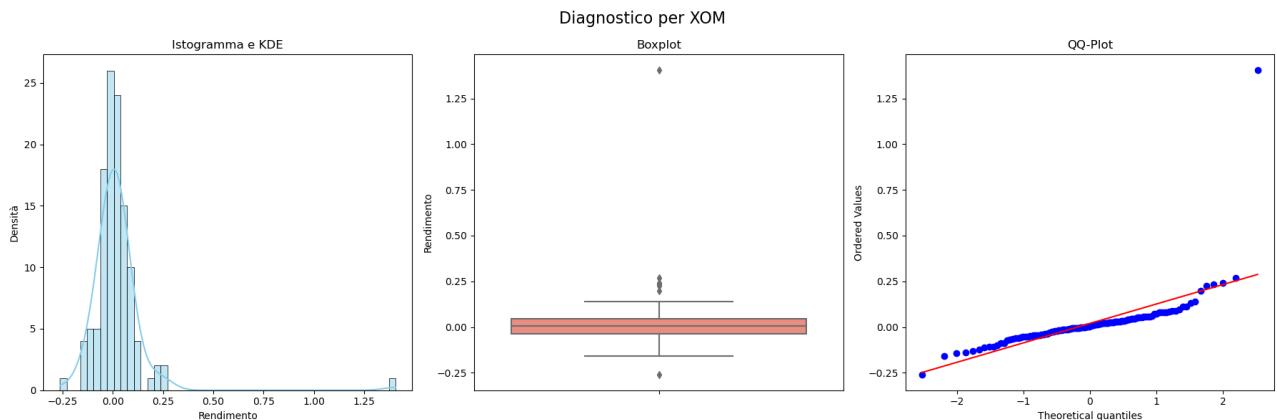
Anche in questo caso dall'istogramma si deduce che tutti i rendimenti si distribuiscono attorno alla media, e anche in questo caso è presente un'asimmetria nella parte negativa dei rendimenti, rappresentata dalla barra isolata a sinistra. Dal boxplot invece deduciamo la presenza di outlier, in particolare di un rendimento particolarmente negativo. In questo caso i dati si adattano meno alla retta del qq-plot rispetto a quelli di GOOGL, indicandoci che si adattino più con difficoltà a una distribuzione normale.



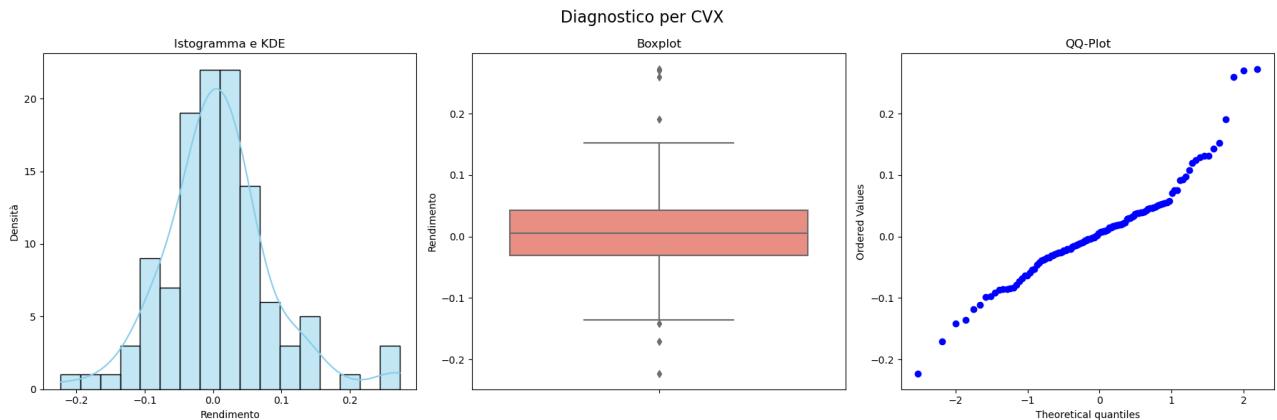
Nel caso di JNJ, vediamo un istogramma caratterizzato da un'asimmetria a destra, cosa confermata dal boxplot, che evidenzia un'outlier particolarmente positivo. I dati, inoltre, si adattano abbastanza bene alla retta del qq-plot, quindi possiamo affermare che i rendimenti della JNJ seguono all'incirca una distribuzione normale.



In questo caso, l'istogramma presenta un'asimmetria a sinistra, confermata dall'outlier particolarmente negativo presente nel boxplot. Per quanto riguarda la normalità, possiamo dire dal qq-plot che i rendimenti di PFE si distribuiscono in modo meno normale rispetto al concorrente JNJ.



Nel caso di XOM, abbiamo un istogramma con un'asimmetria a sinistra, che identifico nell'outlier positivo nel boxplot. Per quanto riguarda il qq-plot, i dati sembrano distribuiti in modo abbastanza normale.



Qui l'istogramma non presenta asimmetrie pronunciate come negli altri casi, e tutti i rendimenti si distribuiscono attorno alla media. Nel boxplot gli outlier sono distribuiti nello stesso modo sia positivamente che negativamente. Per quanto riguarda la normalità, nel qq-plot i dati si distribuiscono intorno alla retta anche se in modo meno omogeneo rispetto agli altri casi.

Calcolo ora le statistiche descrittive per quanto riguarda i rendimenti.

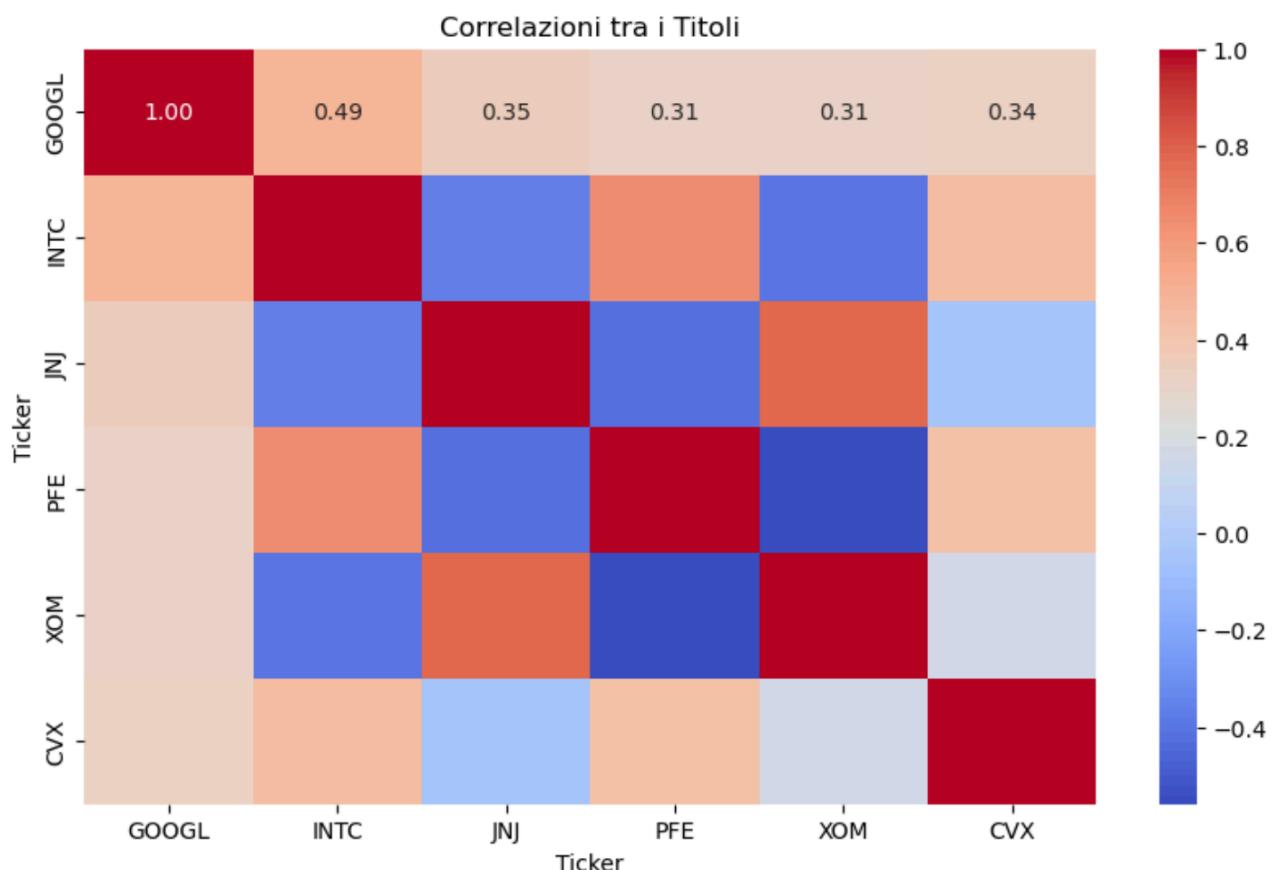
	Mean	Std	Max	Min	Asimm
Ticker					
GOOGL	0.000834	0.017649	0.162584	-0.116341	0.205959
INTC	0.000175	0.027188	0.195213	-0.862373	-12.978078
JNJ	0.000658	0.019131	0.766442	-0.100379	25.879943
PFE	-0.000110	0.022724	0.108552	-0.877425	-23.183066
XOM	0.000940	0.032627	1.369436	-0.122248	29.842500
CVX	0.000323	0.019730	0.227407	-0.305878	-1.661859

Il titolo con il rendimento più basso è PFE, mentre quello con il rendimento più alto è JNJ. Il titolo con la deviazione standard più alta è XOM, mentre quello con la deviazione standard più bassa è GOOGL. Infine, il titolo con asimmetria negativa più alta è PFE, mentre quello con asimmetria positiva più alta è XOM, e quello con asimmetria più bassa è GOOGL.

	Day	Week	Month	Quarter	Year
Ticker					
GOOGL	21.168011	32.726451	-17.681674	-178.078681	-1.550172
INTC	155.325846	-59.816959	32.020450	3.483231	-301.551582
JNJ	29.054152	81.815344	-4.084246	-26.144366	-8.231662
PFE	-206.199421	13.268947	27.519925	4.810056	12.464589
XOM	34.699508	-447.184003	-4.650671	-11.071296	9.330113
CVX	61.001313	182.939848	-7.974338	14.175380	-9.440409

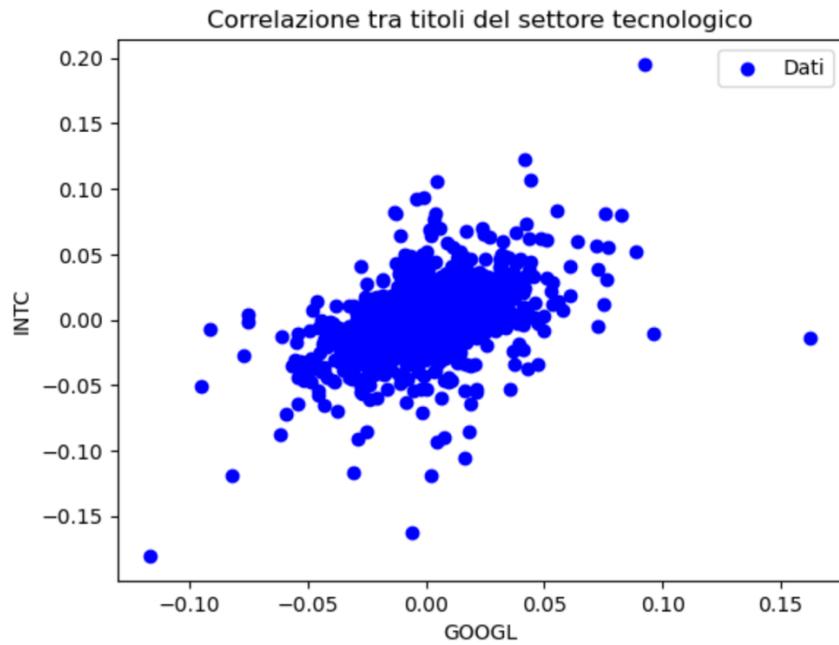
Questa tabella rappresenta come per ogni titolo aumenti la volatilità passando dai rendimenti mensili, fino a quelli annuali. Vediamo per quasi tutti i titoli come la volatilità si stabilizzi passando dai rendimenti giornalieri a quelli annui. Vediamo ora la matrice di correlazione riguardante i rendimenti dei titoli, e li rappresentiamo con un'heatmap.

Ticker	GOOGL	INTC	JNJ	PFE	XOM	CVX
Ticker						
GOOGL	1.000000	0.486688	0.349658	0.307817	0.309032	0.335254
INTC	0.486688	1.000000	-0.360374	0.641665	-0.399935	0.462173
JNJ	0.349658	-0.360374	1.000000	-0.425515	0.785231	-0.047926
PFE	0.307817	0.641665	-0.425515	1.000000	-0.558832	0.428301
XOM	0.309032	-0.399935	0.785231	-0.558832	1.000000	0.154236
CVX	0.335254	0.462173	-0.047926	0.428301	0.154236	1.000000

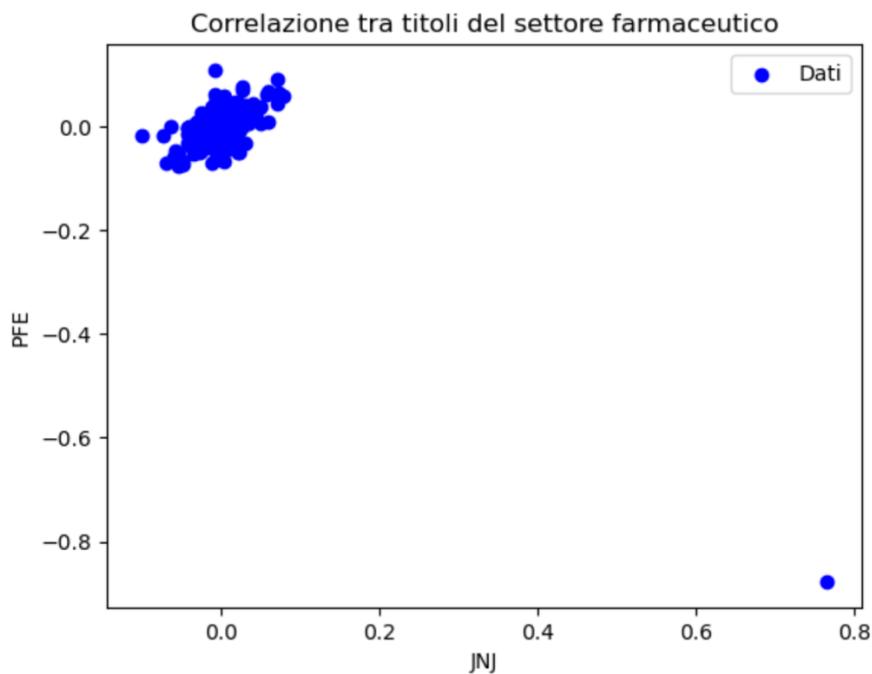


La correlazione positiva più alta è tra XOM e JNJ, mentre la correlazione negativa più bassa è tra PFE e XOM. I titoli meno correlati sono CVX e JNJ.

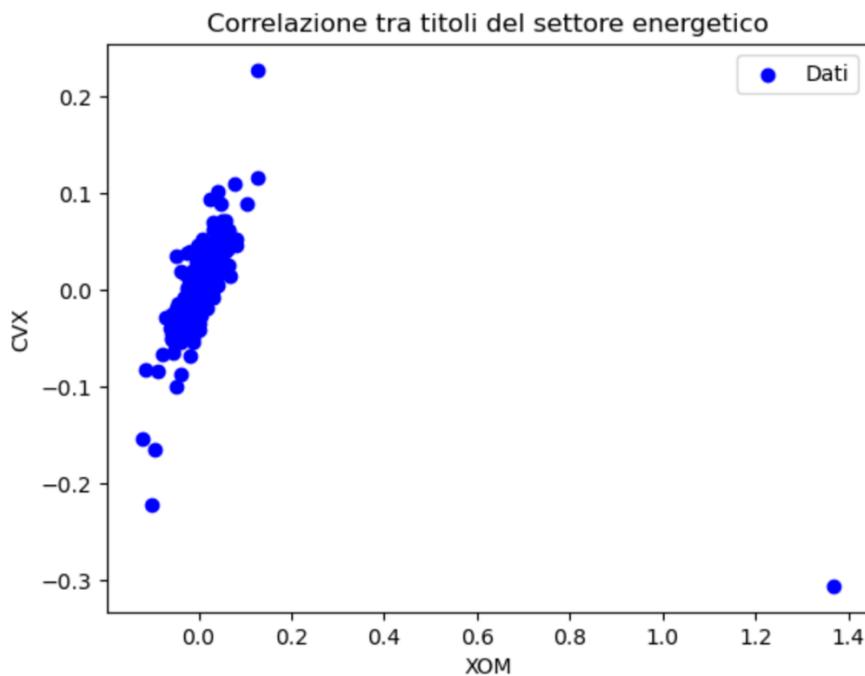
Analizzo ora la correlazione tra i titoli appartenenti allo stesso settore.



Vediamo in questo caso la correlazione tra INTC e GOOGL. Sebbene possiamo notare una correlazione positiva, la dispersione dei punti indica che non è perfetta.



Questo grafico presenta dei dati che non esprimono una particolare correlazione, essendo disposti in modo piuttosto casuale attorno allo zero. Inoltre, l'outlier in basso a destra, potrebbe ulteriormente distorcere la correlazione tra i due titoli.

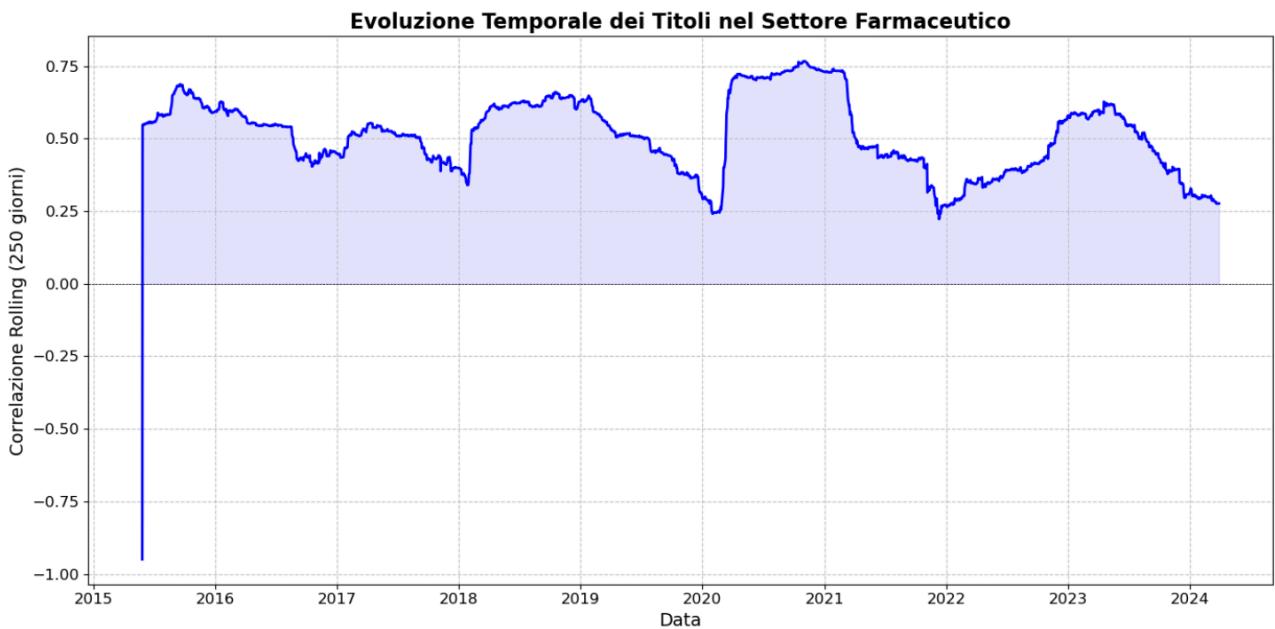


In quest'ultimo caso invece, vediamo come i rendimenti di CVX e XOM tendono a disporsi in una retta con intercetta positiva: questo indica una consistente correlazione positiva, seppur l'outlier in basso a destra potrebbe influire sui risultati.

Vediamo ora come varia l'andamento nel tempo delle correlazioni fra i titoli dello stesso settore:



In questo caso la correlazione tra i titoli del settore tecnologico non è rimasta molto costante, seppur sempre positiva. In particolare, abbiamo la correlazione più alta nel 2021 e più bassa tra il 2015 e il 2016.



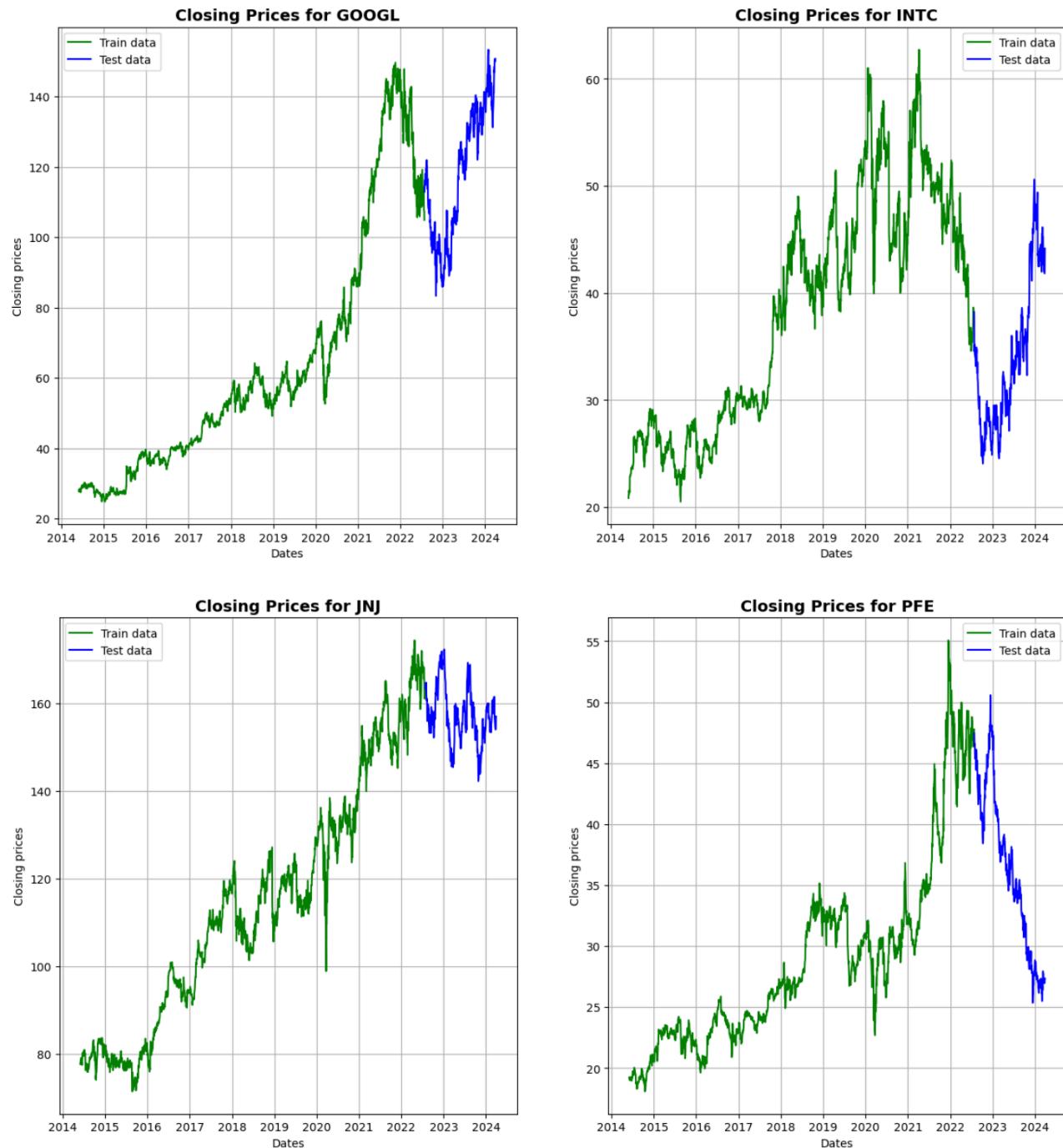
In questo caso, la correlazione dei titoli nel settore farmaceutico si è tenuta più cosate negli anni precedenti al 2020, per poi diventare sempre più variabile. Abbiamo una correlazione sempre positiva, più alta nel 2021, e più bassa nel 2020, ignorando il picco iniziale.



Nel caso del settore farmaceutico, la correlazione è sempre positiva e piuttosto costante con un leggero picco nel 2018, ignorando anche qui il picco iniziale.

ANALISI DI PREVISIONE

Costruisco ora un modello di previsione ARIMA per prevedere i rendimenti di ciascun titolo. Parto dividendo le serie temporali di ogni titolo in test set e training set.





Per tutti i titoli abbiamo utilizzato un modello di addestramento ARIMA(1, 1, 1). Vediamo ora i risultati dell'addestramento per ogni titolo.

GOOGL:

```
SARIMAX Results
=====
Dep. Variable:                      y     No. Observations:                  2474
Model:                 ARIMA(1, 1, 1)   Log Likelihood:           -4536.672
Date:             Tue, 18 Jun 2024   AIC:                         9079.345
Time:                            14:06:12   BIC:                         9096.784
Sample:                           0 - 2474   HQIC:                        9085.680
Covariance Type:                opg
=====
            coef    std err      z   P>|z|      [0.025]     [0.975]
-----
ar.L1      0.8554    0.053   16.187   0.000      0.752      0.959
ma.L1     -0.8883    0.047  -18.898   0.000     -0.980     -0.796
sigma2     2.2958    0.029   79.922   0.000      2.240      2.352
Ljung-Box (L1) (Q):                   0.11   Jarque-Bera (JB):        7902.46
Prob(Q):                           0.74   Prob(JB):                  0.00
Heteroskedasticity (H):               23.29   Skew:                     -0.41
Prob(H) (two-sided):                 0.00   Kurtosis:                  11.72
=====
Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Vediamo che il coefficiente autoregressivo è significativo, indicando una correlazione positiva nel primo lag, e vediamo che anche il coefficiente di media mobile è significativo, indicando una correlazione negativa con il termine d'errore. Il test di Ljung-Box indica che i residui non sono significativamente correlati con il lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale. L'eteroschedasticità indica la presenza di una varianza non costante.

INTC:

SARIMAX Results						
Dep. Variable:	y	No. Observations:	2474			
Model:	ARIMA(1, 1, 1)	Log Likelihood	-3133.049			
Date:	Tue, 18 Jun 2024	AIC	6272.099			
Time:	14:06:34	BIC	6289.538			
Sample:	0 - 2474	HQIC	6278.433			
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
ar.L1	-0.3569	0.067	-5.327	0.000	-0.488	-0.226
ma.L1	0.2105	0.071	2.979	0.003	0.072	0.349
sigma2	0.7378	0.008	94.199	0.000	0.722	0.753
Ljung-Box (L1) (Q):		0.05	Jarque-Bera (JB):	23542.96		
Prob(Q):		0.81	Prob(JB):	0.00		
Heteroskedasticity (H):		6.74	Skew:	-0.99		
Prob(H) (two-sided):		0.00	Kurtosis:	17.99		

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

In questo caso il coefficiente autoregressivo è significativo, indicando una correlazione negativa nel primo lag. Anche il coefficiente di media mobile è significativo, indicando una correlazione positiva con il termine d'errore. Il test Ljung-Box non è significativo, quindi i residui non sono significativamente correlati al lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale, e infine l'eteroschedasticità indica la presenza di varianza non costante.

JNJ:

SARIMAX Results						
Dep. Variable:	y	No. Observations:	2474			
Model:	ARIMA(1, 1, 1)	Log Likelihood	-4341.713			
Date:	Tue, 18 Jun 2024	AIC	8689.426			
Time:	14:06:58	BIC	8706.865			
Sample:	0 - 2474	HQIC	8695.761			
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
ar.L1	-0.3354	0.122	-2.754	0.006	-0.574	-0.097
ma.L1	0.2666	0.125	2.134	0.033	0.022	0.511
sigma2	1.9609	0.024	80.266	0.000	1.913	2.009
Ljung-Box (L1) (Q):		0.01	Jarque-Bera (JB):	8091.90		
Prob(Q):		0.91	Prob(JB):	0.00		
Heteroskedasticity (H):		4.29	Skew:	-0.41		
Prob(H) (two-sided):		0.00	Kurtosis:	11.82		

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Il coefficiente autoregressivo è significativo, indicando una correlazione negativa nel primo lag. Anche il coefficiente di media mobile è significativo, indicando una correlazione positiva con il termine d'errore. Il test Ljung-Box non è significativo, quindi i residui non sono significativamente correlati al lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale. L'eteroschedasticità indica la presenza di varianza non costante.

PFE:

SARIMAX Results						
Dep. Variable:	y	No. Observations:	2474			
Model:	ARIMA(1, 1, 1)	Log Likelihood	-1684.046			
Date:	Tue, 18 Jun 2024	AIC	3374.091			
Time:	14:07:35	BIC	3391.531			
Sample:	0	HQIC	3380.426			
	- 2474					
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
ar.L1	-0.0004	0.540	-0.001	0.999	-1.059	1.058
ma.L1	-0.0217	0.540	-0.040	0.968	-1.080	1.036
sigma2	0.2286	0.003	75.559	0.000	0.223	0.234
Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	5466.16			
Prob(Q):	1.00	Prob(JB):	0.00			
Heteroskedasticity (H):	7.57	Skew:	0.36			
Prob(H) (two-sided):	0.00	Kurtosis:	10.25			

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Il coefficiente autoregressivo non è significativo, indicando che questo termine non contribuisce significativamente alla previsione. Anche coefficiente di media mobile non è significativo. Il test Ljung-Box non è significativo, quindi i residui non sono significativamente correlati al lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale. L'eteroschedasticità indica la presenza di varianza non costante.

XOM:

SARIMAX Results						
Dep. Variable:	y	No. Observations:	2474			
Model:	ARIMA(1, 1, 1)	Log Likelihood	-3711.141			
Date:	Tue, 18 Jun 2024	AIC	7428.282			
Time:	14:08:08	BIC	7445.722			
Sample:	0	HQIC	7434.617			
	- 2474					
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
ar.L1	0.8129	0.228	3.561	0.000	0.366	1.260
ma.L1	-0.8243	0.222	-3.707	0.000	-1.260	-0.389
sigma2	1.1776	0.020	59.484	0.000	1.139	1.216
Ljung-Box (L1) (Q):	1.13	Jarque-Bera (JB):	1506.00			
Prob(Q):	0.29	Prob(JB):	0.00			
Heteroskedasticity (H):	5.04	Skew:	-0.14			
Prob(H) (two-sided):	0.00	Kurtosis:	6.81			

Warnings:

[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Il coefficiente autoregressivo è significativo, indicando una correlazione positiva nel primo lag. Anche il coefficiente di media mobile è significativo, indicando una correlazione negativa con il termine d'errore. Il test Ljung-Box non è significativo, quindi i residui non sono significativamente correlati al lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale, e infine l'eteroschedasticità indica la presenza di varianza non costante.

CVX:

SARIMAX Results						
Dep. Variable:	y	No. Observations:	2474			
Model:	ARIMA(1, 1, 1)	Log Likelihood	-4879.704			
Date:	Tue, 18 Jun 2024	AIC	9765.409			
Time:	14:08:43	BIC	9782.849			
Sample:	0 - 2474	HQIC	9771.744			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.5077	0.217	-2.341	0.019	-0.933	-0.083
ma.L1	0.4752	0.221	2.147	0.032	0.041	0.909
sigma2	3.0299	0.041	74.013	0.000	2.950	3.110
Ljung-Box (L1) (Q):		0.06	Jarque-Bera (JB):		5858.00	
Prob(Q):		0.81	Prob(JB):		0.00	
Heteroskedasticity (H):		5.68	Skew:		-0.46	
Prob(H) (two-sided):		0.00	Kurtosis:		10.48	
=====						
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

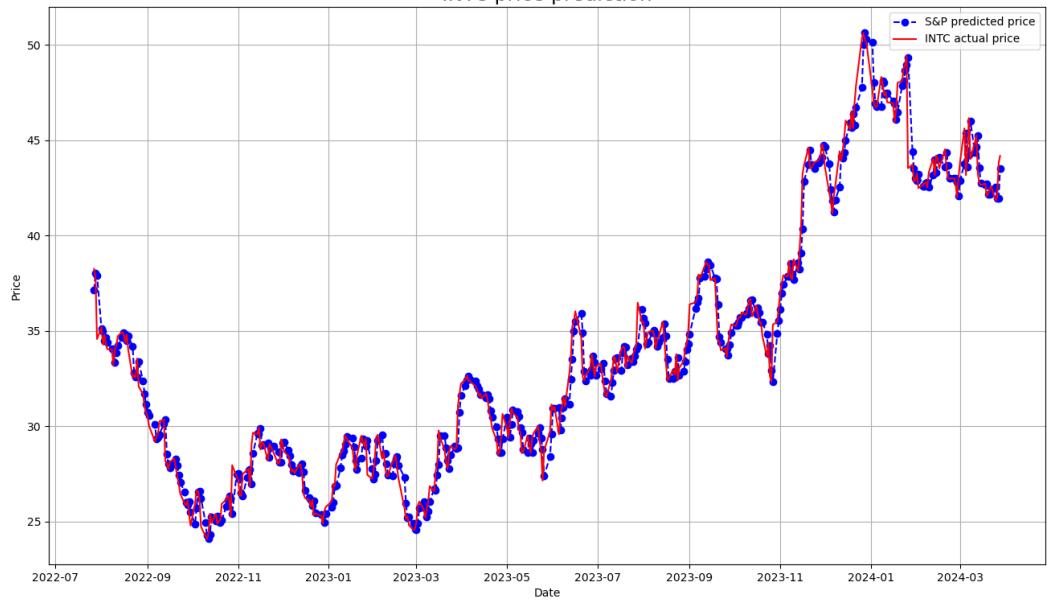
Infine, anche in questo caso il coefficiente di autoregressione è significativo, indicando una correlazione negativa nel primo lag, e anche il coefficiente di media mobile è significativo, indicando una correlazione positiva con i termini di errore. Il test Ljung-Box non è significativo, quindi i residui non sono significativamente correlati al lag 1. Il test Jarque-Bera suggerisce che i residui non seguono una distribuzione normale. L'eteroschedasticità indica la presenza di varianza non costante.

Rappresentiamo ora un confronto tra i prezzi predetti dal modello e i prezzi attuali per ogni titolo.

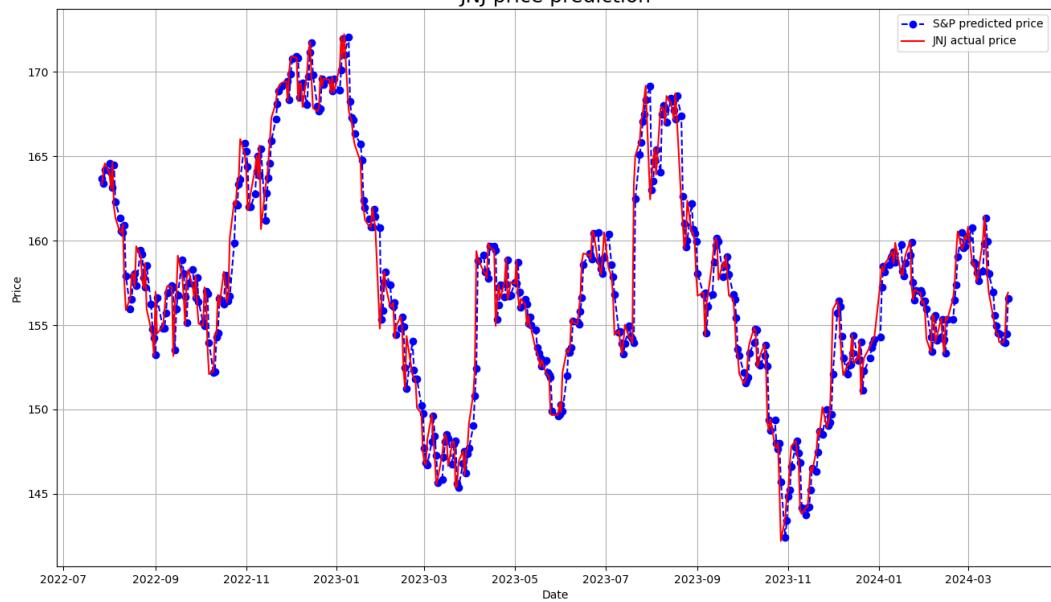
GOOGL price prediction



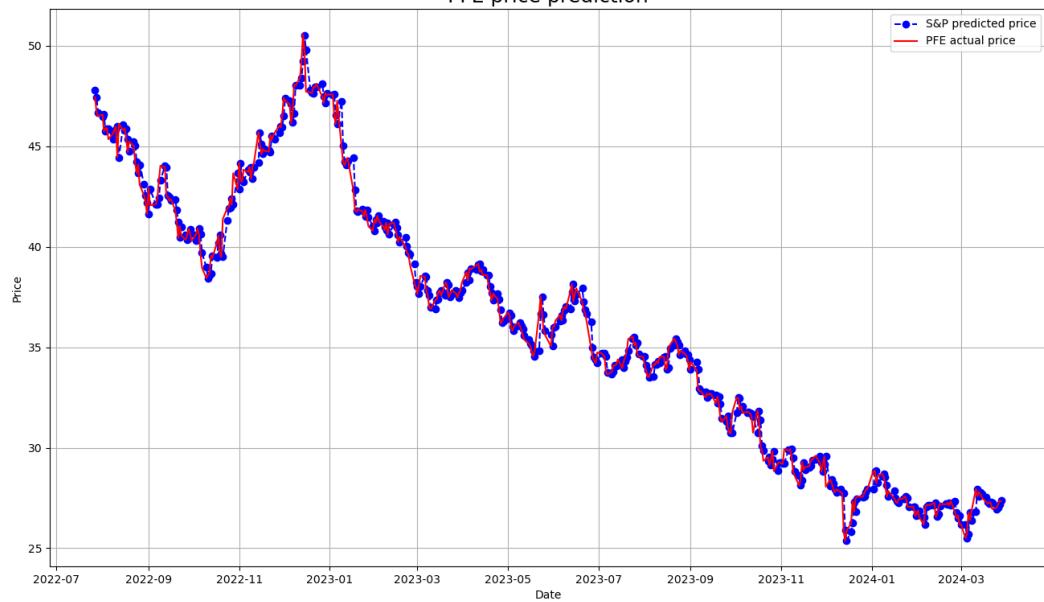
INTC price prediction

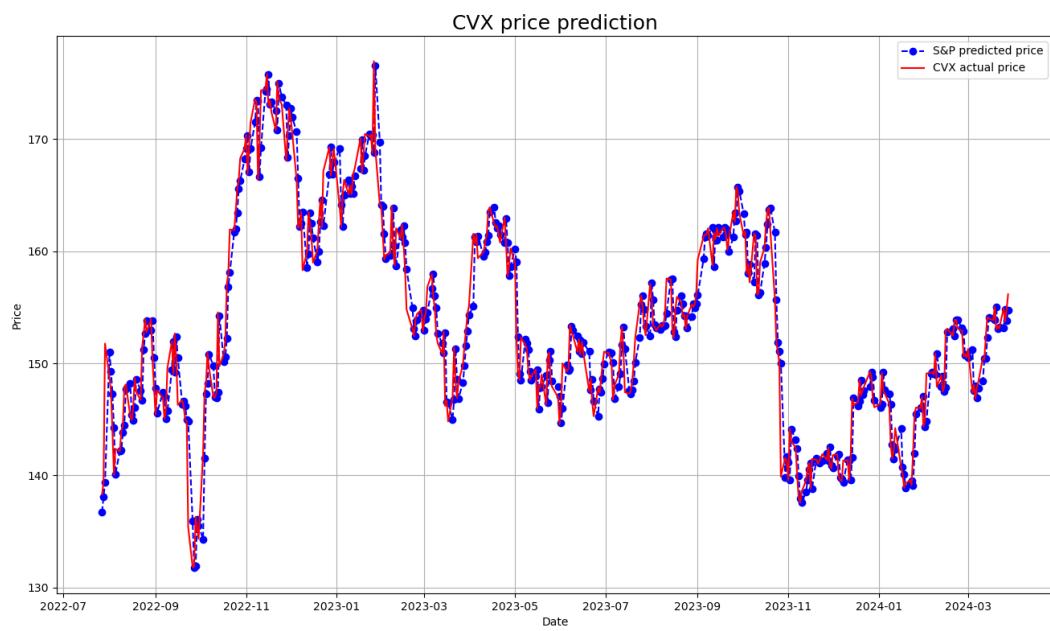
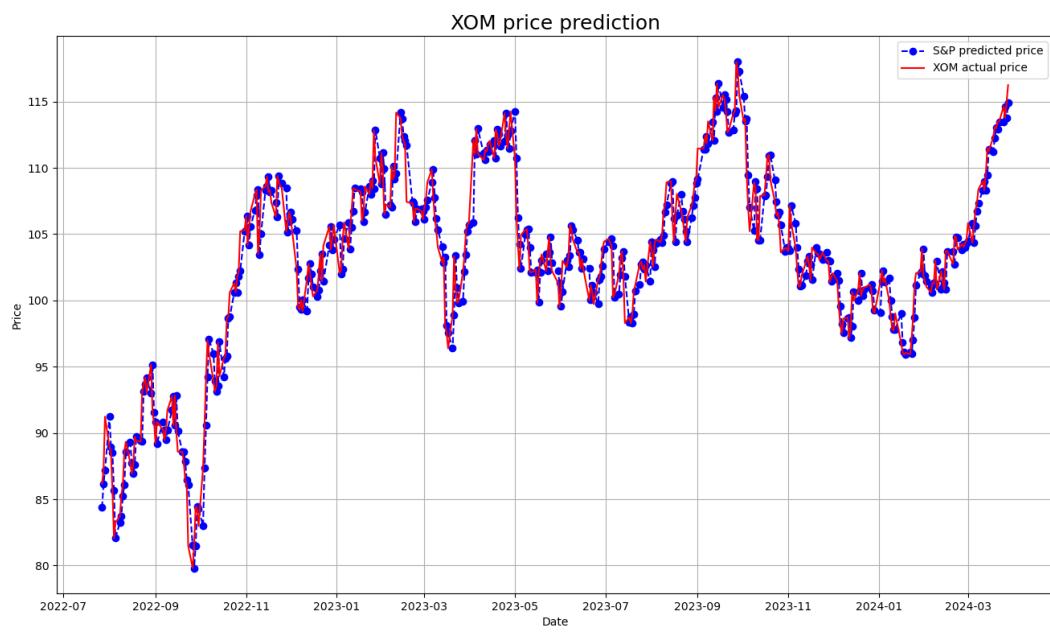


JNJ price prediction



PFE price prediction

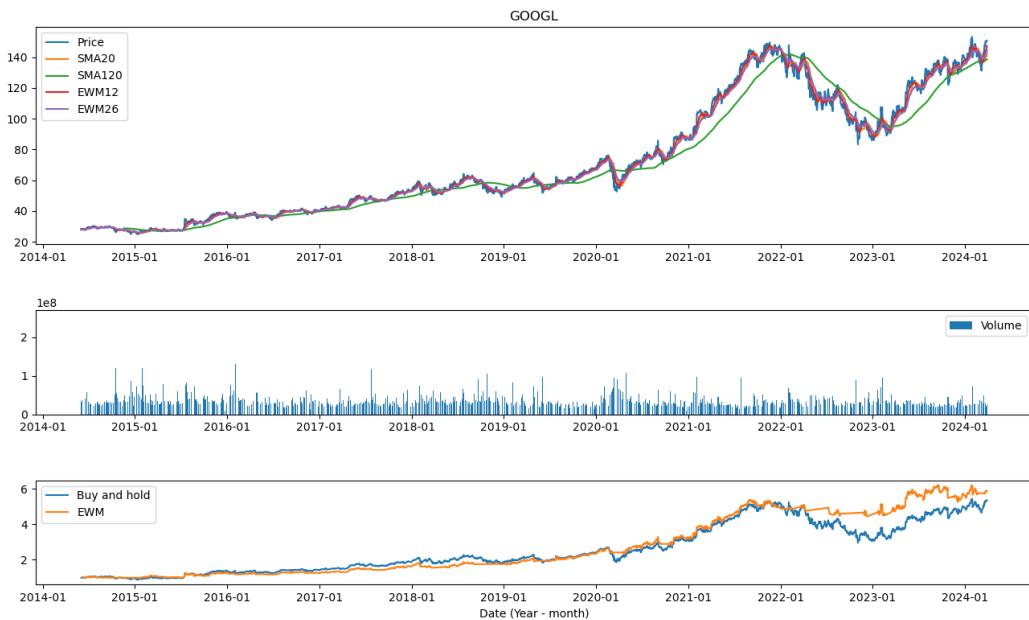




In generale, il modello scelto si è rivelato adatto alla previsione dei rendimenti, essendo quasi tutti i coefficienti ottenuti significativi.

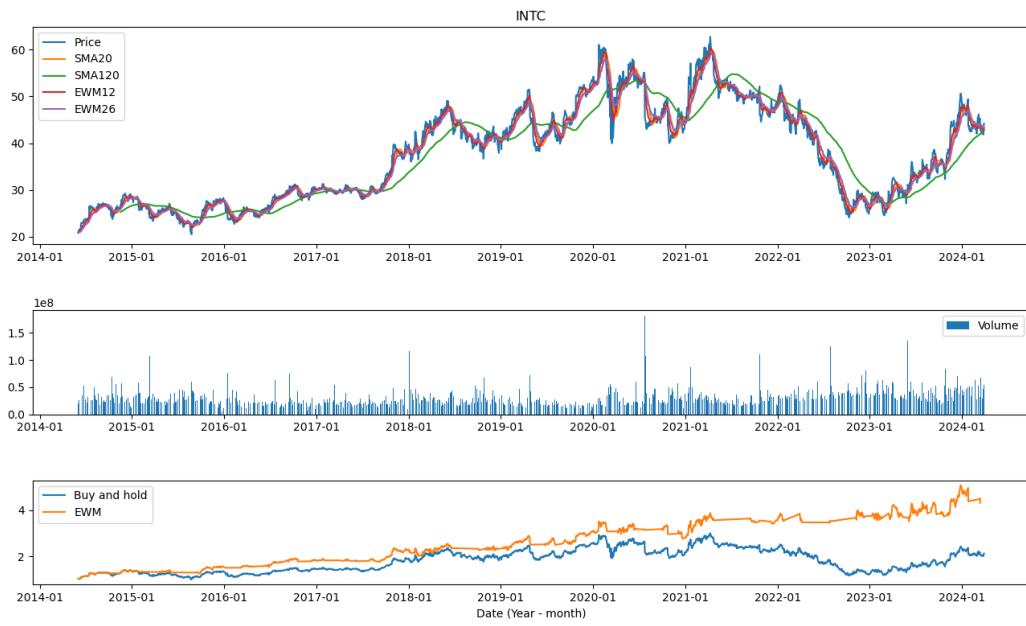
STRATEGIE DI TRADING E BACKTESTING

Come strategia di trading ho utilizzato la MACD (Moving Averages Convergence Divergence). Ho creato per prima cosa ho creato le strategie con la MACD e con il “Buy and Hold” e le ho confrontate in un grafico per ogni titolo.

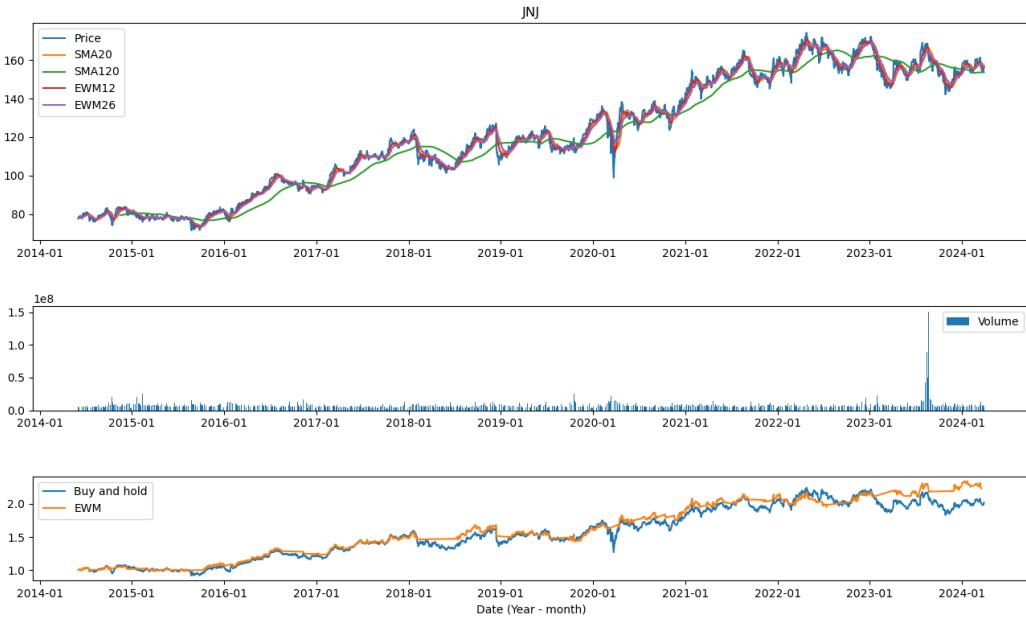


Il grafico in alto contiene il prezzo delle azioni di GOOGL (linea blu), la media mobile semplice a 20 giorni (linea rossa), la media mobile semplice a 120 giorni (linea verde), la media mobile esponenziale a 12 giorni (linea arancione), e la media mobile esponenziale a 26 giorni (linea magenta). Vediamo come tutte le strategie sono più o meno sulla stessa linea.

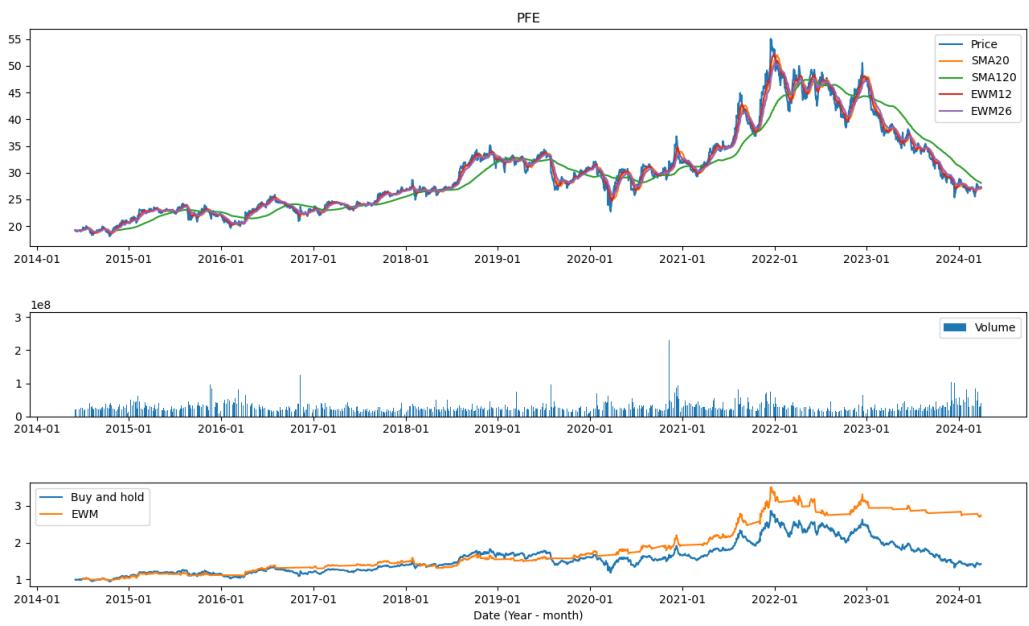
Il grafico in mezzo rappresenta il volume degli scambi giornalieri delle azioni di GOOGL. Infine, l'ultimo grafico confronta la strategia di investimento basata sulle media mobile esponenziale (linea arancione) e la strategia del “Buy and Hold” (linea blu). Vediamo come le due strategie hanno più o meno lo stesso andamento fino al 2022, dove la strategia EWM inizia ad avere rendimenti più alti rispetto all'altra.



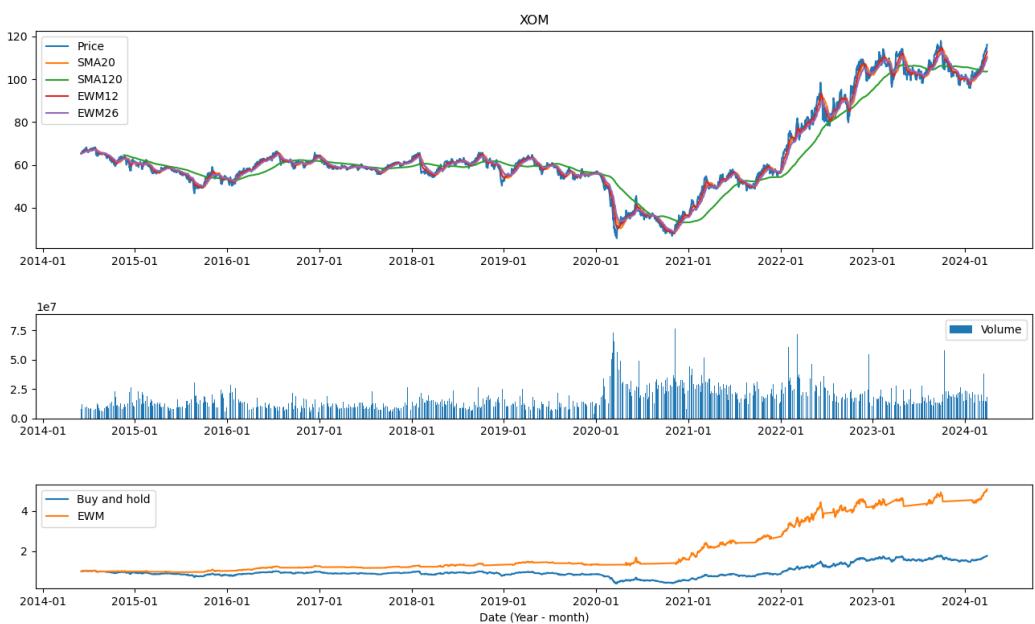
Nel caso di INTC vediamo come nel primo grafico tutte le strategie seguono più o meno lo stesso andamento, mentre nell'ultimo la strategia EWM ha sempre un rendimento più alto rispetto alla strategia del “Buy and Hold”, seppur all'inizio in modo moderato, dopo il 2021 in modo più evidente.



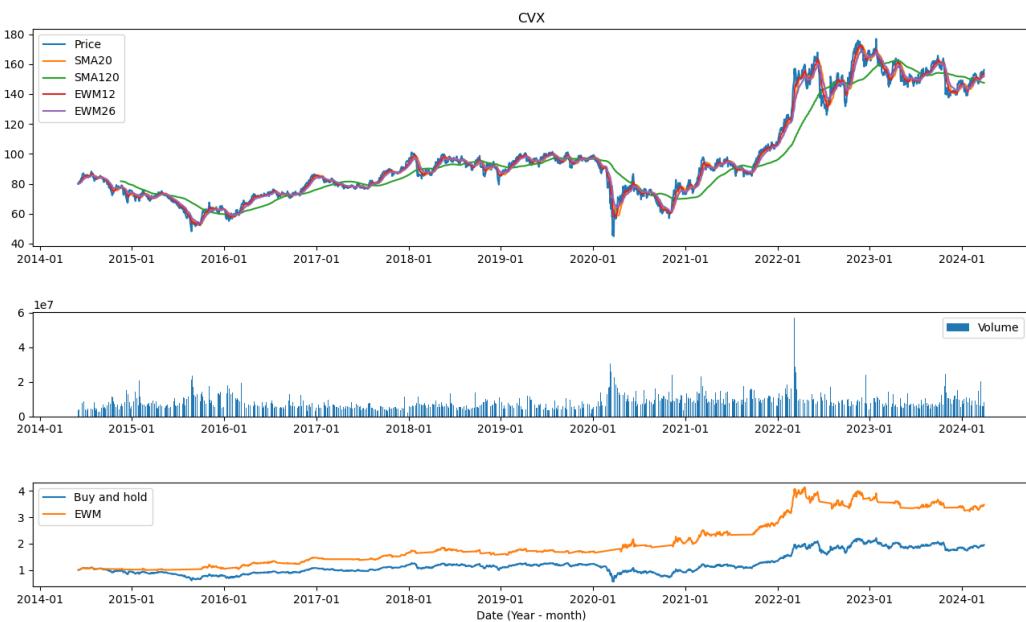
Anche qui, nel caso di JNJ, nel primo grafico l'andamento dei rendimenti è similare per tutte le strategie, mentre nell'ultimo le strategie dell'EWM e del “Buy and Hold” hanno più o meno gli stessi rendimenti, con un leggero discostamento nel 2023.



Nel caso di PFE, vediamo sempre una certa coerenza tra rendimenti delle strategie nel primo grafico, mentre nell'ultimo vediamo dei rendimenti superiori della strategia di EWM rispetto a quella del “Buy and Hold” dopo l'anno 2020.



Anche nel caso di XOM vediamo come i rendimenti delle strategie nel primo grafico seguono lo stesso pattern dei prezzi, mentre nell'ultimo grafico abbiamo un netto aumento dei rendimenti della strategia EMW rispetto alla “Buy and Hold” a partire dal 2020.



Infine, per il titolo CVX vediamo come nel primo grafico tutti rendimenti delle strategie rispecchiano i prezzi delle azioni, mentre nell'ultimo grafico vi è sempre un rendimento superiore da parte della strategia EWM, seppur all'inizio moderato, via via sempre più consistente.

Vediamo ora i risultati di ogni strategia per ogni titolo.

```

Buy and hold strategy return: GOOGL 5.34890306326272
SMA return: GOOGL 3.512923290644119
EWM return: GOOGL 5.888997253617071
Buy and hold strategy return: INTC 2.1178117564936705
SMA return: INTC 1.2280672800190044
EWM return: INTC 4.320711954137239
Buy and hold strategy return: JNJ 2.019371138513618
SMA return: JNJ 1.0477431613032393
EWM return: JNJ 2.2284638603024485
Buy and hold strategy return: PFE 1.4212713722347354
SMA return: PFE 0.8674323418390258
EWM return: PFE 2.7358539427379647
Buy and hold strategy return: XOM 1.780061591112362
SMA return: XOM 1.853469340187038
EWM return: XOM 5.09167905303656
Buy and hold strategy return: CVX 1.9511926534840873
SMA return: CVX 1.060434199118035
EWM return: CVX 3.4777939192452942

```

In generale, per ogni titolo la EWM ha mostrato i rendimenti migliori, suggerendo che una media mobile esponenziale potrebbe essere più efficace nel catturare i movimenti rispetto alle altre strategie. La SMA ha prodotto i rendimenti più bassi.

CAPM

Nella prima parte calcolo il beta di ciascun titolo rispetto al mercato.

{ 'GOOGL': 1.030312456718585, 'INTC': 0.9676085632346906, 'JNJ': 0.5751957722949503, 'PFE': 0.6560140266103617, 'XOM': 0.9569651223403852, 'CVX': 1.0901782874935355}

Successivamente utilizzo il beta per calcolare il rendimento atteso annuo del titolo.

Rendimento atteso annuo dei titoli:

GOOGL: 8.18%

INTC: 7.81%

JNJ: 5.45%

PFE: 5.94%

XOM: 7.74%

CVX: 8.54%

Il titolo con il rendimento annuo più alto è GOOGL, mentre quello con il rendimento annuo più basso è JNJ.

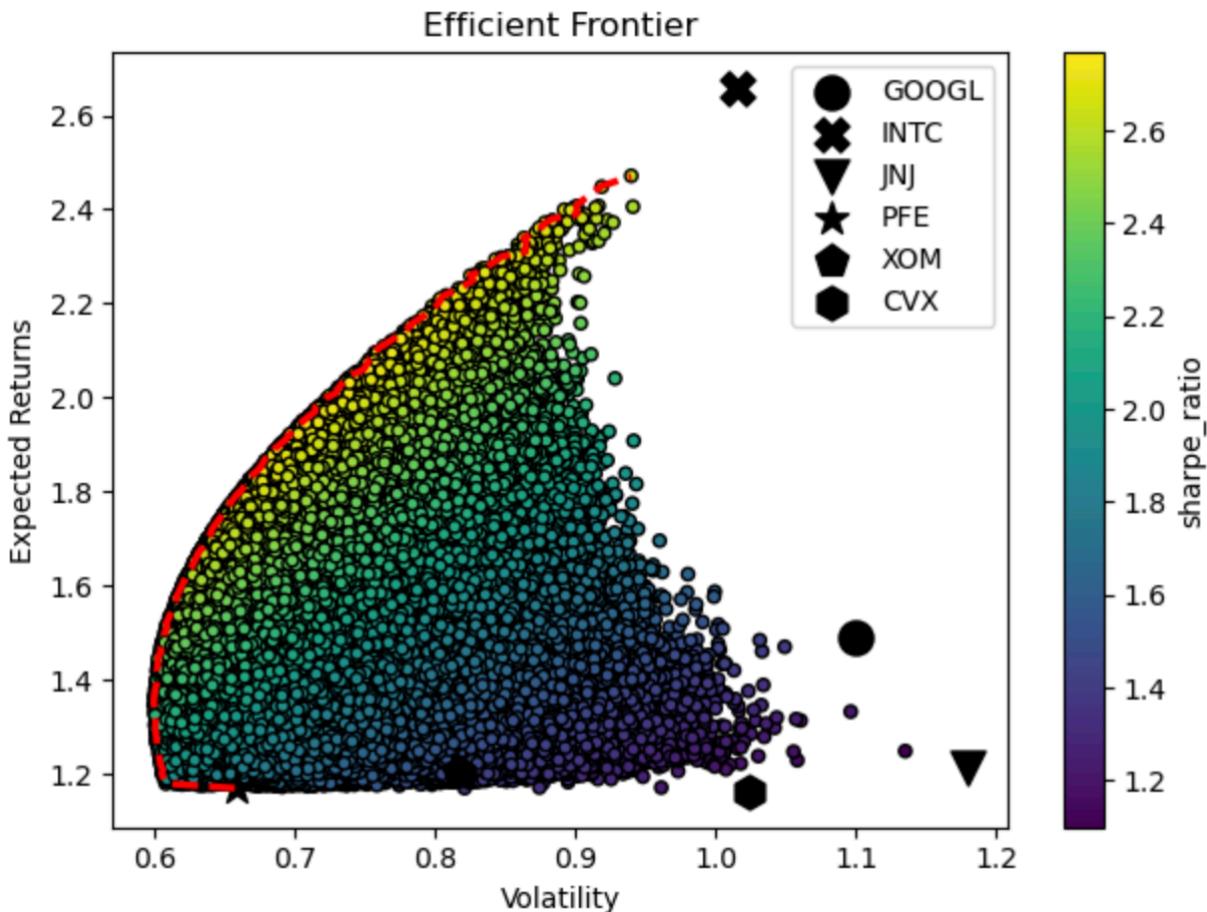
Ora calcolo l'esposizione a rischio di ogni titolo con i fattori di Fama e French.

Dai risultati ottenuti, concludiamo che:

- GOOGL mostra una maggiore esposizione al rischio di mercato rispetto al mercato stesso, essendo il coefficiente di mkt (Market Factor) maggiore di 1 (1.0722)
- INTC ha una minore esposizione al rischio di mercato rispetto al mercato stesso, avendo un Market Factor minore di 1 (0.8983)
- JNJ presenta una minore esposizione al rischio di mercato rispetto al mercato stesso, avendo un Market Factor minore di 1 (0.5890)
- PFE ha una minore esposizione al rischio di mercato rispetto al mercato stesso, avendo un Market Factor minore di 1 (0.6802)
- XOM mostra una minore esposizione al rischio di mercato rispetto al mercato stesso, avendo un Market Factor minore di 1 (0.8971)
- CVX presenta una maggiore esposizione al rischio di mercato dispetto al mercato stesso, essendo il suo Market Factor maggiore di 1 (1.0438).

COSTRUZIONE DI PORTAFOGLIO

Per prima cosa costruisco un portafoglio ottimale in termini di media varianza utilizzando i primi 108 mesi del nostro periodo di tempo totale utilizzando i rendimenti passati.



In questo grafico è rappresentata la frontiera efficiente per un portafoglio di titoli (linea rossa tratteggiata). I punti colorati sul grafico rappresentano diversi portafogli simulati. Il colore dei punti varia in base al rapporto di Sharpe, con una scala cromatica che va dal blu (rapporto di Sharpe basso) al giallo (rapporto di Sharpe alto). I portafogli lungo la frontiera efficiente sono considerati ottimali perché offrono il loro massimo rendimento per un dato livello di rischio.

Ora vediamo il portafoglio con il massimo rapporto di Sharpe con il metodo simulativo. Il portafoglio con il massimo rapporto di Sharpe è progettato per ottenere il massimo rendimento per unità di rischio. Con un rapporto di Sharpe del 276.74%, questo portafoglio offre un rendimento elevato rispetto alla sua volatilità.

Maximum Sharpe Ratio portfolio

Performance

returns: 209.54% volatility: 75.72% sharpe_ratio: 276.74%

Weights

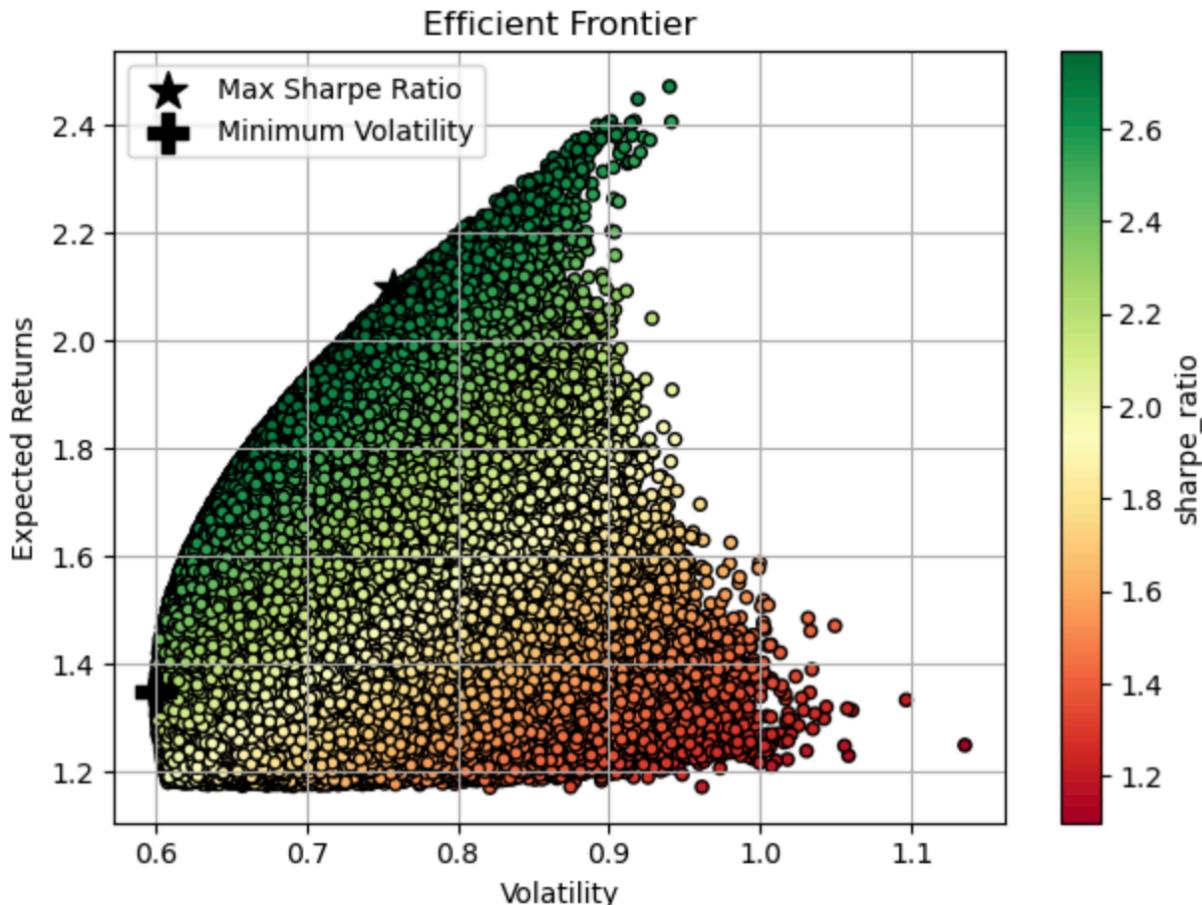
CVX: 2.85% GOOGL: 61.39% INTC: 0.19% JNJ: 25.94% PFE: 8.79% XOM: 0.84%

Ora invece vediamo il portafoglio a minima volatilità con metodo simulativo. Il portafoglio a minima volatilità è progettato per ridurre al minimo il rischio, mantenendo al contempo un rendimento positivo. Con una volatilità del 59.95% e un rapporto di Sharpe del 224.61%, questo portafoglio offre un buon equilibrio tra rischio e rendimento.

```

Minimum Volatility portfolio ----
Performance
returns: 134.66% volatility: 59.95% sharpe_ratio: 224.61%
Weights
CVX: 0.35% GOOGL: 11.49% INTC: 0.31% JNJ: 56.42% PFE: 17.96% XOM: 13.47%

```



Questo grafico rappresenta la frontiera efficiente e il collocamento del portafoglio con il massimo rapporto di Sharpe e del portafoglio con minima volatilità.

Ora vediamo il portafoglio con il massimo rapporto di Sharpe con il metodo analitico:

Maximum Sharpe Ratio portfolio ----

Performance

Return: 205.10% Volatility: 73.94% Sharpe Ratio: 271.96%

Weights

CVX: 5.06% GOOGL: 57.92% INTC: 0.00% JNJ: 27.57% PFE: 9.46% XOM: 0.00%

Questo portafoglio ha un'ottima performance, con un ritorno del 205.10% e un Sharpe Ratio del 271.96%. Questi risultati indicano una gestione molto efficace delle risorse, con GOOGL e JNJ che costituiscono la maggior parte degli investimenti. Tuttavia, la volatilità elevata del 73.94% sottolinea un rischio significativo associato a questo portafoglio.

Infine, invece vediamo il portafoglio a minima volatilità con metodo analitico:

Minimum Volatility portfolio ----

Performance

Return: 31.19% Volatility: 102.42% Sharpe Ratio: 30.45%

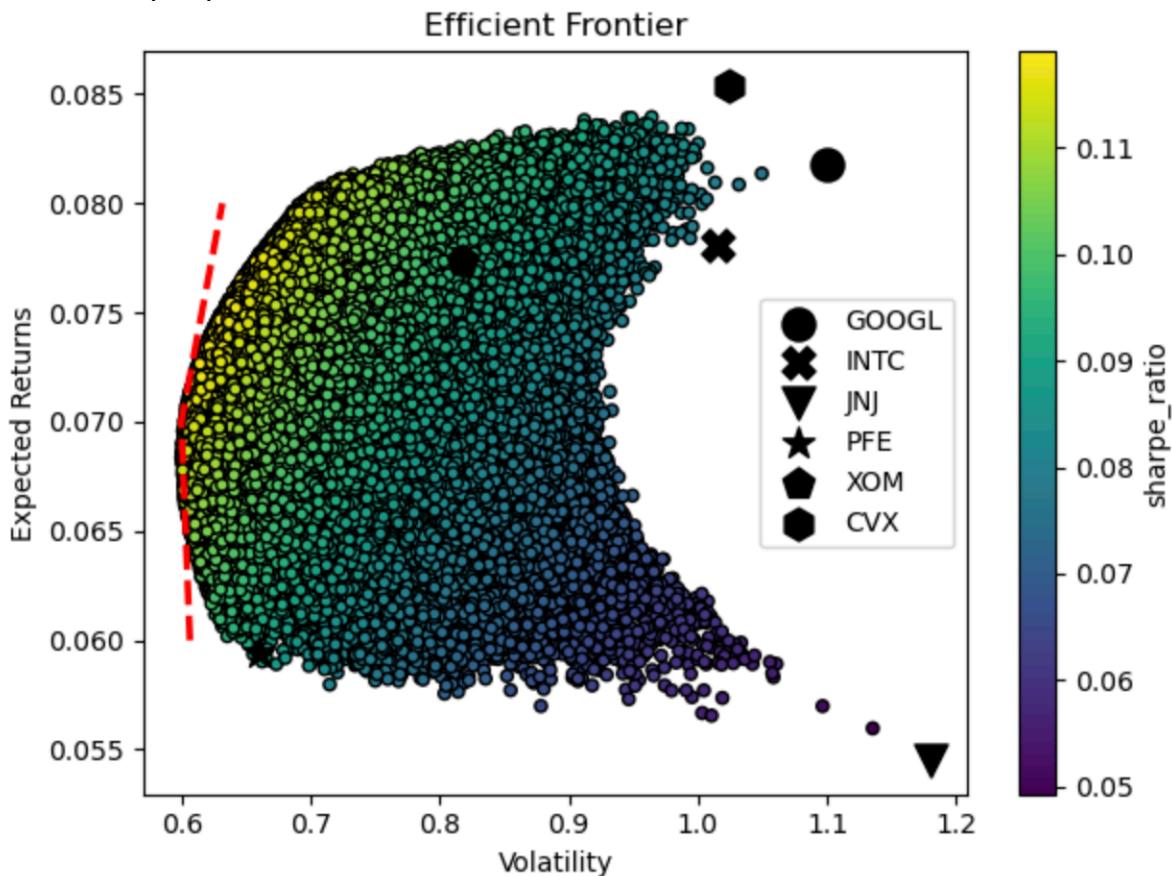
Weights

CVX: 0.00% GOOGL: 0.00% INTC: 0.00% JNJ: 0.00% PFE: 0.00% XOM: 100.00%

Un ritorno del 31.19% è molto alto, il che indica che il portafoglio ha avuto un'ottima performance nel periodo considerato. La volatilità del 102.42% è estremamente alta. Questo livello di volatilità indica che il prezzo delle azioni di XOM è stato molto variabile. Un portafoglio con una volatilità così alta non può essere considerato a bassa volatilità. Un Sharpe Ratio del 30.45% è notevole, suggerendo che il ritorno per unità di rischio è molto alto. Il 100% del portafoglio è investito in Exxon Mobil. Questo approccio non è diversificato e comporta un rischio specifico molto elevato, poiché tutta la performance del portafoglio è legata a un singolo titolo.

Ora costruisco un portafoglio ottimale in termini di media varianza utilizzando i primi 108 mesi del nostro periodo di tempo totale utilizzando i rendimenti attesi.

Costruisco per prima cosa la frontiera efficiente.



Come prima, è rappresentata in questo grafico la frontiera efficiente di un portafoglio di titoli.

Calcoliamo un portafoglio con il massimo rapporto di Sharpe:

Maximum Sharpe Ratio portfolio

Performance

returns: 7.43% volatility: 62.40% sharpe_ratio: 11.90%

Weights

GOOGL: 0.94% INTC: 16.64% JNJ: 0.10% PFE: 28.17% XOM: 31.62% CVX: 22.53%

Questo portafoglio, costruito con metodo simulativo, ha conseguito un ritorno del 7.43%, accompagnato da una volatilità del 62.40% e un Sharpe Ratio del 11.90%. Questi risultati indicano una performance complessivamente buona, con un bilanciamento tra diversi asset. I settori rappresentati suggeriscono una diversificazione settoriale, il che può ridurre il rischio complessivo del portafoglio.

Successivamente calcoliamo un portafoglio con minima volatilità:

Minimum Volatility portfolio ----

Performance

returns: 6.83% volatility: 59.95% sharpe_ratio: 11.40%

Weights

GOOGL: 0.35% INTC: 11.49% JNJ: 0.31% PFE: 56.42% XOM: 17.96% CVX: 13.47%

Questo portafoglio, sempre costruito con metodo simulativo, è progettato per minimizzare la volatilità, con un ritorno del 6.83% e una volatilità del 59.95%. Il Sharpe Ratio di 11.40% indica che il portafoglio offre un buon rendimento in relazione al rischio assunto. La maggior parte degli investimenti è concentrata in PFE (56.42%) e XOM (17.96%), mentre gli altri titoli hanno pesi minori.

Costruiamo ora un portafoglio con il metodo analitico:

Minimum Volatility portfolio ----

Performance

Return: 6.76% Volatility: 59.95% Sharpe Ratio: 11.28%

Weights

GOOGL: 0.00% INTC: 10.49% JNJ: 1.81% PFE: 57.67% XOM: 18.03% CVX: 12.00%

Questo portafoglio con minima volatilità ha un ritorno del 6.76% che possiamo definire moderato. La volatilità è del 59.95%, abbastanza alta, ma minore rispetto a portafogli con rendimenti più alti. Questo indica che il portafoglio ha una certa variabilità nei suoi rendimenti, ma è stato ottimizzato per ridurre le fluttuazioni rispetto a un portafoglio ad alto rendimento. Un Sharpe Ratio di 11.28% indica che il portafoglio offre un buon rendimento aggiustato per il rischio. PFE, con un peso del 57.67%, è l'elemento principale del portafoglio.

Passiamo ora ad un portafoglio con il massimo rapporto di Sharpe:

Maximum Sharpe Ratio portfolio ----

Performance

Return: 7.44% Volatility: 62.48% Sharpe Ratio: 11.91%

Weights

GOOGL: 0.00% INTC: 16.33% JNJ: 0.00% PFE: 27.32% XOM: 33.41% CVX: 22.94%

Questo portafoglio ha un ritorno moderato del 7.44%, che è positivo dato che l'obiettivo principale è ottenere il massimo rendimento possibile per unità di rischio. La volatilità del 62.48% è piuttosto alta, indicando che la variabilità nei rendimenti è significativa. Un Sharpe Ratio di 11.91% è indicativo di un buon equilibrio tra rischio e rendimento. Anche qui PFE è il titolo con il peso più alto, del 27.32%.

Proseguiamo con un confronto tra i portafogli a varianza minima calcolati precedentemente:

Portafoglio a minima varianza metodo simulativo

Annualized Return: 31.19%

Volatility: 102.42%

Sharpe Ratio: 30.45

Beta: 0.87

Equal Weighted Portfolio

Annualized Return: 148.24%

Volatility: 69.26%

Sharpe Ratio: 2.14

Beta: 0.92

Portafoglio a minima varianza metodo analitico

Annualized Return: 31.19%

Volatility: 102.42%

Sharpe Ratio: 30.45

Beta: 0.87

Equal Weighted Portfolio

Annualized Return: 148.24%

Volatility: 69.26%

Sharpe Ratio: 2.14

Beta: 0.92

Portafoglio a minima varianza metodo simulativo con rendimenti calcolati con CAPM

Annualized Return: 6.83%

Volatility: 59.95%

Sharpe Ratio: 11.40

Beta: 0.81

Equal Weighted Portfolio

Annualized Return: 148.24%

Volatility: 69.26%

Sharpe Ratio: 2.14

Beta: 0.92

Portafoglio a minima varianza metodo analitico con rendimenti calcolati con CAPM

Annualized Return: 6.76%

Volatility: 59.95%

Sharpe Ratio: 11.28

Beta: 0.94

Equal Weighted Portfolio

Annualized Return: 148.24%

Volatility: 69.26%

Sharpe Ratio: 2.14

Beta: 0.92

In conclusione, possiamo dire che i portafogli equamente ponderati offrono rendimenti più alti con volatilità più gestibile. I portafogli di minima varianza calcolati con il CAPM offrono rendimenti più bassi ma con volatilità inferiore. I portafogli di minima varianza senza CAPM mostrano Sharpe Ratio estremamente alti, ma comunque seppur più contenuti

anche quelli con CAPM hanno uno Sharpe Ratio alto. Infine, i portafogli di minima varianza tendono ad avere beta inferiori a 1, indicando minore volatilità rispetto al mercato.

CONCLUSIONE

In conclusione, lo studio condotto sulle sei multinazionali Alphabet Inc. Intel Corporation, Johnson & Johnson, Pfizer Inc., Exxon Mobil Corporation e Chevron Corporation, ha evidenziato l'efficacia dei modelli ARIMA nella previsione dei rendimenti e ha dimostrato che la strategia di trading basata sulle medie mobili esponenziali tende a performare meglio rispetto alla strategia "Buy and Hold". Infine la costruzione dei portafogli ottimali ha dimostrato la capacità di ottenere rendimenti elevati con un rischio contenuto, evidenziando l'importanza di una gestione attiva e ottimizzata degli investimenti.