

Documentazione di Progetto

Business Intelligence per i Servizi Finanziari

Tommaso Cammelli
851593

6 giugno 2022

Indice

1	Sommario dei dati utilizzati	2
1.1	Presentazione e descrizione dei titoli selezionati	2
1.2	Funzioni utilizzate per download e fusione	3
1.3	presentazione dei dati	3
2	Statistiche descrittive	4
2.1	Rendimenti semplici e composti	4
2.1.1	Titoli Tecnologici	4
2.1.2	Titoli Militari	5
2.1.3	Titoli Bancari	6
2.2	Istogrammi sui rendimenti e dispersione	8
2.2.1	Titoli tecnologici (Meta/Alphabet)	8
2.2.2	Titoli militari (Raytheon/Lockheed Martin)	8
2.2.3	Titoli bancari (Bank of America/JPMorgan Chase)	9
2.3	Grafici diagnostici a 4 sezioni	10
2.3.1	Grafici Diagnostici per Meta Platforms, Inc. (FB)	10
2.3.2	Grafici Diagnostici per Alphabet Inc. (GOOG)	11
2.3.3	Grafici Diagnostici per Raytheon Technologies Corporation. (RTX)	12
2.3.4	Grafici Diagnostici per Lockheed Martin Corporation. (LMT)	13
2.3.5	Grafici Diagnostici per Bank of America. (BAC)	14
2.3.6	Grafici Diagnostici per JPMorgan Chase. (JPM)	15
2.4	Statistiche descrittive univariate	16
2.5	Matrici di varianze/covarianze dei rendimenti	19
2.6	Matrice di correlazione dei rendimenti	19
2.7	Grafico sulla correlazione dei titoli nel tempo e scatter plot	20
2.7.1	Titoli tecnologici FB/GOOG	20
2.7.2	Titoli militari RTX/LMT	21
2.7.3	Titoli bancari BAC/JPM	22
3	Analisi di previsione	24
3.1	Modello di previsione per Meta (FB)	24
3.2	Modello di previsione per Alphabet (GOOG)	26
3.3	Modello di previsione per Raytheon (RTX)	27
3.4	Modello di previsione per Lockheed Martin (LMT)	28
3.5	Modello di previsione per Bank of America (BAC)	29
3.6	Modello di previsione per JPMorgan Chase (JPM)	30
3.7	Extra: Forecasting usando tecniche Deep Learning (MLP)	31
4	Strategie di trading e backtesting	32
4.1	Strategia buy/sell mediante Bollinger's Bands	32
4.2	Backtesting su titolo Meta (FB)	33
5	Capital Asset Pricing Model - CAPM	34
5.1	Calcolo dell'indice beta	34
5.2	Calcolo esposizione dei titoli con Fama-French	35
5.2.1	Esposizione di Meta (FB)	35
5.2.2	Esposizione di Alphabet (GOOG)	36
5.2.3	Esposizione di Raytheon (RTX)	36
5.2.4	Esposizione di Lockheed Martin (LMT)	37
5.2.5	Esposizione di Bank of America (BAC)	37
5.2.6	Esposizione di JPMorgan Chase (JPM)	38
5.3	Calcolo dei rendimenti attesi usando beta	38
6	Costruzione di un portfolio	39
6.1	Costruzione del portfolio ottimale	39
6.1.1	Costruzione tramite simulazioni di Monte Carlo	39
6.2	Beta dei portafogli ottimali	42
6.2.1	beta del portfolio costituito solo con dati passati	42
6.2.2	beta del portfolio costituito con dati di previsione	43
6.3	Costruzione del portfolio effettivo e confronto	43

6.3.1	Dati reali passati	43
6.3.2	Dati di previsione	43
7	Conclusioni	44

1 Sommario dei dati utilizzati

1.1 Presentazione e descrizione dei titoli selezionati

Per questo progetto sono stati presi in considerazione 6 titoli azionari, appartenenti a 3 settori diversi:

- **Settore tecnologico:** Meta Platforms, Inc. (FB), Alphabet Inc. (GOOG)
- **Settore militare:** Raytheon Technologies Corporation (RTX), Lockheed Martin Corporation (LMT)
- **Settore bancario:** Bank of America Corporation (BAC), JPMorgan Chase & Co. (JPM)

Motivazione per scelta dei titoli

- **Meta Platforms, Inc. (FB):** Questo titolo è stato scelto in quanto è una delle aziende con la *market capitalization* più alta nel mondo¹, e facendo parte del *faang*² ho ritenuto essere un investimento solido. considerando il recente crollo del prezzo potrebbe essere un buon momento per prendere posizioni lunghe sul titolo³.
- **Alphabet Inc. (GOOG):** Questo titolo insieme a quello precedente fa parte di *faang*², è stato scelto in quanto è interessante da confrontare contro altri titoli del settore tecnologico come FB, soprattutto in momenti critici come la crisi finanziaria causata dall'epidemia di COVID-19, dove GOOG ha subito un crollo del 29% nel primo trimestre⁴.
- **Raytheon Technologies Corporation (RTX):** RTX ha mostrato negli anni un trend in salita abbastanza stabile, nonostante il crollo durante la crisi finanziaria del 2020 RTX è riuscito a recuperare il crollo.⁵
- **Lockheed Martin Corporation (LMT):** LMT è stata scelta in quanto è fa parte anche lui nella categoria militare e permette di confrontare l'andamento di mercato nel confronto di RTX.
- **Bank of America Corporation (BAC):** Come primo titolo finanziario ho scelto BAC in quanto nonostante la volatilità negli ultimi anni a causa dell'epidemia, il titolo ha mostrato un leggero trend di salita negli anni ed il crollo del prezzo potrebbe rivelarsi una opportunità.
- **JPMorgan Chase & Co. (JPM):** JPM come per BAC ha risentito molto dalla crisi finanziaria del 2020, negli anni tuttavia ha mostrato un trend di salita più evidente rispetto a BAC, ritengo che potrebbe essere vantaggioso a lungo termine

¹pagina web di referenza: <https://companiesmarketcap.com/tech/largest-tech-companies-by-market-cap/>

²Acronimo dei cinque top stocks americani nel settore tecnologico, <https://www.investopedia.com/terms/f/faang-stocks.asp>

³<https://finance.yahoo.com/news/good-time-increase-meta-platforms-142109519.html>

⁴articolo riferimento declino, <https://www.investopedia.com/alphabet-googl-sells-off-after-revenue-decline-5072988>

⁵<https://www.investopedia.com/raytheon-technologies-drops-then-pops-on-earnings-beat-5074746>

1.2 Funzioni utilizzate per download e fusione

Per il download dei dati da Yahoo! Finance⁶ è stata utilizzata la nota libreria di python `yfinance`⁷ dove attraverso la funzione `download()` ha permesso di scaricare i dati di interesse nel periodo rilevante per questo progetto.

```
# Esempio di download da Yahoo! Finance dello storico prezzi di FB
import yfinance as yf
```

```
fb_df = yf.download('FB', start='2011-11-30', end='2021-11-30')
```

Relativamente alla fusione dei dati scaricati in un unico `DataFrame` di `Pandas`⁸ è stata utilizzata la funzione `DataFrame()` per creare un nuovo `DataFrame` vuoto, sono stati poi usati i costrutti base di python per popolare il `DataFrame` con i nostri dati di interesse.

```
# Esempio di fusione dei dati da due indici scaricati precedentemente
import pandas as pd
```

```
adj_close_tot = pd.DataFrame()
adj_close_tot["Meta_Price"] = fb_df[["Adj_Close"]]
adj_close_tot["Alphabet_Price"] = goog_df[["Adj_Close"]]
```

1.3 presentazione dei dati

Rappresentiamo i dati ottenuti tramite un grafico a linee che si trova alla figura 1 dove si mostra la variazione di prezzo di tutti gli stock considerati in questo progetto⁹ nel periodo da 30-11-2011 a 30-11-2021.

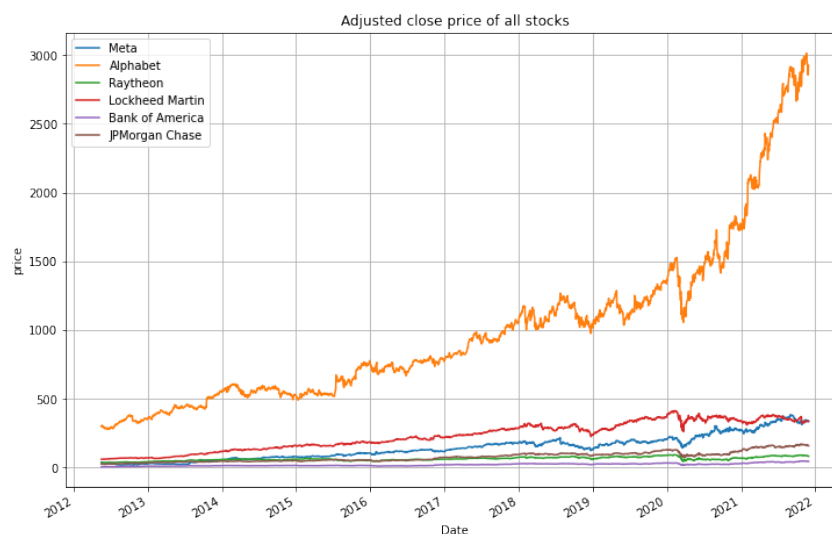


Figura 1: grafico con prezzo degli stock da 18/05/2012 a 30/11/2021

Tutti i grafici del progetto sono stati generati utilizzando la libreria di python `matplotlib`¹⁰ che tramite apposite funzioni ha permesso la quasi totale personalizzazione dei grafici per semplificare la lettura dei dati.

Rappresentiamo ora alla figura 2 le prime 10 righe della tabella che contiene il prezzo combinato di tutti gli stock considerati (stessa tabella utilizzata per il plot del grafico qui sopra), fusi in un solo `DataFrame` grazie a `Pandas`.

Nota: Meta Platforms, Inc. (FB) è stata quotata in borsa solo a partire dal 18/05/2012, a causa di ciò i dati aggregati partono solo da quella data.

⁶<https://finance.yahoo.com>

⁷Libreria FOSS per download di dati finanziari da Yahoo! finance, <https://pypi.org/project/yfinance/>

⁸Libreria per data analysis e manipulation, <https://pandas.pydata.org/>

⁹FB, GOOG, RTX, LMT, BAC, JPM

¹⁰Libreria per creare visualizzazioni dei dati anche interattive in Python, <https://matplotlib.org>

	Meta Price	Alphabet Price	Raytheon Price	Lockheed Martin Price	Bank of America Price	JPMorgan Chase Price
Date						
2012-05-18	38.230000	299.078979	36.082355	60.931610	6.052373	25.427305
2012-05-21	34.029999	305.908386	36.740398	61.557266	5.888562	24.683229
2012-05-22	31.000000	299.278229	36.860043	61.601433	6.017887	25.822113
2012-05-23	32.000000	303.592072	36.919865	61.351143	6.181696	26.011929
2012-05-24	33.029999	300.702881	36.640682	61.365898	6.155832	25.791744
2012-05-25	31.910000	294.660553	36.401409	60.880070	6.164454	25.434896
2012-05-29	28.840000	296.060303	37.433334	61.579334	6.414482	25.533600
2012-05-30	28.190001	293.016693	36.760330	61.683636	6.215916	25.024893
2012-05-31	29.600000	289.345459	36.944771	61.683636	6.345415	25.169157
2012-06-01	27.719999	284.423920	35.902882	60.506592	6.060519	24.242868

Figura 2: tabella con Adjusted Close degli stock da 18/05/2012 a 30/11/2021 (prime 10 righe)

2 Statistiche descrittive

2.1 Rendimenti semplici e composti

Dato che il prezzo degli asset spesso non è *stazionario* (media e varianza cambia nel tempo), ci viene comodo trasformare i prezzi in rendimenti per cercare di rendere la serie temporale stazionaria (la proprietà desiderata per la modellazione statistica).

2.1.1 Titoli Tecnologici

Per i titoli GOOG e FB sono stati calcolati i rendimenti semplici netti (grafico a figura 3) e i rendimenti composti (grafico a figura 4) e posti a confronto si nota una generale correlazione nei rendimenti.

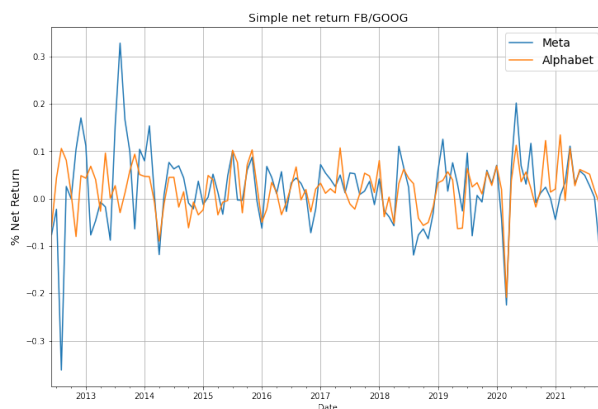


Figura 3: Rendimenti semplici netti FB e GOOG

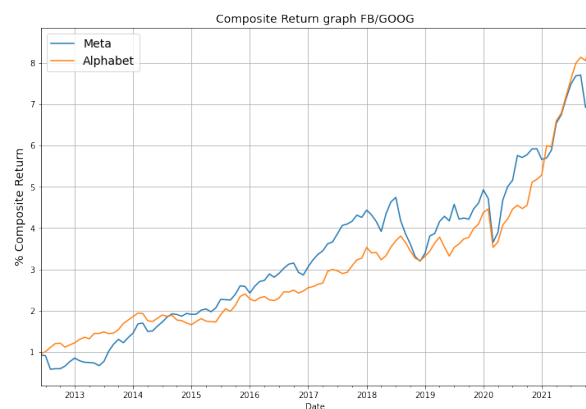


Figura 4: Rendimenti composti FB e GOOG

Note sui titoli tecnologici

Confrontando le serie storiche di GOOG e FB nel grafico a figura 1, utilizzando la funzione `.corr()` di pandas si mostra un forte correlazione positiva di 0.962272 (figura 5) tra i due titoli, come dimostrato anche dai simili rendimenti composti (ad eccezione per alcuni eventi).

	Meta Price	Alphabet Price
Meta Price	1.000000	0.962272
Alphabet Price	0.962272	1.000000

Figura 5: tabella con correlazione titoli GOOG e FB (metodo di Pearson)

Note sui rendimenti di Meta (FB)

Osservando il grafico in figura 3 relativamente ai rendimenti semplici di FB, notiamo 3 eventi di notevole discostamento dalla media.

Nel primo trimestre del 2012 si nota un significativo crollo, analizzando notizie e articoli si può ricondurre il crollo a scetticismo che c'è stato tra gli investitori¹¹ in quanto Facebook (ora Meta) all'epoca era stata appena quotata in borsa, il crollo è stato circa del 35% solo nei primi mesi.

Nel terzo trimestre del 2013 tuttavia viene evidenziato una crescita sostanziale rispetto alla media, nonostante il crollo capitato poco dopo la quotazione in borsa nel 2012, a metà anno il prezzo di FB è riuscito a raggiungere di nuovo il prezzo di IPO¹² raggiungendo poi verso il terzo trimestre la quota record del tempo di 50\$, in base a vari articoli¹³ si evidenzia come la crescita possa essere attribuita dalla introduzione delle pubblicità su dispositivi mobile (precedentemente la pubblicità era mostrata solo da sito web), a luglio 2013 facebook ha infatti annunciato che la pubblicità su mobile ha contribuito al 41% delle loro vendite nel secondo semestre.

Tra il primo ed il secondo trimestre del 2020, come evidenziato dal grafico c'è stato un grande discostamento dalla media, inizialmente molto negativo ma poco dopo molto positivo, la caduta iniziale del prezzo la si può attribuire alla crisi finanziaria del 2020, scatenata dalla epidemia da covid-19¹⁴.

Tuttavia grazie alla natura digitale del servizio, dunque non esposta alla epidemia come altre attività e all'annuncio di *Facebook Shops* il prezzo ha velocemente raggiunto il valore pre-crollo arrivando addirittura ad un record del tempo (20-05-2020) di \$230.75¹⁵

Note sui rendimenti di Alphabet (GOOG)

Sul titolo GOOG si notano distaccamenti negativi sul rendimento di entità notevolmente inferiore rispetto a FB, tuttavia si hanno distaccamenti positivi più numerosi (il net return raggiunge 0.1 più volte rispetto a FB) anche se spesso non oltre 0.1.

Come per FB, ad aprile 2020 c'è stato un decremento significativo del prezzo (che si è riflettuto sul rendimento) a causa della epidemia da covid-19¹⁶, ma sempre per il fatto che Alphabet Inc. fornisce prevalentemente servizi digitali il prezzo è ritornato al valore pre-crollo velocemente.

2.1.2 Titoli Militari

Per i titoli RTX e LMT sono stati calcolati i rendimenti semplici netti (grafico a figura 6) e i rendimenti composti (grafico a figura 7) e posti a confronto si nota una generale correlazione nei rendimenti, anche se di entità inferiore rispetto a FB/GOOG.

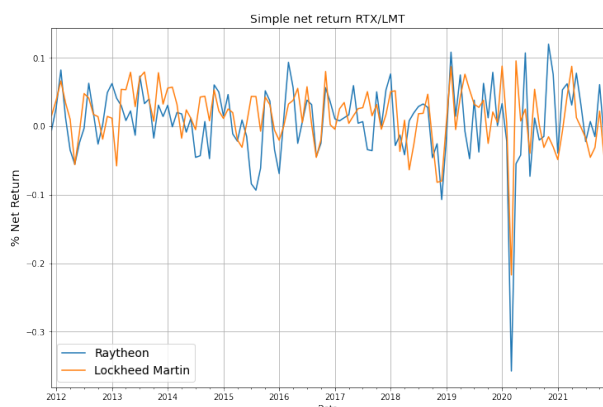


Figura 6: Rendimenti semplici netti RTX e LMT

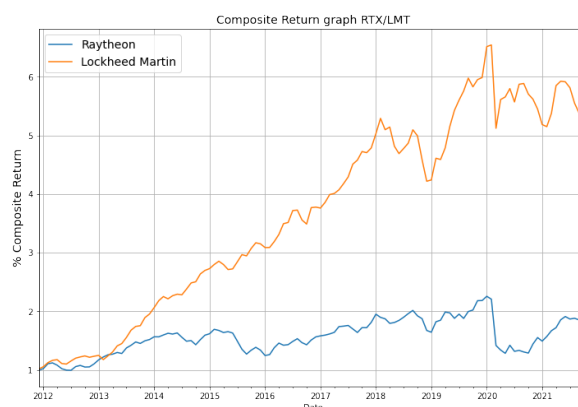


Figura 7: Rendimenti composti RTX e LMT

¹¹ <https://money.cnn.com/2012/05/22/markets/facebook-shares/index.htm>

¹² Initial Public Offering, <https://www.investopedia.com/terms/i/ipo.asp>

¹³ <https://www.reuters.com/article/us-facebook-ipoprice-idUSBRE96T1C120130730>,

<https://money.cnn.com/2013/09/26/investing/facebook-stock/index.html>

¹⁴ <https://www.investopedia.com/facebook-stock-crashes-into-bear-market-territory-4800598>

¹⁵ <https://www.cnn.com/2020/05/20/facebook-shares-reach-all-time-high-after-announcing-online-shopping-feature.html>

¹⁶ <https://www.investopedia.com/alphabet-stock-crashes-into-bear-market-territory-4800600>

Confrontando i titoli RTX e LMT usando sempre il grafico a figura 1 e utilizzando la funzione `.corr()` di pandas viene mostrata anche in questo caso una **forte correlazione positiva** di 0.836831 (figura 8).

	Raytheon Price	Lockheed Martin Price
Raytheon Price	1.000000	0.836831
Lockheed Martin Price	0.836831	1.000000

Figura 8: tabella con correlazione titoli RTX e LMT (metodo di Pearson)

Note sui rendimenti di Raytheon (RTX)

Osservando il grafico in figura 6 relativamente ai rendimenti semplici di RTX, si nota come il nel tempo è stato molto fluido, troviamo **3** eventi notevoli.

Sul titolo RTX è stato più complicato la ricerca di notizie coerenti con il crollo del prezzo in quanto è notevolmente meno famoso rispetto a FB o GOOG, tuttavia relativamente al primo ed al secondo crollo si può assumere siano in relazione alla politica di export delle armi americane.

Il crollo più notevole è accaduto durante la crisi finanziaria del 2020, in base a un articolo¹⁷ la motivazione del crollo è stata a causa della temporanea debolezza nel settore commerciale aerospaziale, causata dalla epidemia da COVID-19, si assume comunque che essendo una compagnia incentrata nel settore aerospaziale sarà comunque una migliore scelta rispetto alla competizione per il settore della difesa.

Note sui rendimenti di Lockheed Martin (LMT)

Il titolo LMT ha subito variazioni molto meno significanti rispetto a RTX e gli altri titoli (come si può vedere dal grafico a figura 6), analizzando varie notizie nel web si identificano **2** eventi significativi.

Tra fine 2018 e l'inizio del 2019 si è registrato un crollo del 18.4% relativo allo stock LMT, in base ad un articolo¹⁸ le cause del crollo sono molteplici, una delle cause identificate è una continua diminuzione del budget dalla casa bianca, una altra è la dimissione improvvisa del CFO¹⁹ in quanto era ben visto dagli investitori per le sue abilità comunicative.

Il crollo più importante del prezzo è stato durante la crisi finanziaria del 2020, dopo il crollo nel primo trimestre tuttavia c'è stato un notevole recupero poco dopo messo poi a rischio verso ottobre a causa di problemi nella supply chain della produzione²⁰ causato dal tilt nelle fabbriche e nei trasporti.

Nello stesso articolo si evidenzia inoltre come la motivazione della instabilità verso la fine del 2020 sia anche causata da una incertezza politica, in quanto si assumeva che una vittoria democratica avrebbe tagliato il budget alla difesa, anche se il rischio di guerra nucleare dovrebbe evitare un crollo del titolo.

2.1.3 Titoli Bancari

Per i titoli BAC e JPM sono stati calcolati i rendimenti semplici netti (grafico a figura 9) e i rendimenti composti (grafico a figura 10) e posti a confronto anche per questi due titoli si nota una forte correlazione nei rendimenti.

Confrontando i titoli BAC e JPM usando il grafico a figura 1 e utilizzando la funzione `.corr()` di pandas viene mostrata una **forte correlazione positiva** di 0.909522 (figura 11).

Note sui rendimenti di Bank of America (BAC)

Il titolo BAC ha in multiple occasioni dei sostanziali aumenti di prezzo ma anche dei crolli, identifichiamo almeno **5** eventi notevoli.

Nel primo trimestre del 2012 si nota un netto crollo del prezzo, facendo delle ricerche ho trovato un articolo²¹ che spiega come BAC ha accettato di pagare 2.3 milioni di dollari agli investitori per patteggiare una causa legale per gli eventi legati alla crisi finanziaria del 2008. Essendo un periodo vicino al 2008 si assume anche che la alta volatilità si anche data dalla poca fiducia degli investitori in questo titolo.

¹⁷<https://www.fool.com/investing/2020/07/02/heres-why-raytheon-technologies-shares-are-down-34.aspx>

¹⁸<https://www.fool.com/investing/2019/01/11/heres-why-lockheed-martin-lost-184-in-2018.aspx>

¹⁹Chief Financial Officer, https://en.wikipedia.org/wiki/Chief_financial_officer

²⁰<https://www.investopedia.com/lockheed-martin-lmt-sells-off-despite-strong-quarter-5083204>

²¹<https://www.reuters.com/article/uk-bofa-lawsuit-idUKBRE88R11X20120928>

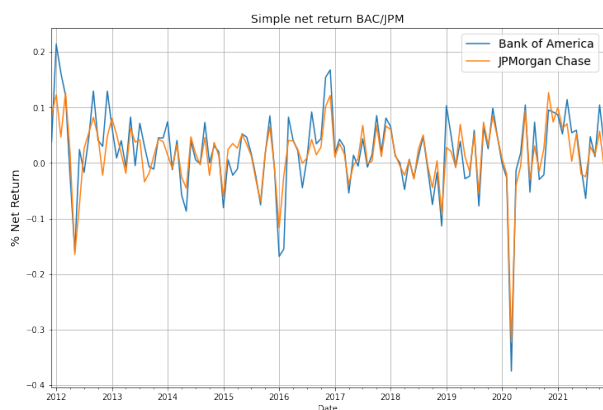


Figura 9: Rendimenti semplici netti BAC e JPM

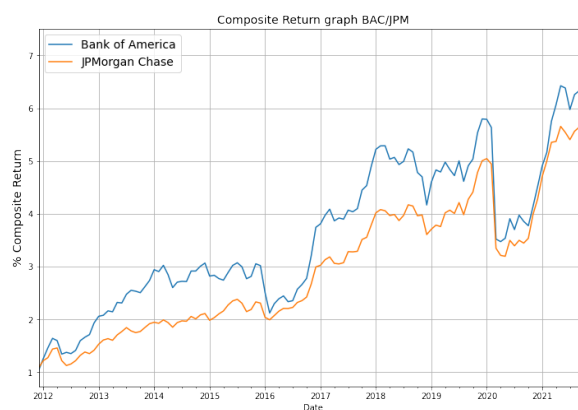


Figura 10: Rendimenti composti BAC e JPM

	Bank of America Price	JPMorgan Chase Price
Bank of America Price	1.000000	0.909522
JPMorgan Chase Price	0.909522	1.000000

Figura 11: tabella con correlazione titoli BAC e JPM (metodo di Pearson)

Nel 2016 si registra un altro periodo di instabilità per questo titolo, risultato poi verso la fine dell'anno con un aumento del 34% sul prezzo, secondo questo articolo²² il declino del prezzo all'inizio dell'anno è stato a causa del crollo dei prezzi relativi all'energia, successivamente i prezzi del settore energetico ci hanno messo poco tempo a recuperare, tuttavia nonostante il recupero a metà anno il prezzo di BAC è comunque sceso, la causa della seconda declinazione dei prezzi, si può attribuire alla improvvisa decisione dell'Inghilterra di uscire dalla unione europea (brexit), che ha causato un momento di incertezza finanziaria, soprattutto con Bank of America in quanto a causa di brexit avrebbe dovuto spostare parte delle proprie operazioni in un altro paese europeo.

Verso la fine dell'anno poi c'è stata una crescita notevole e inaspettata del prezzo, secondo l'articolo questa crescita può essere stata a causa della vittoria del partito Repubblicano delle elezioni presidenziali, in quanto il candidato presidente aveva promesso la rimozione di regolatorie instaurate nel 2010 dal precedente Presidente.

Nel 2020 come per i titoli mostrati in precedenza c'è stato un grosso crollo del prezzo di BAC a causa della crisi finanziaria del 2020, secondo questo articolo²³ nonostante l'ottimismo del CEO riguardante il futuro recupero dal crollo, i mesi successivi sono rimasti altamente instabili e non hanno portato come per i titoli tecnologici a un aumento stabile.

Note sui rendimenti di JPMorgan Chase (JPM)

Controllando i grafici relativi al prezzo e rendimenti, si può notare come il titolo JPM sia stato molto correlato con BAC, in quanto operano nello stesso settore e sono entrambi titoli finanziari.

Relativamente agli eventi del 2012 e 2016 si può assumere che gli stessi eventi di BAC abbiano causato il crollo sui rendimenti e quindi periodi di instabilità.

Relativamente alla crisi finanziaria del 2020, secondo questo articolo²⁴ oltre al crollo del prezzo a causa della epidemia, c'è stato un enorme movimento di 6.8 milioni di dollari alla riserva di credito della banca, movimento che si assume sia stato effettuato per proteggere la banca da un aumento di default relativi alle attività che hanno effettuato prestiti con JPM, questo movimento ha causato un ulteriore crollo del prezzo che si è propagato nei mesi successivi.

²²<https://www.fool.com/investing/2016/12/29/why-bank-of-americas-stock-climbed-34-in-2016.aspx>

²³<https://www.investopedia.com/bank-of-america-earnings-4770948>

²⁴<https://www.cnbc.com/2020/04/14/jpmorgan-chase-jpm-earnings-q1-2020.html>

2.2 Istogrammi sui rendimenti e dispersione

Per misurare la dispersione sui titoli ci torna utile il concetto di deviazione standard, grazie ad essa si può avere una idea della volatilità associata al titolo, un dato utile per effettuare investimenti e/o strategie di trading²⁵.

2.2.1 Titoli tecnologici (Meta/Alphabet)

L'istogramma del rendimento relativo ai titoli GOOG e FB può essere visto a figura 12, mentre la distribuzione dei rendimenti si trova nel grafico in figura 13. Dai dati si evidenzia come la maggior parte dei rendimenti avviene tra 0.0 e 0.1, oltre al fatto che alphabet nel periodo di interesse ha avuto rendimenti più alti rispetto a Meta.

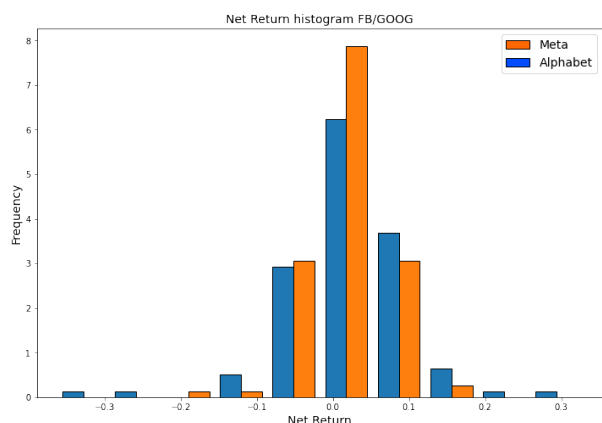


Figura 12: Istogramma rendimenti FB e GOOG

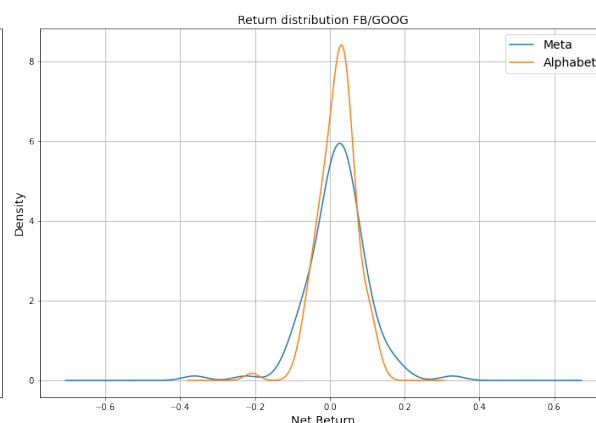


Figura 13: Distribuzione di FB e GOOG

Inoltre é stata calcolata grazie a *Pandas* la Deviazione Standard, dove il risultato si può vedere nella tabella a figura 14.

```
Deviazione Standard
Rendimenti Netti Meta      0.080678
Rendimenti Netti Alphabet  0.050795
dtype: float64
```

Figura 14: Deviazione Standard dei titoli FB e GOOG

2.2.2 Titoli militari (Raytheon/Lockheed Martin)

L'istogramma del rendimento relativo ai titoli RTX e LMT può essere visto nel grafico a figura 15, mentre la distribuzione dei rendimenti si trova nel grafico a figura 16.

Dai dati si evidenzia che i rendimenti si trovano per la maggior parte tra -0.05 e 0.5, Lockheed martin ha avuto rendimenti più alti rispetto a Raytheon.

Sempre con *pandas* è stata calcolata la Deviazione Standard, dove il risultato è riportato nella tabella a figura 17.

²⁵<https://www.investopedia.com/terms/s/standarddeviation.asp>

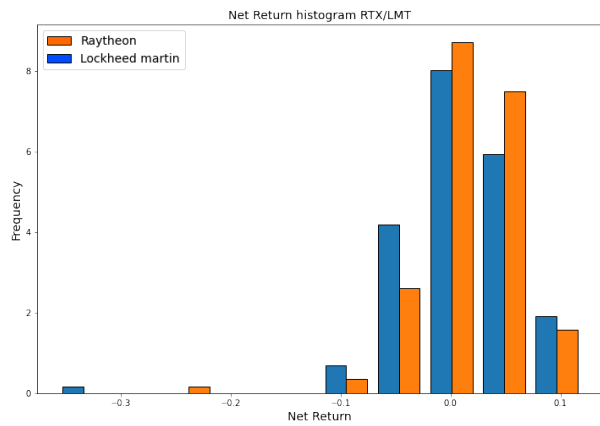


Figura 15: Istogramma rendimenti RTX e LMT

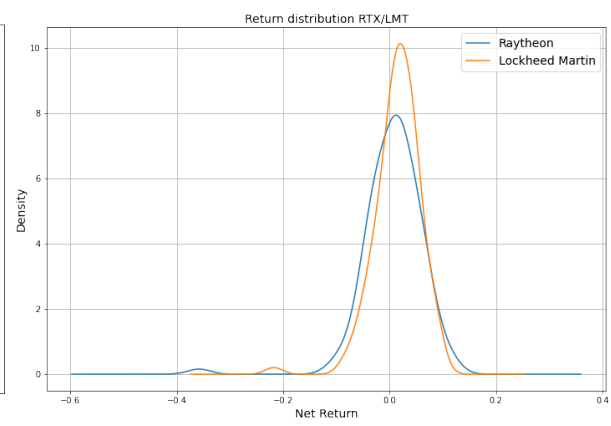


Figura 16: Distribuzione di RTX e LMT

```

Deviazione Standard
Rendimenti Netti Raytheon      0.055370
Rendimenti Netti Lockheed Martin 0.042682
dtype: float64

```

Figura 17: Deviazione Standard dei titoli RTX e LMT

2.2.3 Titoli bancari (Bank of America/JPMorgan Chase)

L'istogramma per il rendimento dei titoli finanziari BAC e JPM si può vedere nel grafico nella figura 18, mentre la distribuzione dei rendimenti si può vedere nel grafico in figura 19.

Dai dati si evidenzia che i rendimenti si trovano per la maggior parte tra -0.01 e 0.1 , Bank of America ha avuto rendimenti più alti rispetto a JPMorgan Chase.

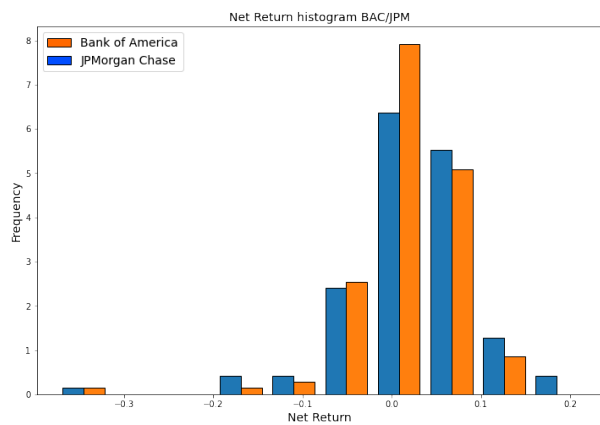


Figura 18: Istogramma rendimenti BAC e JPM

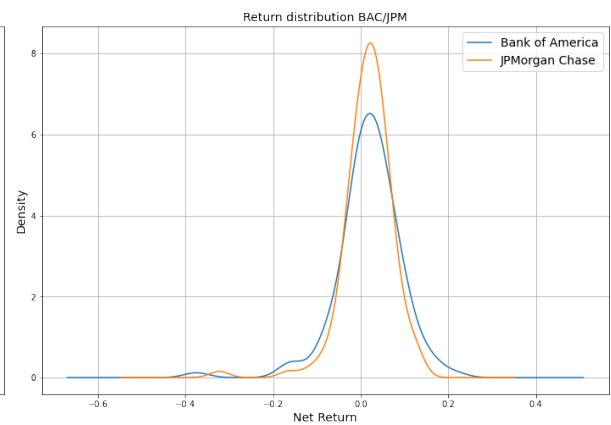


Figura 19: Distribuzione di BAC e JPM

Con *pandas* è stata calcolata la Deviazione Standard, dove il risultato è nella tabella a figura 20.

Deviazione Standard	
Rendimenti Netti Bank of America	0.072910
Rendimenti Netti JPMorgan Chase	0.056944
dtype: float64	

Figura 20: Deviazione Standard dei titoli BAC e JPM

2.3 Grafici diagnostici a 4 sezioni

Vengono mostrati per i titoli considerati la serie di 4 grafici diagnostici (Istogramma, kernel density, Q-Q plot e boxplot).

Questi grafici rappresentano 4 modi diversi per rappresentare la **Distribuzione Normale**, fondamentale per studiare la distribuzione sui rendimenti dei titoli.

Questi grafici ci permettono inoltre di identificare gli **outliers**, osservazioni che sono significativamente differenti dalla maggioranza, è fondamentale prima di lavorare con algoritmi di machine learning identificare tali outliers in quanto possono influenzare i risultati portando a risultati incorretti o con bias.

2.3.1 Grafici Diagnostici per Meta Platforms, Inc. (FB)

Per Meta, possiamo trovare i grafici alla figura 21, si può notare come i rendimenti sono distribuiti normalmente e simmetricamente, Si notano inoltre degli outliers: due tra 0.2 e 0.35 e due tra -0.2 e -0.4.

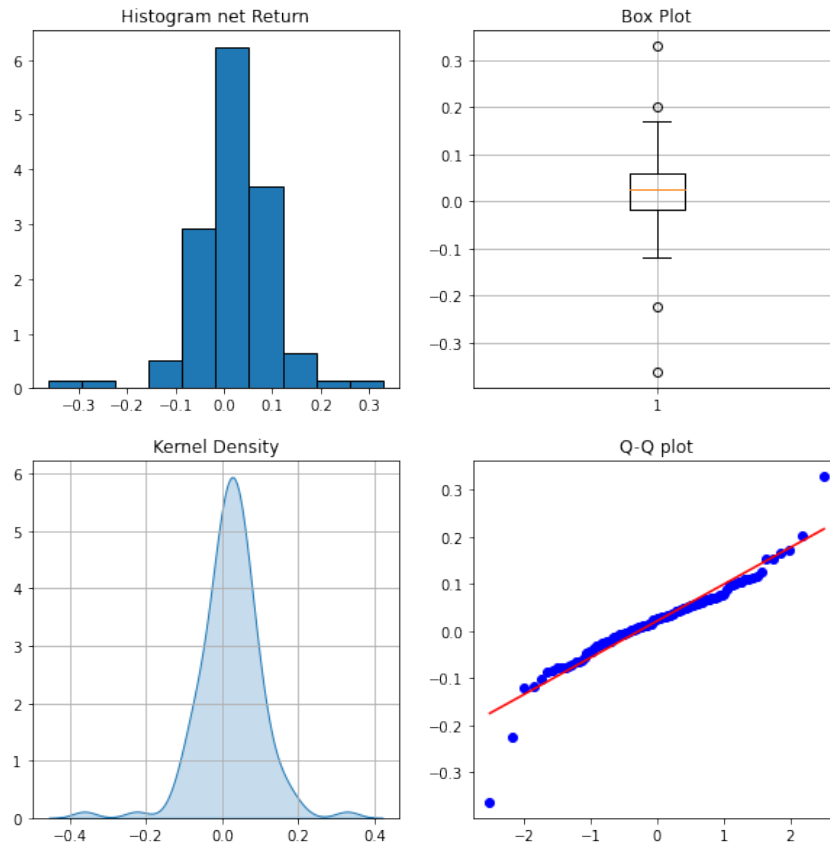


Figura 21: Grafici diagnostici per Meta (FB)

2.3.2 Grafici Diagnostici per Alphabet Inc. (GOOG)

Per Alphabet, possiamo trovare i grafici alla figura 22, possiamo notare anche qui che i rendimenti sono distribuiti normalmente, con una notevole inclinazione positiva rispetto a FB, Anche per questo titolo si nota un outliers: tra -0.20 e -0.25 .

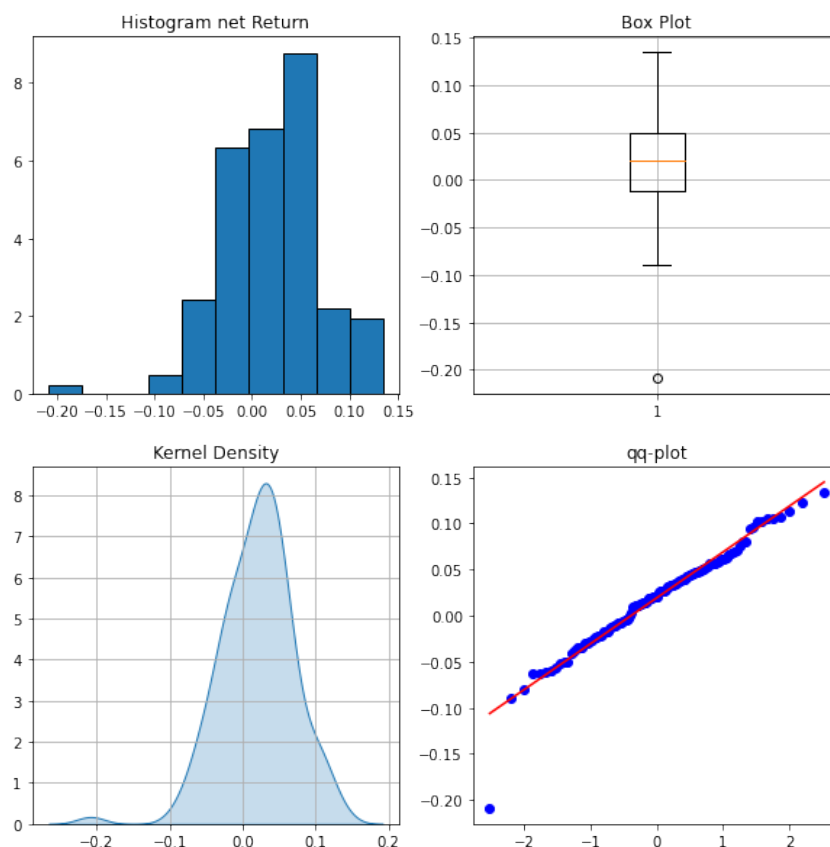


Figura 22: Grafici diagnostici per Alphabet (GOOG)

2.3.3 Grafici Diagnostici per Raytheon Technologies Corporation. (RTX)

Per Raytheon, i grafici si trovano alla figura 23, possiamo notare che qui i rendimenti sono distribuiti normalmente e simmetricamente. Per questo titolo troviamo un outliers: tra -0.35 e -0.4 .

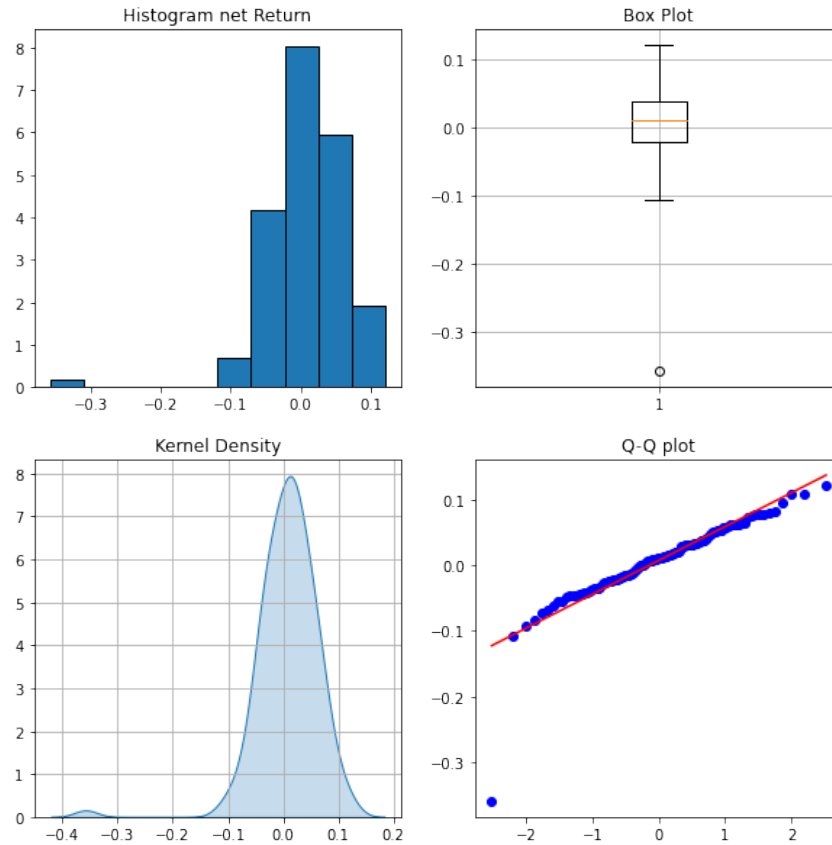


Figura 23: Grafici diagnostici per Raytheon (RTX)

2.3.4 Grafici Diagnostici per Lockheed Martin Corporation. (LMT)

Per Lockheed Martin, i grafici si trovano alla figura 24, anche qui i rendimenti sono distribuiti normalmente e in maniera simmetrica (la curva è più translata verso destra). Per questo titoli abbiamo tre outliers: due tra -0.05 e -0.10 , e uno tra -0.20 e -0.25 .

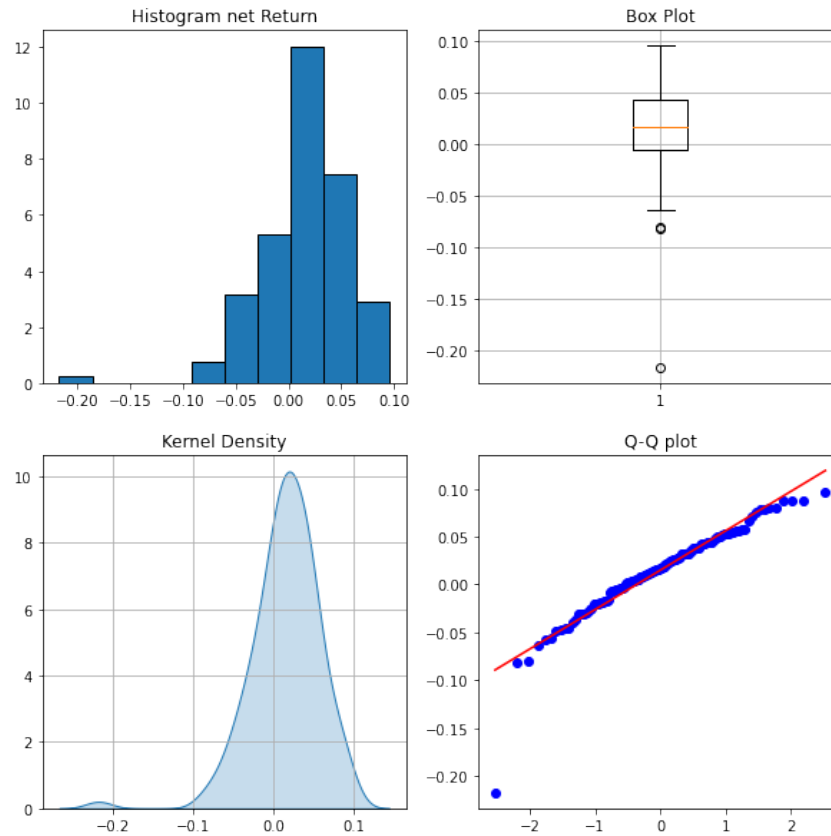


Figura 24: Grafici diagnostici per Lockheed Martin (LMT)

2.3.5 Grafici Diagnostici per Bank of America. (BAC)

Per Bank of America, i grafici diagnostici sono alla figura 25, anche qui sono distribuiti generalmente normalmente e simmetricamente. Per BAC abbiamo numerosi outliers: tre tra 0.15 e 0.25, quattro o più tra -0.1 e -0.2 e uno tra -0.35 e -0.4.

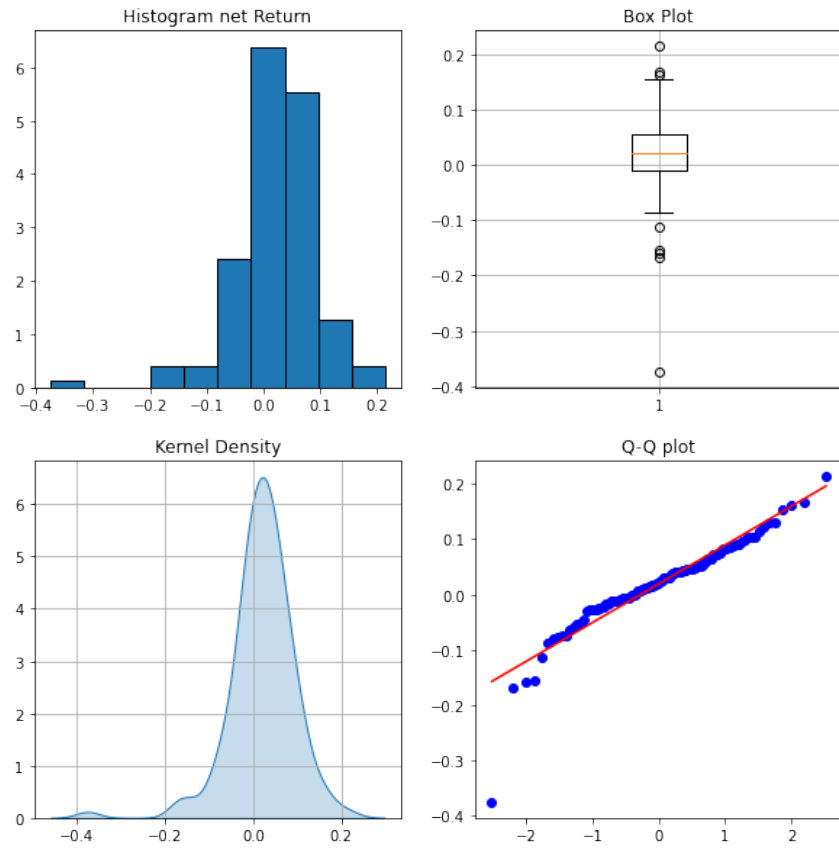


Figura 25: Grafici diagnostici per Bank of America (BAC)

2.3.6 Grafici Diagnostici per JPMorgan Chase. (JPM)

Per JPMorgan Chase, troviamo i grafici diagnostici alla figura 26, la distribuzione è anche qui normale e leggermente inclinata verso destra. Per questo titolo abbiamo 3 outliers: due tra -0.1 e -0.2 e uno tra -0.3 e -0.35 .

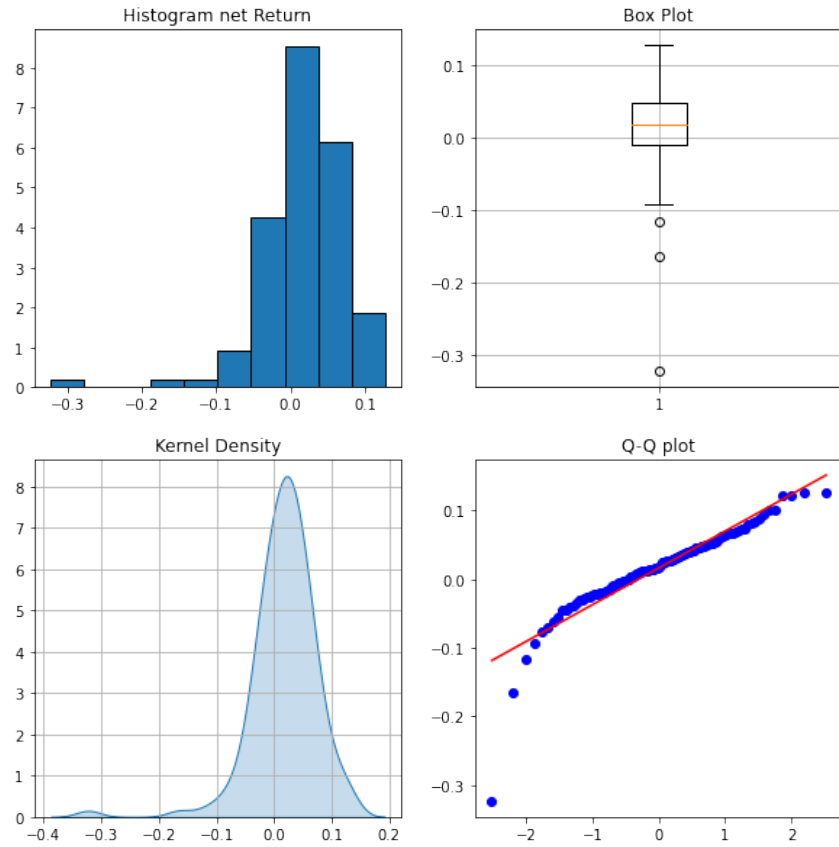


Figura 26: Grafici diagnostici per JPMorgan Chase (JPM)

2.4 Statistiche descrittive univariate

Per ogni serie di rendimenti sono state considerate le seguenti statistiche univariate: media, varianza, deviazione standard, asimmetria e curtosi.

Tali statistiche univariate servono per [TODO]

Statistiche per Meta (FB)

Per Meta identifichiamo una volatilità del 36.44%, i grafici e la tabella sono a figura 27 e 28.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0207	0.0065	0.0807	-0.5917	5.5871

Figura 27: Statistiche univariate per Meta (FB)

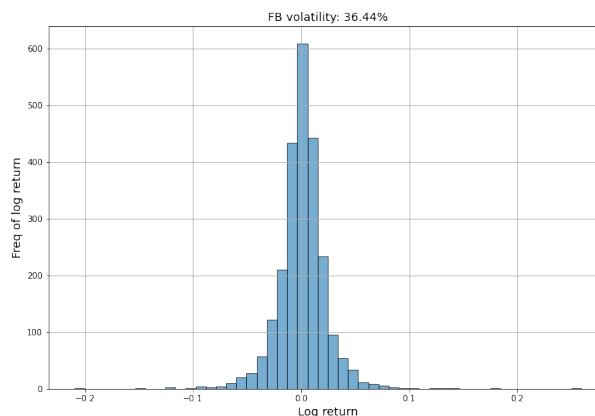


Figura 28: Grafico volatilità di Meta (FB)

Statistiche per Alphabet (GOOG)

Per Alphabet identifichiamo una volatilità del 25.11%, i grafici e la tabella sono a figura 29 e 30.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0192	0.0025	0.0501	-0.6988	2.6747

Figura 29: Statistiche univariate per Alphabet (GOOG)

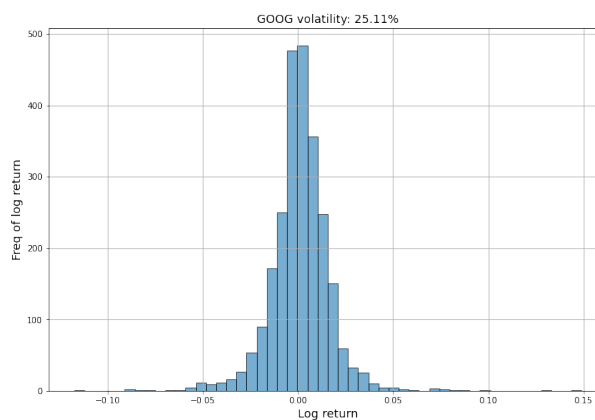


Figura 30: Grafico volatilità di Alphabet (GOOG)

Statistiche per Raytheon (RTX)

Per Raytheon identifichiamo una volatilità del 25.49%, i grafici e la tabella sono a figura 31 e 32.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0072	0.0031	0.0554	-2.3533	14.7837

Figura 31: Statistiche univariate per Raytheon (RTX)

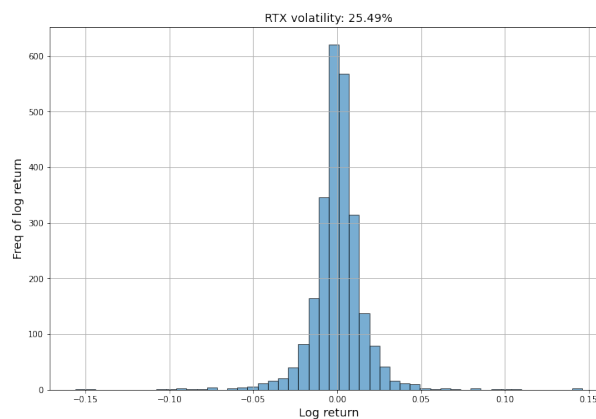


Figura 32: Grafico volatilità di Raytheon (RTX)

Statistiche per Lockheed Martin (LMT)

Per Lockheed Martin identifichiamo una volatilità del 21.1%, i grafici e la tabella sono a figura 33 e 34.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0148	0.0018	0.0427	-1.4478	6.3315

Figura 33: Statistiche univariate per Lockheed Martin (LMT)

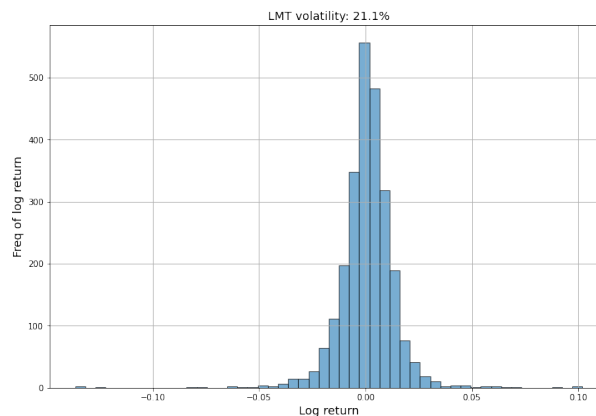


Figura 34: Grafico volatilità di Lockheed Martin (LMT)

Statistiche per Bank of America (BAC)

Per Bank of America identifichiamo una volatilità del 31.74%, i grafici e la tabella sono a figura 35 e 36.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0194	0.0053	0.0729	-1.3789	6.9176

Figura 35: Statistiche univariate per Bank of America (BAC)

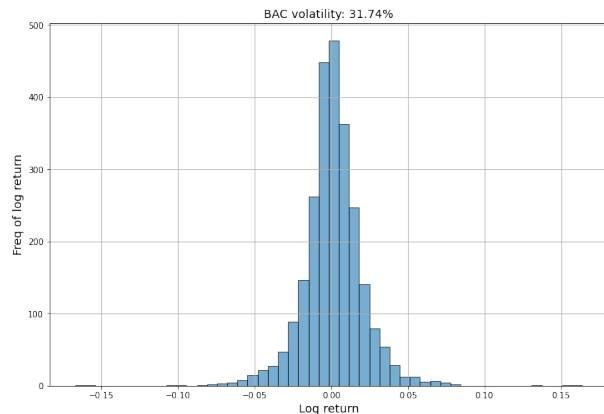


Figura 36: Grafico volatilità di Bank of America (BAC)

Statistiche per JPMorgan Chase (JPM)

Per JPMorgan Chase identifichiamo una volatilità del 27.04%, i grafici e la tabella sono a figura 37 e 38.

	Media	Varianza	Deviazione Standard	Asimmetria	Curtosi
Adj Close	0.0166	0.0032	0.0569	-1.9897	10.3171

Figura 37: Statistiche univariate per JPMorgan Chase (JPM)

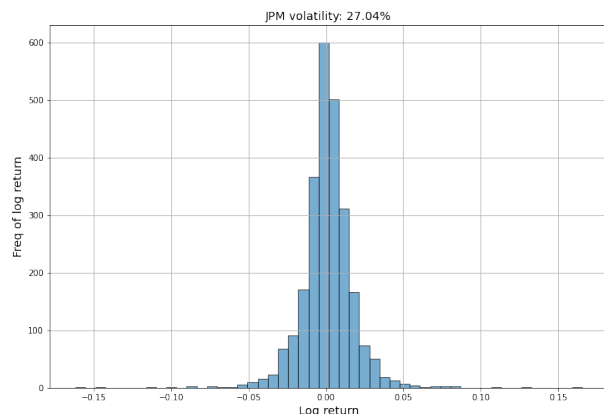


Figura 38: Grafico volatilità di JPMorgan Chase (JPM)

Note sulle statistiche presentate qui sopra

Basandosi sulle statistiche descrittive univariate presentate qui sopra possiamo concludere che:

- Utilizzando la **media** si è notato che, **Meta (FB)** e **Bank of America (BAC)** hanno il rendimento più alto, mentre **Raytheon (RTX)** e **Lockheed Martin (LMT)** hanno il rendimento più basso.
- Il titolo **Meta (FB)** ha una distribuzione dei rendimenti più vicina alla distribuzione normale, mentre **Raytheon (RTX)** è più lontano dalla normale.
- La volatilità più alta è stata riscontrata tra le azioni di **Meta (FB)** e **Bank of America (BAC)**, invece la più bassa tra **Lockheed Martin (LMT)** e **Alphabet (GOOG)**.
- La deviazione standard più bassa (indice di rischio per un asset) la hanno **Lockheed Martin (LMT)** e **Alphabet (GOOG)**, mentre quella più alta la hanno **Meta (FB)** e **Bank of America (BAC)**.

2.5 Matrice di varianze/covarianze dei rendimenti

La matrice di varianze/covarianze è una particolare matrice dove sulla diagonale abbiamo la varianza del relativo titoli e nelle altre posizioni abbiamo il valore di covarianza della coppia di titoli, questa matrice ci permette quindi di osservare questi due valori in una sola matrice.

Covarianza: La covarianza cov è una misura che rappresenta come si muove la media di due variabili aleatorie. Nella finanza è utilizzata per determinare la relazione direzionale tra i rendimenti di due asset, una covarianza positiva indica che i rendimenti dei due asset si muovono insieme mentre una covarianza negativa indica che si muovono inversamente.

$$cov(x, y) = \frac{(x_i - x_m) \cdot (y_i - y_m)}{n - 1}$$

dove

x_i = Rendimento i-esimo del titolo X

x_m = La media del titolo X del dataset

y_i = Rendimento i-esimo del titolo Y

y_m = La media del titolo Y del dataset

Varianza: La varianza σ^2 è una misura statistica utilizzata per determinare la distanza dei numeri dalla loro media in un dataset, in finanza è utilizzata per determinare la volatilità di un asset.

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} = cov(x, x)$$

$$mat_{var/cov} = \begin{bmatrix} \sigma_1^2 & cov(1, 2) & \cdots & cov(1, j) \\ cov(2, 1) & \sigma_2^2 & \cdots & cov(2, j) \\ \vdots & \vdots & \ddots & \vdots \\ cov(i, 1) & cov(i, 2) & \cdots & \sigma_i^2 \end{bmatrix}$$

	BAC	FB	GOOG	JPM	LMT	RTX
BAC	3.822654	1.275829	1.315323	2.887410	1.123326	1.861807
FB	1.275829	5.347862	1.843259	1.093739	0.780299	0.956870
GOOG	1.315323	1.843259	2.536485	1.147189	0.775151	1.099412
JPM	2.887410	1.093739	1.147189	2.790345	1.057372	1.725901
LMT	1.123326	0.780299	0.775151	1.057372	1.799854	1.207929
RTX	1.861807	0.956870	1.099412	1.725901	1.207929	2.616264

Figura 39: Matrice di Varianza/Covarianza su tutti i titoli

Considerando la tabella in figura 39, si può notare come JPM e BAC hanno un valore di Covarianza relativamente alto del 2.88%. Tale valore indica uno stesso comportamento in caso di aumento del prezzo o crollo, questa considerazione implica un maggior rischio nel nostro portfolio.

Si nota inoltre che GOOG e LMT hanno una covarianza del 0.77%, questo valore più vicino allo zero rispetto alle altre coppie implica un minor rischio nel portfolio tra questi due titoli.

2.6 Matrice di correlazione dei rendimenti

La matrice di correlazione dei rendimenti ci permette di osservare per ogni coppia di titoli il valore di correlazione sui rendimenti.

Correlazione: La correlazione r è un termine statistico che descrive il grado di coordinazione tra due variabili, se le due variabili si muovono nella stessa direzione la coordinazione è positiva, se si muovono in direzione opposta allora è negativa.

Il valore massimo è 1.0 mentre quello minimo è -1.0, quando ci avviciniamo agli estremi, si dice che la correlazione è **forte**, altrimenti se il valore è vicino allo zero si parla di correlazione **bassa** (o non esistente).

$$r = \frac{n \cdot (\sum(X, Y) - (\sum(X) \cdot \sum(Y)))}{\sqrt{(n \cdot \sum(X^2) - \sum(X)^2) \cdot (n \cdot \sum(Y^2) - \sum(Y)^2)}}$$

dove

r = Coefficiente di correlazione

n = Numero di osservazioni

Nel caso dei titoli finanziari la correlazione è un valore importante in quanto ci permette di effettuare predizioni sui trend futuri e quindi gestire il rischio del nostro portfolio.

	BAC	FB	GOOG	JPM	LMT	RTX
BAC	1.000000	0.282176	0.422409	0.884091	0.428257	0.588724
FB	0.282176	1.000000	0.500472	0.283136	0.251509	0.255813
GOOG	0.422409	0.500472	1.000000	0.431211	0.362786	0.426780
JPM	0.884091	0.283136	0.431211	1.000000	0.471824	0.638773
LMT	0.428257	0.251509	0.362786	0.471824	1.000000	0.556649
RTX	0.588724	0.255813	0.426780	0.638773	0.556649	1.000000

Figura 40: Matrice di correlazione sui rendimenti

Dalla matrice di correlazione a figura 40 notiamo che i titoli più correlati sono JPM e BAC con un valore di correlazione pari a 0.88. I titoli meno correlati sono invece FB e LMT con un valore di correlazione pari a 0.251.

2.7 Grafico sulla correlazione dei titoli nel tempo e scatter plot

2.7.1 Titoli tecnologici FB/GOOG

Date	FB	GOOG
2020-01-31	0.072747	0.071436
2020-02-29	-0.042180	0.019112
2020-03-31	-0.201236	-0.188317
2020-04-30	0.067908	0.038494
2020-05-31	0.223418	0.119090
2020-06-30	0.071565	0.036155
2020-07-31	0.030879	0.057668
2020-08-31	0.123631	0.021388
2020-09-30	-0.008311	-0.017718
2020-10-31	0.012114	0.018619
2020-11-30	0.024188	0.130280
2020-12-31	0.000812	0.013687

Figura 41: Ritorni logaritmici mensili di FB/GOOG per l'anno 2020

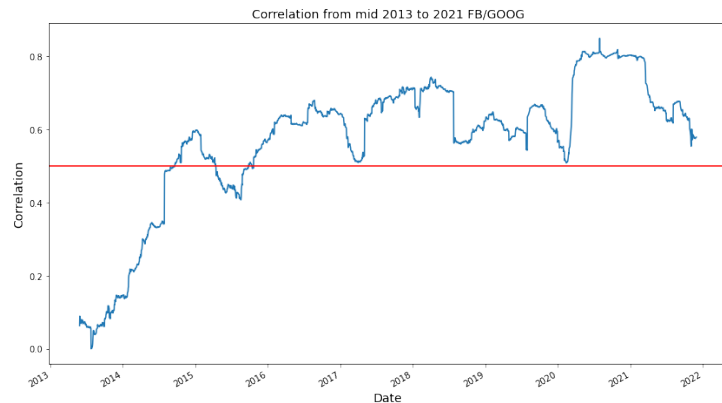


Figura 42: Correlazione nel tempo di FB/GOOG

Dal grafico in figura 42 si evince che c'è sempre stata tra i due titoli una correlazione positiva, dopo il 2015 viene mostrato come la correlazione supera 0.5 e per tutta la durata successiva rimane oltre questo valore. La tabella a figura 41 la forte correlazione in quanto in tutto l'anno i rendimenti hanno avuto lo stesso andamento.

Lo scatter plot a figura 42 evidenzia e conferma la forte correlazione tra i due titoli (seppur ci sono alcuni punti lontani dal centro).

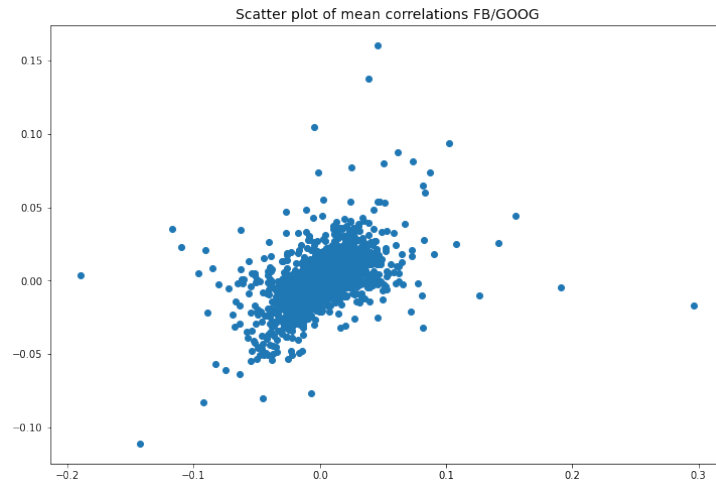


Figura 43: Dispersione rendimenti FB/GOOG

2.7.2 Titoli militari RTX/LMT

	LMT	RTX
Date		
2017-01-31	-0.004217	0.011036
2017-02-28	0.026029	0.008094
2017-03-31	0.035454	0.012663
2017-04-30	0.004365	0.017049
2017-05-31	0.015709	0.061372
2017-06-30	0.026246	0.004578
2017-07-31	0.027679	0.006915
2017-08-31	0.052033	-0.033448
2017-09-30	0.015584	-0.034925
2017-10-31	0.032425	0.051791
2017-11-30	-0.004186	-0.000643
2017-12-31	0.017443	0.054748

Figura 44: Ritorni logaritmici mensili di RTX/LMT per l'anno 2017

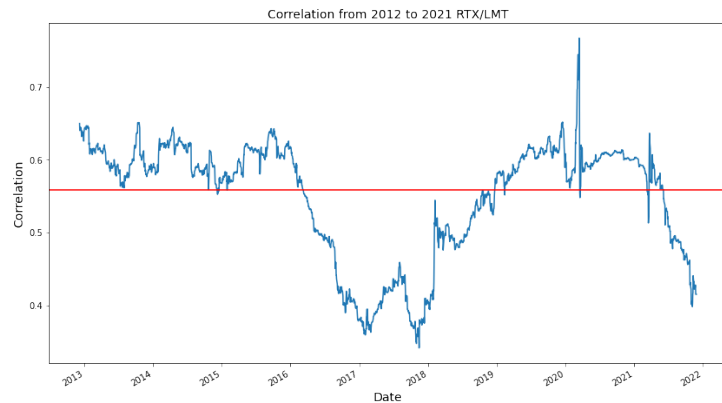


Figura 45: Correlazione nel tempo di RTX/LMT

Dal grafico in figura 45 viene mostrato come ci sono stati nel tempo momenti di correlazione alta, e momenti dove c'è stato un crollo nella correlazione, nel dettaglio fino al 2016 la correlazione è rimasta stabile tra 0.55 e 0.65, dal 2016 c'è stato un notevole crollo che ha permesso raggiungere valori di correlazione inferiori a 0.4, verso il 2018 la correlazione è ritornata ai valori prima del crollo, con addirittura uno spike poco dopo il 2020, presumibilmente a causa delle crisi finanziaria del 2020.

Nella tabella a figura 44 vengono mostrati i rendimenti logaritmici per l'anno 2017, si può notare come siano nettamente meno simili i rendimenti tra i due titoli, a figura 46 viene mostrato invece i rendimenti per l'anno 2020, in questo caso si nota come i rendimenti sono più "vicini".

Lo scatter plot a figura 47 dimostra come la correlazione sia più debole rispetto ai titoli precedenti.

	LMT	RTX
Date		
2020-01-31	0.092100	0.033832
2020-02-29	0.004387	-0.021824
2020-03-31	-0.195262	-0.300745
2020-04-30	0.100500	-0.053140
2020-05-31	0.007897	-0.041124
2020-06-30	0.025524	0.113340
2020-07-31	-0.038608	-0.070383
2020-08-31	0.055741	0.012215
2020-09-30	0.002273	-0.019823
2020-10-31	-0.030346	-0.014668
2020-11-30	-0.015057	0.127738
2020-12-31	-0.029530	0.079163

Figura 46: Ritorni logaritmici mensili di RTX/LMT per l'anno 2020

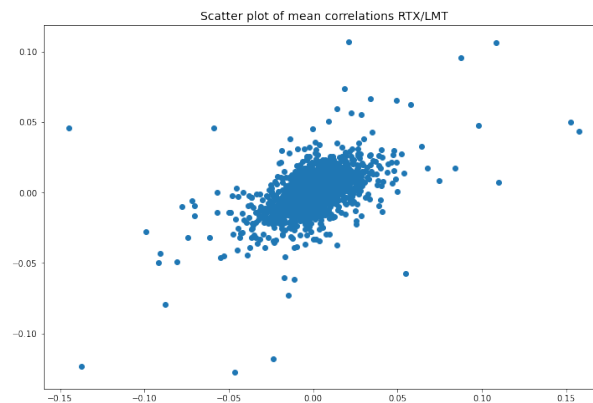


Figura 47: Dispersione rendimenti RTX/LMT

2.7.3 Titoli bancari BAC/JPM

	BAC	JPM
Date		
2020-01-31	-0.001261	0.008060
2020-02-29	-0.026479	-0.019927
2020-03-31	-0.312586	-0.275772
2020-04-30	-0.013475	-0.039523
2020-05-31	0.018518	-0.005196
2020-06-30	0.110186	0.099117
2020-07-31	-0.051169	-0.029457
2020-08-31	0.076336	0.031229
2020-09-30	-0.029254	-0.014791
2020-10-31	-0.021067	0.026514
2020-11-30	0.100144	0.134992
2020-12-31	0.096068	0.076549

Figura 48: Ritorni logaritmici mensili di BAC/JPM per l'anno 2020

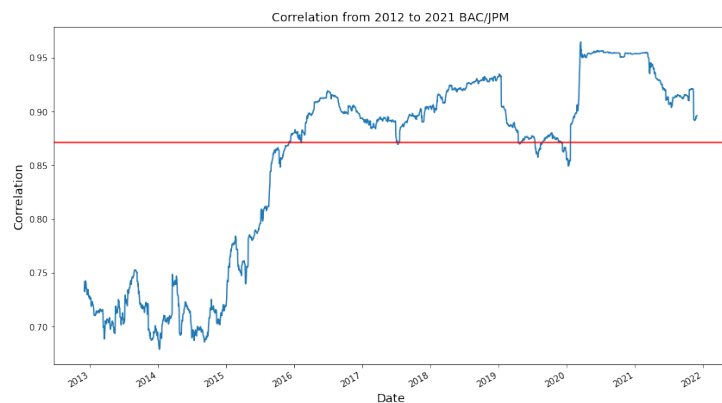


Figura 49: Correlazione nel tempo di BAC/JPM

Dal grafico a figura 49 si nota come la correlazione tra i due titoli sia sempre stata piuttosto alta, con valori mai sotto 0.6. Dal 2015 viene evidenziato una notevole crescita nella correlazione che dopo il 2016 non raggiungerà valori inferiori a 0.85.

Dallo scatter plot mostrato a figura 50 viene confermata la forte correlazione.

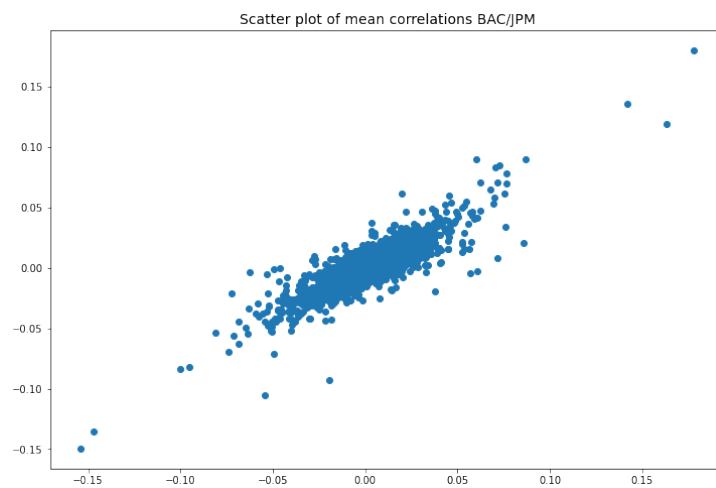


Figura 50: Dispersione rendimenti BAC/JPM

3 Analisi di previsione

Per la costruzione del modello di previsione ho deciso di utilizzare il modello di analisi statistica chiamato **ARIMA**²⁶ (**AutoRegressive Integrated Moving Average**). Questo modello utilizza i dati delle serie temporali per permettere di capire meglio il dataset e per effettuare predizioni su trend futuri.

Questo sistema è una forma di *analisi di regressione* che calcola la forza di una variabile aleatoria indipendente relativamente ai cambiamenti di altre variabili. L'obiettivo di questo modello è di effettuare predizioni sul futuro delle securities esaminando le differenze tra i valori nelle serie temporali invece di utilizzare valori attuali.

Possiamo evidenziare le componenti fondamentali di ARIMA scomponendo il suo acronimo:

- *Autoregression (AR)*: Si riferisce ad un modello che mostra una variabile che cambia in base alla regressione del suo passato o dei suoi valori precedenti.
- *Integrated (I)*: Rappresenta la differenziazione di osservazioni grezze per permettere alla serie temporale di diventare stazionaria. (permettendo la applicazione della auto regressione e media mobile *ARMA*).
- *Moving average (MA)*: Incorpora la dipendenza tra una osservazione ed l'errore residuo dai modelli di media mobile applicato alle osservazioni passate.

Nel modello ARIMA ogni componente visto qui sopra funziona come parametro avente una notazione standard, un esempio di notazione standard sarebbe ARIMA con p, d e q , dove valori interi sostituiscono i parametri per indicare la tipologia del modello ARIMA utilizzato. I parametri possono essere definiti come:

- p : Il numero di osservazioni passate nel modello.
- d : Il numero di volte che le osservazioni grezze sono state differenziate, chiamato anche grado di differenziazione.
- q : la dimensione della finestra relativa alla media mobile, chiamato anche ordine della media mobile.

Impostando i parametri visti qui sopra possiamo ottenere dei *casi particolari*, utili per le analisi:

- ARIMA(0, 0, 0): Rumore bianco.
- ARIMA(0, 1, 0) senza costante: Passeggiata aleatoria²⁷.
- ARIMA(p , 0, q): ARMA(p , q).
- ARIMA(p , 0, 0): modello AR(p).
- ARIMA(0, 0, q): modello MA(q).
- ARIMA(0, 1, 2): modello "smorzato" di Holt.²⁸
- ARIMA(0, 1, 1): senza costante: modello SES.
- ARIMA(0, 2, 2): metodo lineare di Holt²⁸, con errori aggiuntivi.

Una delle più note debolezze del modello ARIMA nel contesto finanziario è la incapacità di catturare il clustering della volatilità che si osserva nella maggior parte degli asset finanziari.

Per questo progetto al fine di identificare i parametri migliori p, d e q da applicare è stata utilizzata la libreria *pmdarima*²⁹, che grazie alla funzione `auto_arima` mediante una ricerca a griglia riesce a trovare valori che minimizzano il valore AIC³⁰.

3.1 Modello di previsione per Meta (FB)

Per Meta, è stato identificato come valore dei parametri migliore $p = 3, d = 1, q = 1$ con un errore AIC pari a 1836.767 (figura 52).

Dalla predizione a figura 51 si nota che data la alta volatilità del titolo ci sono stati momenti dove il prezzo predetto è stato molto lontano da quello reale, come per esempio nel 2018 o verso la fine nell'anno

²⁶<https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp>

²⁷Serie con passi in direzioni casuali, https://it.wikipedia.org/wiki/Passeggiata_aleatoria

²⁸<https://otexts.com/fpp3/holt.html>

²⁹<http://alkaline-ml.com/pmdarima/1.8.3/about.html>

³⁰Stimatore per gli errori di predizione, https://en.wikipedia.org/wiki/Akaike_information_criterion

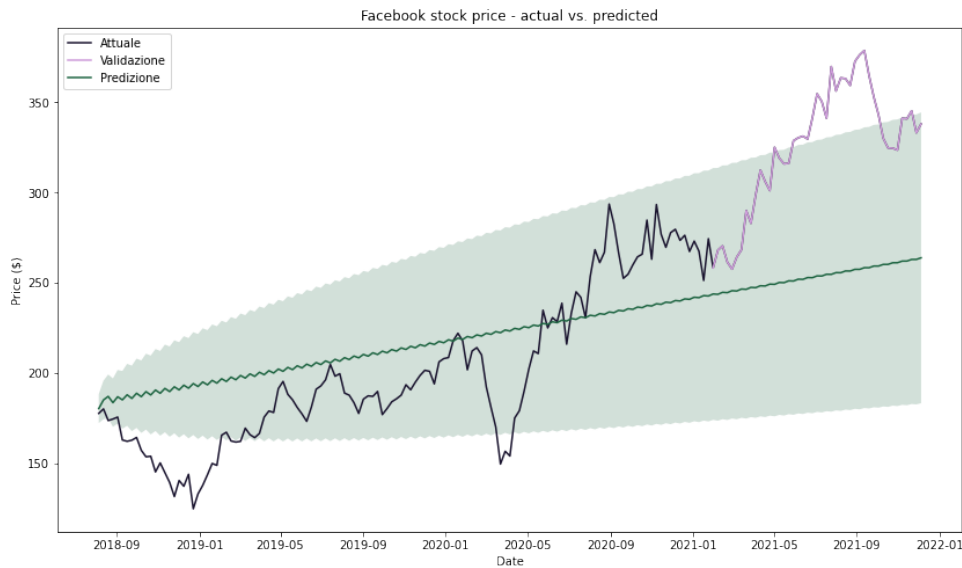


Figura 51: Predizione sul prezzo di FB usando ARIMA

del 2021. Tuttavia nel periodo centrale la predizione è stata sempre abbastanza vicina al prezzo predetto, rimanendo nell'intervallo di confidenza.

A figura 53 si può prendere visione a cadenza mensile nel periodo finale di 10 mesi la differenza tra il prezzo reale e quello predetto, si nota come sia molto distante il vero prezzo a causa della elevata volatilità.

SARIMAX Results

Dep. Variable:	y	No. Observations:	324
Model:	SARIMAX(3, 1, 1)	Log Likelihood	-912.383
Date:	Tue, 24 May 2022	AIC	1836.767
Time:	16:34:12	BIC	1859.433
Sample:	0	HQIC	1845.815
	- 324		
Covariance Type:	opg		

Figura 52: Parametri di ARIMA identificati per FB

	Real Price	Prediction
Date		
2021-01-31	258.329987	243.891500
2021-02-28	257.619995	245.708086
2021-03-31	283.019989	247.525226
2021-04-30	301.130005	249.342899
2021-05-31	328.730011	251.057075
2021-06-30	341.369995	252.898650
2021-07-31	369.790009	254.739760
2021-08-31	372.630005	257.528447
2021-09-30	352.959991	259.348681
2021-10-31	323.570007	261.180255
2021-11-30	333.119995	263.019499
2021-12-31	338.029999	263.901003

Figura 53: Prezzo reale vs predizione FB

3.2 Modello di previsione per Alphabet (GOOG)

Per Alphabet, è stato identificato come valore dei parametri migliore $p = 3, d = 1, q = 2$ con un errore AIC pari a 3107.301 (figura 55).

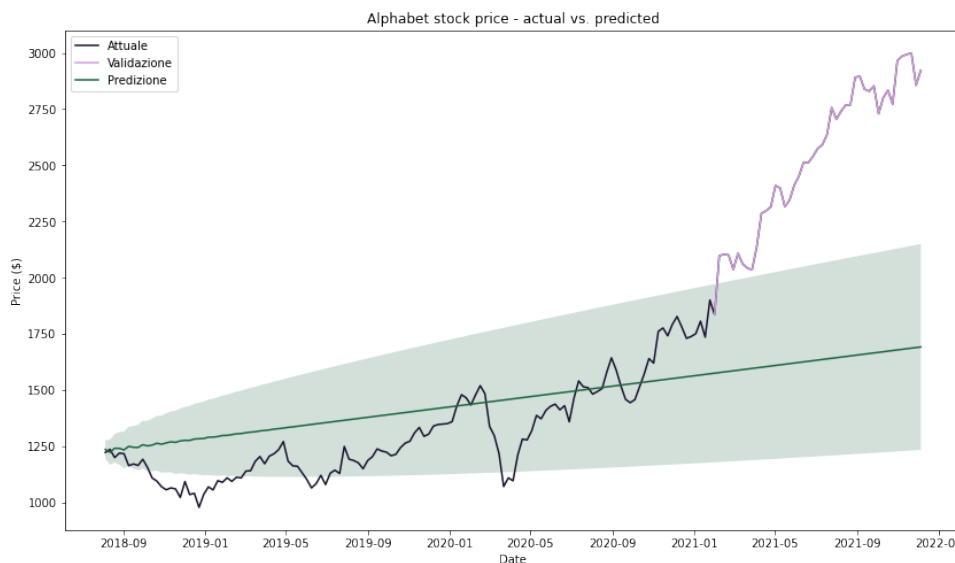


Figura 54: Predizione sul prezzo di GOOG usando ARIMA

Dal grafico di predizione a figura 54 si nota come il prezzo predetto sia stato vicino a quello reale solo nella parte centrale del grafico, verso la fine a causa della elevata volatilità si nota come il prezzi si sia distaccato enormemente. Analizzando la tabella a figura 56 con il prezzo reale e quello predetto nel periodo di valutazione si può notare come la differenza sia parecchio elevata, rendendo evidente le debolezze di questo modello statistico.

SARIMAX Results

Dep. Variable:	y	No. Observations:	348
Model:	SARIMAX(3, 1, 2)	Log Likelihood	-1546.650
Date:	Tue, 24 May 2022	AIC	3107.301
Time:	16:39:05	BIC	3134.246
Sample:	0	HQIC	3118.029
	- 348		
Covariance Type:	opg		

Figura 55: Parametri di ARIMA identificati per GOOG

	Real Price	Prediction
Date		
2021-01-31	1835.739990	1575.988479
2021-02-28	2036.859985	1586.587168
2021-03-31	2035.550049	1597.185860
2021-04-30	2315.300049	1607.784465
2021-05-31	2411.560059	1621.032806
2021-06-30	2539.899902	1631.631479
2021-07-31	2756.320068	1642.230140
2021-08-31	2891.010010	1655.478467
2021-09-30	2852.659912	1666.077117
2021-10-31	2965.409912	1679.325447
2021-11-30	2856.120117	1689.924103
2021-12-31	2922.280029	1692.573771

Figura 56: Prezzo reale vs predizione GOOG

3.3 Modello di previsione per Raytheon (RTX)

Per Raytheon, è stato identificato come valore dei parametri migliore $p = 1, d = 1, q = 1$ con un errore AIC pari a 1201.826 (figura 58).

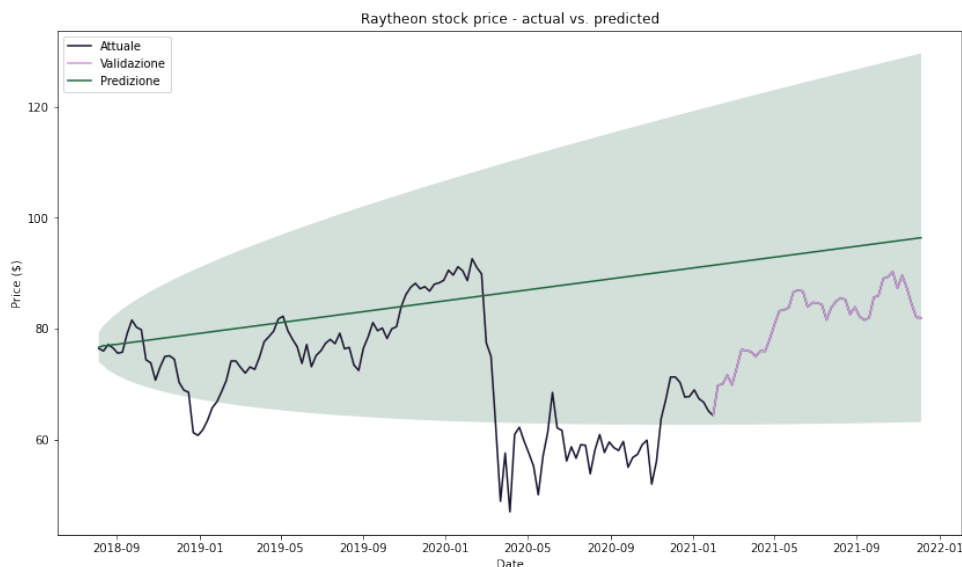


Figura 57: Predizione sul prezzo di RTX usando ARIMA

Nel caso di RTX si nota dal grafico a figura 57 come nel primo periodo il prezzo predetto è stato simile a quello reale, tuttavia dal 2020 si è visto un improvviso distaccamento, causato presumibilmente dalla crisi finanziaria del 2020, verso il 2021 nel periodo di validazione si è notato come dimostrato dalla tabella a figura 59 che il prezzo è tornato ad avvicinarsi al quello predetto.

SARIMAX Results

Dep. Variable:	y	No. Observations:	348
Model:	SARIMAX(1, 1, 1)	Log Likelihood	-596.913
Date:	Tue, 24 May 2022	AIC	1201.826
Time:	16:49:36	BIC	1217.224
Sample:	0	HQIC	1207.957
	- 348		
Covariance Type:	opg		

Figura 58: Parametri di ARIMA identificati per RTX

	Real Price	Prediction
Date		
2021-01-31	64.404472	91.438565
2021-02-28	69.908676	91.889467
2021-03-31	75.910011	92.340370
2021-04-30	78.211502	92.791272
2021-05-31	86.663994	93.354900
2021-06-30	84.719887	93.805802
2021-07-31	83.830887	94.256704
2021-08-31	83.931099	94.820332
2021-09-30	85.660828	95.271234
2021-10-31	87.331573	95.834862
2021-11-30	82.050888	96.285764
2021-12-31	81.991577	96.398490

Figura 59: Prezzo reale vs predizione RTX

3.4 Modello di previsione per Lockheed Martin (LMT)

Per Lockheed Martin, è stato identificato come valore dei parametri migliore $p = 0, d = 1, q = 2$ con un errore AIC pari a 1996.251 (figura 61).

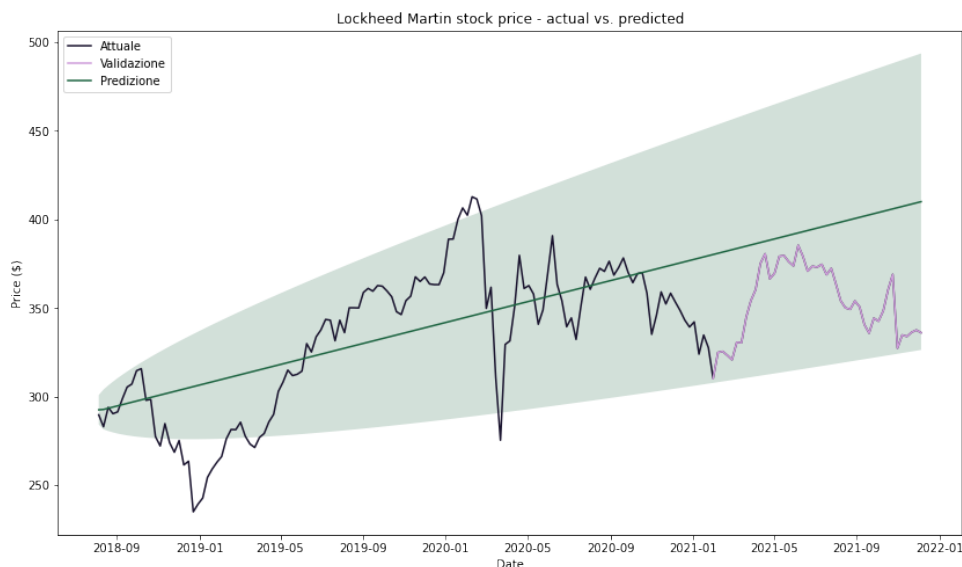


Figura 60: Predizione sul prezzo di LMT usando ARIMA

Nel grafico di predizione a figura 60 è reso evidente come in alcuni momenti la predizione sul prezzo è stata vicina a quello reale, seppur ci sono stati due momenti dove il prezzo è uscito dall'intervallo di confidenza. Dalla tabella a figura 62 si può notare nel periodo di validazione come il prezzo verso la metà si è avvicinato a quello predetto.

SARIMAX Results

Dep. Variable:	y	No. Observations:	348
Model:	SARIMAX(0, 1, 2)	Log Likelihood	-994.125
Date:	Tue, 24 May 2022	AIC	1996.251
Time:	16:57:17	BIC	2011.648
Sample:	0	HQIC	2002.381
	- 348		
Covariance Type:	opg		

Figura 61: Parametri di ARIMA identificati per LMT

	Real Price	Prediction
Date		
2021-01-31	310.229340	380.194987
2021-02-28	320.796753	382.905049
2021-03-31	354.270355	385.615110
2021-04-30	366.490265	388.325172
2021-05-31	373.768280	391.712750
2021-06-30	373.582489	394.422811
2021-07-31	372.369812	397.132873
2021-08-31	354.004120	400.520451
2021-09-30	344.281616	403.230512
2021-10-31	327.339417	406.618090
2021-11-30	337.583557	409.328151
2021-12-31	336.115906	410.005667

Figura 62: Prezzo reale vs predizione LMT

3.5 Modello di previsione per Bank of America (BAC)

Per Bank of America, è stato identificato come valore dei parametri migliore $p = 1, d = 1, q = 1$ con un errore AIC pari a 584.994 (figura 64).

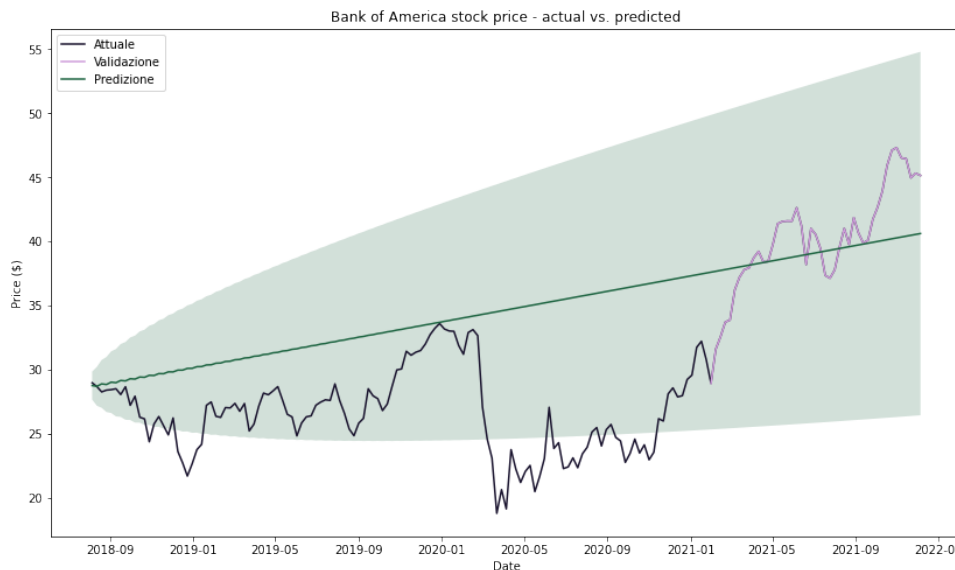


Figura 63: Predizione sul prezzo di BAC usando ARIMA

Nel grafico a figura 63 viene evidenziato come la predizione è stata vicina al prezzo reale nel periodo 2019-2020, successivamente c'è stato un grosso distacco sempre come visto precedentemente a causa della crisi finanziaria del 2020. verso il 2021 nel periodo di validazione come evidenziato dalla tabella a figura 65, il prezzo si è avvicinato a quello predetto, rimanendo sempre nell'intervallo di confidenza.

SARIMAX Results

Dep. Variable:	y	No. Observations:	348
Model:	SARIMAX(1, 1, 1)	Log Likelihood	-288.497
Date:	Tue, 24 May 2022	AIC	584.994
Time:	17:14:17	BIC	600.392
Sample:	0	HQIC	591.125
			- 348
Covariance Type:	opg		

Figura 64: Parametri di ARIMA identificati per BAC

	Real Price	Prediction
Date		
2021-01-31	28.948441	37.612453
2021-02-28	33.888714	37.886182
2021-03-31	37.952354	38.159924
2021-04-30	38.442947	38.433676
2021-05-31	41.592564	38.775089
2021-06-30	41.008949	39.048961
2021-07-31	37.146503	39.322828
2021-08-31	41.866177	39.665656
2021-09-30	41.734093	39.939445
2021-10-31	47.319763	40.281319
2021-11-30	45.319218	40.555164
2021-12-31	45.150856	40.623935

Figura 65: Prezzo reale vs predizione BAC

3.6 Modello di previsione per JPMorgan Chase (JPM)

Per Bank of America, è stato identificato come valore dei parametri migliore $p = 0, d = 1, q = 0$ (caso particolare di ARIMA) con un errore AIC pari a 1336.117 (figura 67).

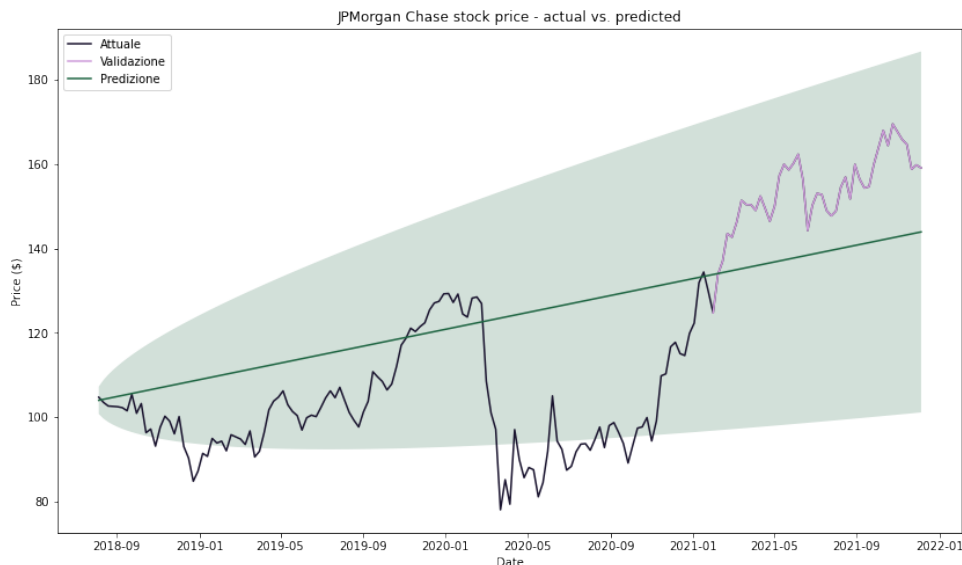


Figura 66: Predizione sul prezzo di JPM usando ARIMA

Il grafico a figura 66 si comporta in maniera molto simile a ciò che è stato visto per BAC, tale comportamento si può attribuire alla correlazione veramente alta tra i due titoli. Nella tabella a figura 68, si può vedere come nel periodo di verifica il prezzo sia sempre stato abbastanza vicino a quello predetto, e sempre nell'intervallo di confidenza (in maniera simile a BAC).

SARIMAX Results

Dep. Variable:	y	No. Observations:	348
Model:	SARIMAX(0, 1, 0)	Log Likelihood	-666.058
Date:	Tue, 24 May 2022	AIC	1336.117
Time:	17:27:36	BIC	1343.815
Sample:	0	HQIC	1339.182
	- 348		
Covariance Type:	opg		

Figura 67: Parametri di ARIMA identificati per JPM

	Real Price	Prediction
Date		
2021-01-31	124.743980	133.821034
2021-02-28	142.679504	134.738259
2021-03-31	150.357834	135.655484
2021-04-30	146.464935	136.572709
2021-05-31	160.166458	137.719241
2021-06-30	150.229202	138.636466
2021-07-31	147.750320	139.553691
2021-08-31	159.922256	140.700222
2021-09-30	159.912445	141.617447
2021-10-31	167.635147	142.763978
2021-11-30	159.780792	143.681203
2021-12-31	159.090073	143.910510

Figura 68: Prezzo reale vs predizione JPM

3.7 Extra: Forecasting usando tecniche Deep Learning (MLP)

Nelle sezioni qui sopra è stata ottenuta una previsione sul prezzo mediante modelli statistici, in questa sezione per il titolo FB viene fatta una previsione con il sistema **MLP** (Multilayer **p**erceptrons) e verrà messa a confronto con la previsione ottenuta tramite modello **ARIMA** del punto 3.1.

I Multilayer perceptrons sono una delle architetture base dei neural networks, sono composti da tre componenti:

- **input layer:** vettore di *features*.
- **hidden layers:** Ogni livello "nascosto" è composto da N neuroni.
- **output layer:** Questo livello è l'output del network, dipende dalla task (regressione/classificazione).

è stata utilizzata la libreria *pytorch* per il forecasting.

Rispetto ad **ARIMA** uno dei benefici dell'utilizzo del deep learning è che non si fanno assunzioni sui dati soggiacenti, inoltre non c'è alcun bisogno che la serie sia stazionaria. MLP è un modello di neural network feedforward, è più semplice rispetto ad altre tipologie di modelli deep learning, per questo motivo è sempre raccomandato prima di utilizzare modelli più sofisticati.

4 Strategie di trading e backtesting

Il backtesting ci permette di valutare la performance di una strategia di trading utilizzando indicatori euristici o tecnici e applicandoli a dati storici.

Grazie al backtesting possiamo simulare la strategia di trading utilizzando dati storici e quindi generare risultati ed analizzare il rischio oltre ad il potenziale profitto prima di investire vero capitale.

Nel contesto di questo progetto, è stato creato un algoritmo di trading usando le bande di bollinger per il signaling.

4.1 Strategia buy/sell mediante Bollinger's Bands

Le bande di Bollinger sono un metodo statistico per derivare informazioni sul prezzo e volatilità di un asset nel tempo.

Per ottenere le bande di Bollinger abbiamo bisogno di calcolare la media mobile e la deviazione standard della serie temporale, utilizzando una finestra specifica (tipicamente 20 giorni). Successivamente dobbiamo importare la banda superiore/inferiore a K volte (tipicamente 2) la deviazione standard mobile sopra o sotto la media mobile.

Per la strategia di trading scelta è stata impostata come finestra 20 giorni e come fattore di deviazione 2.

La strategia prevede si creare un segnale di *buy* nel caso di crossover da parte del prezzo della banda inferiore, per quanto riguarda invece il segnale *sell* ci deve essere un crossover nella banda superiore.

Di seguito il codice per la impostazione dei signal:

```
# Impostazione del buy signal
self.buy_signal = bt.ind.CrossOver(self.datas[0],
                                   self.b_band.lines.bot)

# Impostazione del sell signal
self.sell_signal = bt.ind.CrossOver(self.datas[0],
                                    self.b_band.lines.top)
```

Se è presente già una posizione aperta non verranno aperte altre posizioni in caso di ulteriori cross-over, la posizione viene chiusa automaticamente appena viene generato un segnale di sell.

L'algoritmo utilizzato non prevede short-selling, a causa di ciò per effettuare una operazione di sell deve essere presente una posizione aperta sul titolo, di seguito il codice che effettua l'ordine di acquisto o vendita:

```
if not self.position:
    if self.buy_signal > 0:
        size = int(self.broker.getcash() / self.datas[0].open)
        self.log(f'BUY CREATED — Size: {size}, \
                  Cash: {self.broker.getcash():.2f}, \
                  Open: {self.data_open[0]}, \
                  Close: {self.data_close[0]}')
        self.buy(size=size)
    else:
        if self.sell_signal < 0:
            self.log(f'SELL CREATED — Size: {self.position.size}')
            self.sell(size=self.position.size)
```

4.2 Backtesting su titolo Meta (FB)

Per effettuare il backtesting della strategia scelta, è stato preso in considerazione il titolo Meta (FB) nell'anno 2020, si può trovare il grafico a figura 69, come valore iniziale del portfolio sono stati impostati 10000\$ e per rendere i dati più reali è stata impostata anche una percentuale di commissione pari a 0.001%.

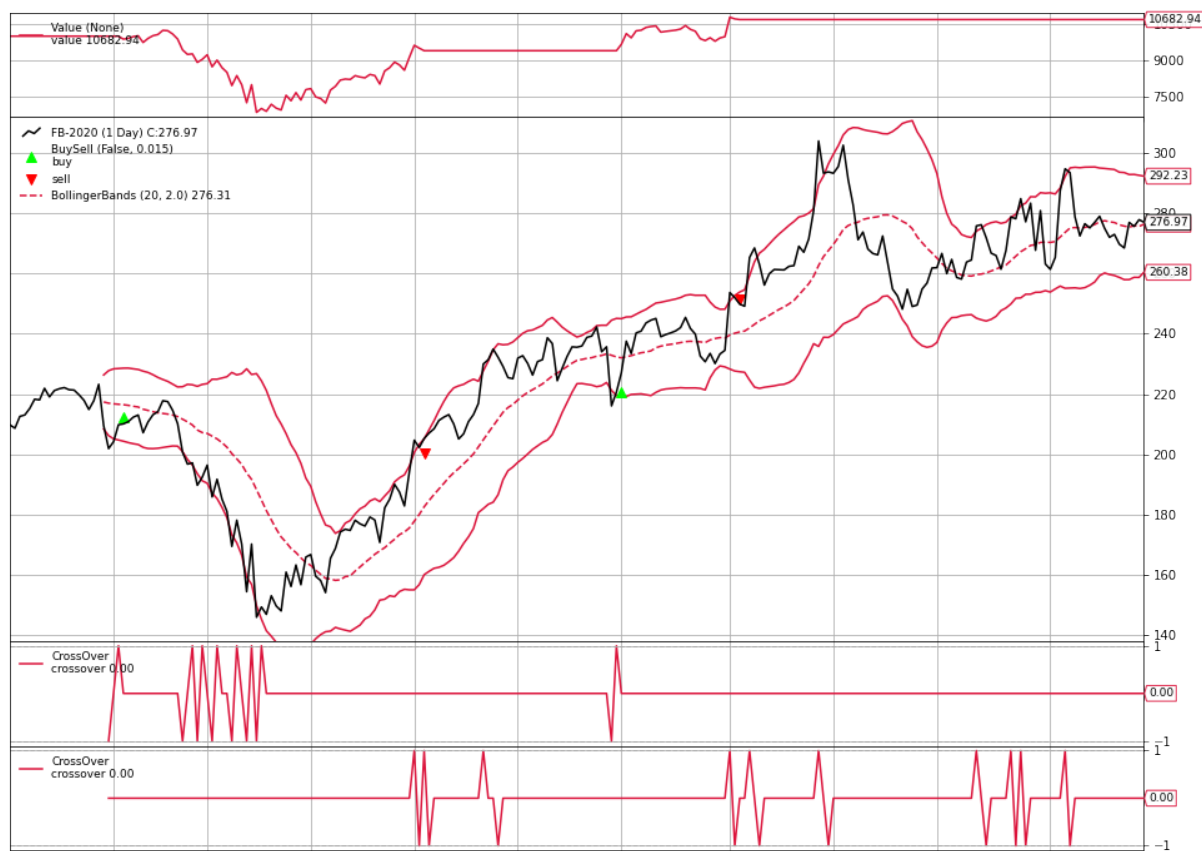


Figura 69: Backtesting della strategia BB su FB

Starting Portfolio Value: 10000.00
 2020-02-05, BUY CREATED --- Size: 47, Cash: 10000.00, Open: 212.51, Close: 210.11
 2020-02-05, BUY EXECUTED --- Price: 212.51, Cost: 9987.97, Commission: 9.99
 2020-05-04, SELL CREATED --- Size: 47
 2020-05-04, SELL EXECUTED --- Price: 200.20, Cost: 9987.97, Commission: 9.41
 2020-05-04, OPERATION RESULT --- Gross: -578.57, Net: -597.97
 2020-06-30, BUY CREATED --- Size: 42, Cash: 9402.03, Open: 220.59, Close: 227.07
 2020-06-30, BUY EXECUTED --- Price: 220.59, Cost: 9264.78, Commission: 9.26
 2020-08-04, SELL CREATED --- Size: 42
 2020-08-04, SELL EXECUTED --- Price: 251.56, Cost: 9264.78, Commission: 10.57
 2020-08-04, OPERATION RESULT --- Gross: 1300.74, Net: 1280.91
 Final Portfolio Value: 10682.94

Figura 70: backtrader log for FB

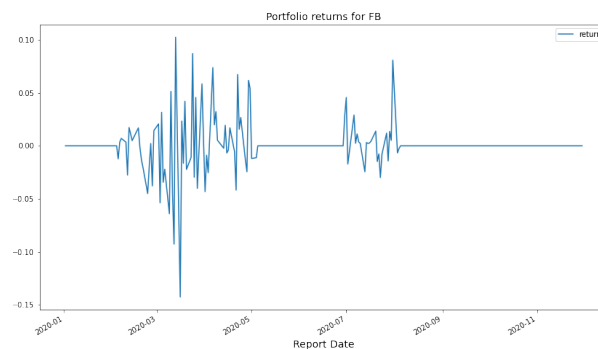


Figura 71: Backtesting portfolio returns

Dai log di **backtrader** a figura 70 si può vedere come la strategia è riuscita ad avere un profitto, passando da 10000.00\$ a 10682,94\$ marcando un profitto di 682,94\$ nel corso di un anno.

Analizzando i dati tecnici possiamo osservare un rendimento annuale/normalizzato pari al 7.47%.

Dal backtest è evidente come la strategia sia lontana dall'essere perfetta, dai due grafici presentati qui sopra si nota come nel primo *sell* ci sia stata una perdita sul valore del portfolio, al secondo *sell* tuttavia la strategia è riuscita ad ottenere un significativo rendimento che ha coperto la perdita del primo *sell*.

5 Capital Asset Pricing Model - CAPM

Il CAPM (Capital Asset Pricing Model) è un modello che rappresenta la relazione tra i rendimenti attesi di un indice rischioso ed il rischio di mercato (chiamato anche rischio sistematico).

Possiamo rappresentare il CAPM mediante la equazione:

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

dove

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

r_f = ratio risk-free

$E(r_m)$ = Rendimento atteso del mercato

β = coefficiente beta

5.1 Calcolo dell'indice beta

L'indice beta si può interpretare come il livello di sensitività dei rendimenti di un asset relativamente al mercato.

- $\beta \leq -1$: L'asset si muove nella direzione opposta al benchmark e in un ammontare superiore rispetto al negativo del benchmark.
- $-1 < \beta < 0$: L'asset si muove nella direzione opposta al benchmark.
- $\beta = 0$: Non esiste correlazione tra il movimento del prezzo dell'asset e il benchmark di mercato.
- $0 < \beta < 1$: L'asset si muove nella stessa direzione del mercato, ma con un ammontare inferiore. (per esempio uno stock di una compagnia che non è molto suscettibile a fluttuazioni giornaliere).
- $\beta = 1$: L'asset ed il mercato si muovono nella stessa direzione e con lo stesso ammontare.
- $\beta \geq 1$: L'asset si muove nella stessa direzione del mercato, ma l'ammontare è maggiore (per esempio nel caso di stock che sono molto suscettibili a cambiamenti giornalieri nelle notizie di mercato).

La equazione del CAPM può essere modificata per ottenere la formula di calcolo per beta

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

Nel calcolo di beta per tutti i titoli considerati in questo progetto verrà considerato l'indice S&P 500 (^GSPC).

Mediante il metodo delle covarianze l'indice beta per ogni titolo è:

- Meta (FB): 1.1146
- Alphabet (GOOG): 1.0138
- Raytheon (RTX): 1.3118
- Lockheed Martin(LMT): 0.835
- Bank of America(BAC): 1.5124
- JPMorgan Chase (JPM): 1.2579

5.2 Calcolo esposizione dei titoli con Fama-French

Il modello a tre fattori Fama-French è stato inventato per espandere il modello CAPM aggiungendo due fattori che aiutano a spiegare l'eccesso di ritorno di un asset o portfolio. I tre fattori considerati sono:

- **Market Factor (MKT)**: Misura l'eccesso di ritorno del mercato, uguale a quello che troviamo nel CAPM.
- **Size factor, SMB(Small Minus Big)**: Misura l'eccesso di ritorno degli stock con un market cap piccolo rispetto agli stock con un grande market cap.
- **Value factor, HML(High Minus Low)**: Misura l'eccesso di ritorno dei **value stock** rispetto ai **growth stock**. I **value stock** hanno un alto ratio book-to-market, mentre i **growth stock** sono caratterizzati da un basso ratio.

Possiamo rappresentare il modello come

$$E(r_i) - r_f = \alpha + \beta_{mkt}MKT + \beta_{smb}SMB + \beta_{hml}HML$$

dove

$E(r_i)$ = Denota il rendimento aspettato dell'asset i

r_f = ratio risk-free

α = costante di intercettamento (equivalente a 0)

Per l'utilizzo di questo modello in questo progetto sono stati scaricati i dati direttamente dal sito del Professor French³¹.

5.2.1 Esposizione di Meta (FB)

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn		R-squared:	0.194		
Model:	OLS		Adj. R-squared:	0.172		
Method:	Least Squares		F-statistic:	8.851		
Date:	Thu, 26 May 2022		Prob (F-statistic):	2.63e-05		
Time:	15:09:13		Log-Likelihood:	112.20		
No. Observations:	114		AIC:	-216.4		
Df Residuals:	110		BIC:	-205.5		
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	0.0100	0.009	1.089	0.278	-0.008	0.028
mkt	1.1343	0.231	4.910	0.000	0.676	1.592
smb	-0.0631	0.366	-0.172	0.863	-0.789	0.662
hml	-0.4505	0.286	-1.578	0.117	-1.016	0.115
=====						
Omnibus:	27.694		Durbin-Watson:	1.867		
Prob(Omnibus):	0.000		Jarque-Bera (JB):	104.260		
Skew:	0.716		Prob(JB):	2.29e-23		
Kurtosis:	7.461		Cond. No.	43.8		

Figura 72: Risultati del modello Fama-French per FB

Dai risultati del modello Fama-French a figura 72 troviamo come fattore SMB -0.0631 e come fattore HML il valore -0.4505 , l'intercept non presenta un valore significante.

³¹http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ftp/F-F_Research_Data_Factors_CSV.zip

5.2.2 Esposizione di Alphabet (GOOG)

OLS Regression Results						
=====						
Dep. Variable:	excess_rtn	R-squared:	0.392			
Model:	OLS	Adj. R-squared:	0.376			
Method:	Least Squares	F-statistic:	24.92			
Date:	Thu, 26 May 2022	Prob (F-statistic):	1.63e-12			
Time:	15:23:11	Log-Likelihood:	191.82			
No. Observations:	120	AIC:	-375.6			
Df Residuals:	116	BIC:	-364.5			
Df Model:	3					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	0.0068	0.005	1.400	0.164	-0.003	0.016
mkt	1.0504	0.122	8.637	0.000	0.810	1.291
smb	-0.5393	0.196	-2.752	0.007	-0.927	-0.151
hml	-0.1268	0.154	-0.825	0.411	-0.431	0.178
=====						
Omnibus:		6.136	Durbin-Watson:	2.133		
Prob(Omnibus):		0.047	Jarque-Bera (JB):	7.037		
Skew:		0.309	Prob(JB):	0.0296		
Kurtosis:		4.012	Cond. No.	44.4		

Figura 73: Risultati del modello Fama-French per GOOG

Dai risultati del modello Fama-French a figura 73 troviamo come fattore SMB -0.504 e come fattore HML il valore -0.1268 , anche in questo caso l'intercept non presenta un valore significante.

5.2.3 Esposizione di Raytheon (RTX)

OLS Regression Results						
Dep. Variable:	excess_rtn	R-squared:	0.606			
Model:	OLS	Adj. R-squared:	0.595			
Method:	Least Squares	F-statistic:	59.36			
Date:	Thu, 26 May 2022	Prob (F-statistic):	2.50e-23			
Time:	15:28:18	Log-Likelihood:	210.65			
No. Observations:	120	AIC:	-413.3			
Df Residuals:	116	BIC:	-402.1			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	-0.0067	0.004	-1.612	0.110	-0.015	0.002
mkt	1.2293	0.104	11.824	0.000	1.023	1.435
smb	-0.0075	0.168	-0.045	0.964	-0.339	0.324
hml	0.4470	0.131	3.401	0.001	0.187	0.707
Omnibus:	12.722	Durbin-Watson:	2.189			
Prob(Omnibus):	0.002	Jarque-Bera (JB):	26.871			
Skew:	0.361	Prob(JB):	1.46e-06			
Kurtosis:	5.203	Cond. No.	44.4			

Figura 74: Risultati del modello Fama-French per RTX

Dai risultati del modello Fama-French a figura 74 troviamo come fattore SMB -0.0075 e come fattore HML il valore 0.4470 , in questo caso l'intercept non presenta un valore significante.

5.2.4 Esposizione di Lockheed Martin (LMT)

OLS Regression Results						
Dep. Variable:	excess_rtn	R-squared:	0.383			
Model:	OLS	Adj. R-squared:	0.367			
Method:	Least Squares	F-statistic:	24.01			
Date:	Thu, 26 May 2022	Prob (F-statistic):	3.70e-12			
Time:	15:34:07	Log-Likelihood:	210.71			
No. Observations:	120	AIC:	-413.4			
Df Residuals:	116	BIC:	-402.3			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0046	0.004	1.103	0.272	-0.004	0.013
mkt	0.8730	0.104	8.402	0.000	0.667	1.079
smb	-0.5501	0.167	-3.285	0.001	-0.882	-0.218
hml	-0.0070	0.131	-0.053	0.958	-0.267	0.253
Omnibus:	4.739	Durbin-Watson:	2.122			
Prob(Omnibus):	0.094	Jarque-Bera (JB):	4.429			
Skew:	-0.469	Prob(JB):	0.109			
Kurtosis:	3.085	Cond. No.	44.4			

Figura 75: Risultati del modello Fama-French per LMT

Dai risultati del modello Fama-French a figura 75 troviamo come fattore SMB -0.5501 e come fattore HML il valore -0.0070 , in questo caso l'intercept non presenta un valore significante.

5.2.5 Esposizione di Bank of America (BAC)

OLS Regression Results						
Dep. Variable:	excess_rtn	R-squared:	0.604			
Model:	OLS	Adj. R-squared:	0.593			
Method:	Least Squares	F-statistic:	58.87			
Date:	Thu, 26 May 2022	Prob (F-statistic):	3.34e-23			
Time:	15:37:12	Log-Likelihood:	178.26			
No. Observations:	120	AIC:	-348.5			
Df Residuals:	116	BIC:	-337.4			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0073	0.005	1.356	0.178	-0.003	0.018
mkt	1.3409	0.136	9.847	0.000	1.071	1.611
smb	0.2108	0.219	0.961	0.339	-0.224	0.645
hml	1.0896	0.172	6.330	0.000	0.749	1.431
Omnibus:	10.439	Durbin-Watson:	2.198			
Prob(Omnibus):	0.005	Jarque-Bera (JB):	17.075			
Skew:	0.375	Prob(JB):	0.000196			
Kurtosis:	4.689	Cond. No.	44.4			

Figura 76: Risultati del modello Fama-French per BAC

Dai risultati del modello Fama-French a figura 76 troviamo come fattore SMB 0.2108 e come fattore HML il valore 1.0896 , in questo caso l'intercept non presenta un valore significante mentre invece il valore HML supera 1.

5.2.6 Esposizione di JPMorgan Chase (JPM)

OLS Regression Results						
Dep. Variable:	excess_rtn	R-squared:	0.691			
Model:	OLS	Adj. R-squared:	0.683			
Method:	Least Squares	F-statistic:	86.28			
Date:	Thu, 26 May 2022	Prob (F-statistic):	2.06e-29			
Time:	15:44:25	Log-Likelihood:	221.60			
No. Observations:	120	AIC:	-435.2			
Df Residuals:	116	BIC:	-424.1			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0058	0.004	1.541	0.126	-0.002	0.013
mkt	1.1199	0.095	11.802	0.000	0.932	1.308
smb	0.0163	0.153	0.107	0.915	-0.287	0.319
hml	1.0124	0.120	8.440	0.000	0.775	1.250
Omnibus:	10.230	Durbin-Watson:	2.123			
Prob(Omnibus):	0.006	Jarque-Bera (JB):	20.512			
Skew:	-0.256	Prob(JB):	3.51e-05			
Kurtosis:	4.959	Cond. No.	44.4			

Figura 77: Risultati del modello Fama-French per JPM

Dai risultati del modello Fama-French a figura 77 troviamo come fattore SMB 0.0163 e come fattore HML il valore 1.0124, in questo caso l'intercept non presenta un valore significativo mentre invece il valore HML come per BAC supera 1.

5.3 Calcolo dei rendimenti attesi usando beta

Effettuiamo il calcolo dei rendimenti attesi utilizzando la formula del modello Fama-French, presentata precedentemente, i valori trovati relativamente ai rendimenti attesi sono:

Meta (FB)

- Rendimento atteso mensile: 0.0156 o circa 1%.
- Rendimento atteso annuale: 0.1877 o circa 18%.

Alphabet (GOOG)

- Rendimento atteso mensile: 0.0139 o circa 1%.
- Rendimento atteso annuale: 0.1664 o circa 16%.

Raytheon (RTX)

- Rendimento atteso mensile: 0.0149 o circa 1%.
- Rendimento atteso annuale: 0.1783 o circa 17%.

Lockheed Martin (LMT)

- Rendimento atteso mensile: 0.0114 o circa 1%.
- Rendimento atteso annuale: 0.1366 o circa 13%.

Bank of America (BAC)

- Rendimento atteso mensile: 0.0149 o circa 1%.
- Rendimento atteso annuale: 0.1782 o circa 17%.

JPMorgan Chase (JPM)

- Rendimento atteso mensile: 0.0123 o circa 1%.
- Rendimento atteso annuale: 0.147 o circa 14%.

6 Costruzione di un portfolio

Per costruzione di un portfolio intendiamo distribuire la quota di investimento su certi asset (che siano stock, opzioni, bonds o qualsiasi altro strumento finanziario), l'obiettivo primario rimane nel bilanciare il rischio ed i potenziali profitti.

Il principio fondamentale utilizzato per la allocazione degli asset nel portfolio che stiamo creando di chiama **Modern Portfolio Theory** (MPT o analisi con media-varianza), MPT è stata creata per aiutare gli investitori nella costruzione di un portfolio che massimizza i rendimenti per un livello di rischio specificato.

MPT è legato al concetto di *diversificazione*, ciò significa che possedere diversi tipi di asset riduce il rischio, in quanto la perdita di rendimento di una particolare security ha meno impatto sulla performance di portfolio. In principio minore è la correlazione tra gli asset nel portfolio, meglio è per la diversificazione.

6.1 Costruzione del portfolio ottimale

6.1.1 Costruzione tramite simulazioni di Monte Carlo

Utilizziamo il metodo Monte Carlo per ottenere un set di portafogli ottimali (nella frontiera di efficienza), cioè:

- Con il rendimento più alto dato un livello di rischio
- Con il più basso livello di rischio dato un livello di rendimento aspettato

Il metodo Monte Carlo è stocastico (utilizza in input dati random), per questo motivo ad ogni esecuzione del codice saranno presentati risultati diversi, tale metodo è una simulazione in quanto è una rappresentazione virtuale di un problema.

Utilizzando solo dati storici

Utilizzando i dati storici (108 mesi) dei sei titoli relativi a questo progetto, eseguendo le simulazioni otteniamo come risultato il grafico a figura 78.

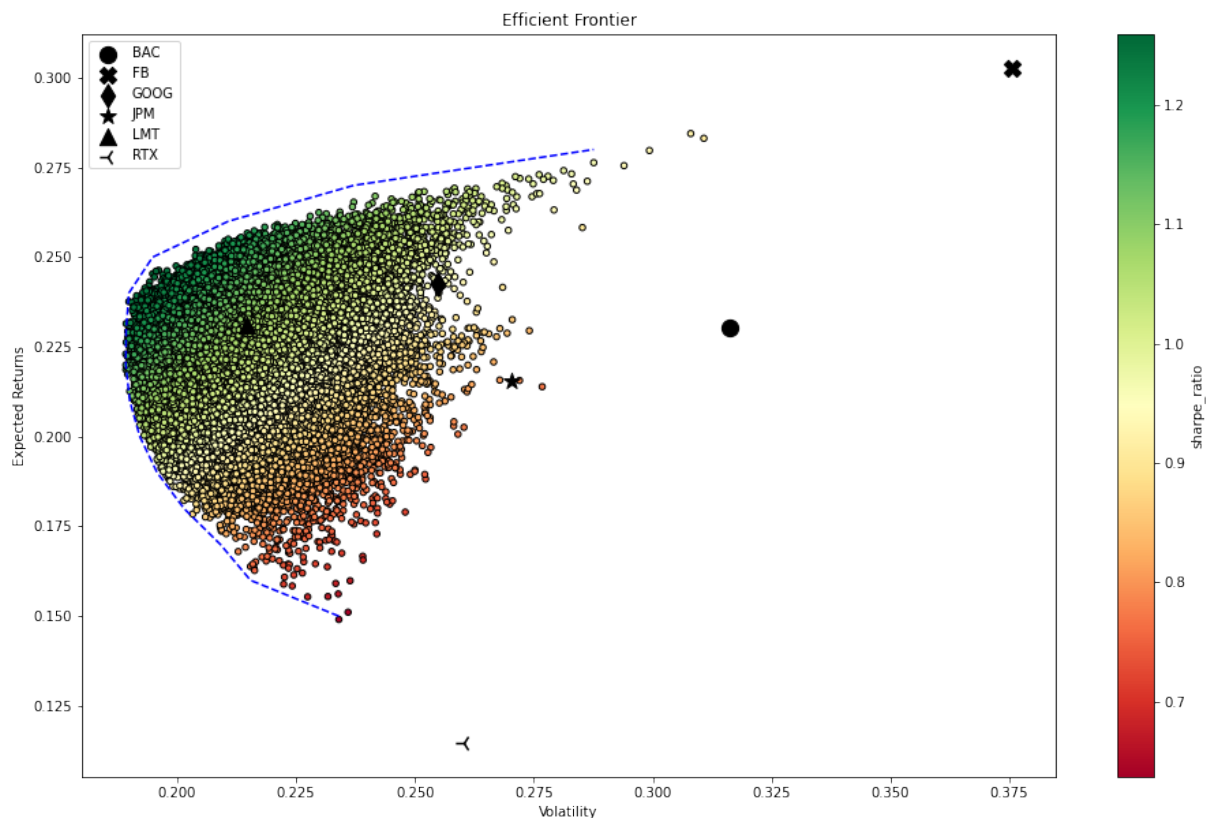


Figura 78: Grafico con tutti i portafogli generati da Monte Carlo

Dalle simulazioni eseguite identifichiamo vari portafogli che sono localizzati nella frontiera di efficienza, a figura 79 possiamo osservare lo stesso grafico ma con evidenziati due portafogli sub-ottimali, uno con il rendimento più alto tra tutti e l'altro con la varianza più bassa.

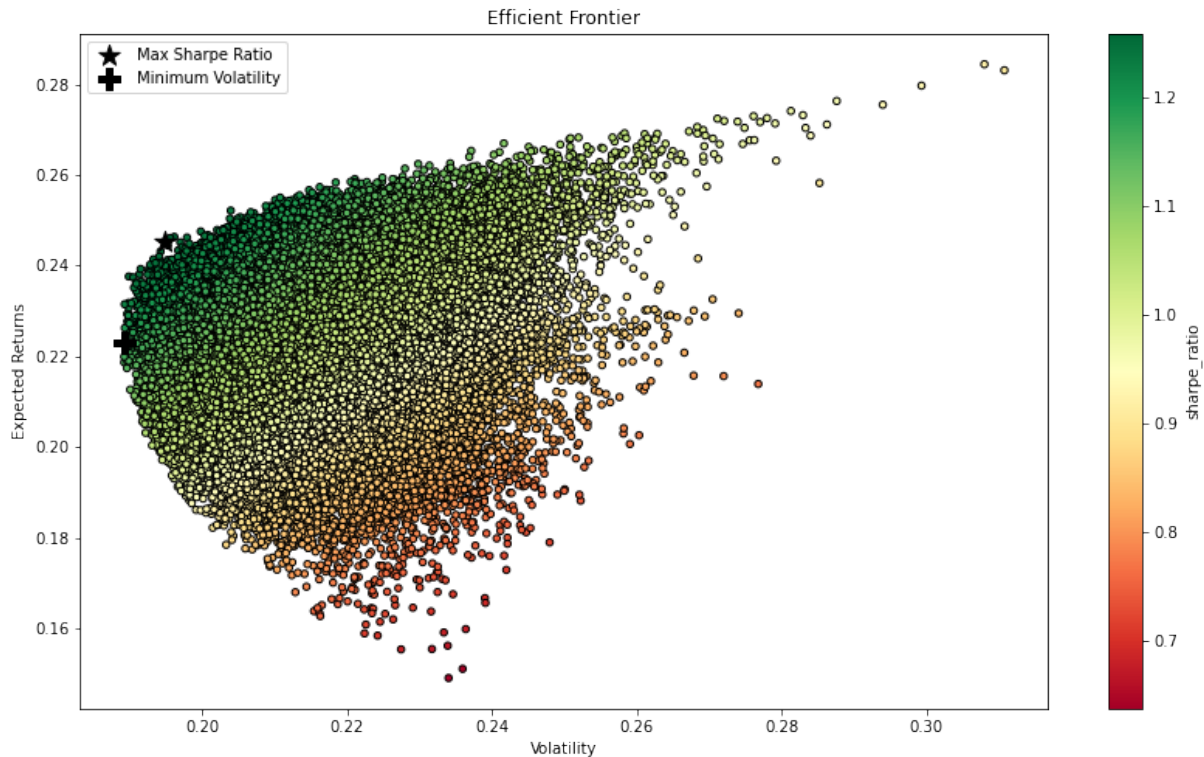


Figura 79: Grafico con evidenza dei due portafogli sub-ottimali

La composizione del portafoglio sub-ottimale con il rendimento più alto la è la seguente (figura 80).

```
Maximum Sharpe ratio portfolio ----
Performance
returns: 24.54% volatility: 19.48% sharpe_ratio: 125.94%
Weights
BAC: 1.55% FB: 15.66% GOOG: 31.05% JPM: 0.10% LMT: 51.17% RTX: 0.48%
```

Figura 80: Composizione con pesi del portafoglio sub-ottimale con rendimento più alto

Utilizzando i dati di previsione (portfolio di previsione)

Applichiamo le simulazioni di Monte Carlo ai dati di previsione identificati nel capitolo 3 (su un periodo di 108 mesi).

La ripartizione della origine dei dati è la seguente:

- Primi 80 mesi: Dati storici
- Dal mese 81 a 108: Dati di previsione

Dalle simulazioni usando oltre ai dati storici quelli di previsione otteniamo una dispersione dei portafogli notevolmente inferiore, a figura 81 abbiamo i portafogli con la frontiera di efficienza evidenziata mentre a figura 82 vengono evidenziati come per il precedente grafico i due portafogli tale per cui uno ha il rendimento più alto ed il secondo ha la volatilità più bassa.

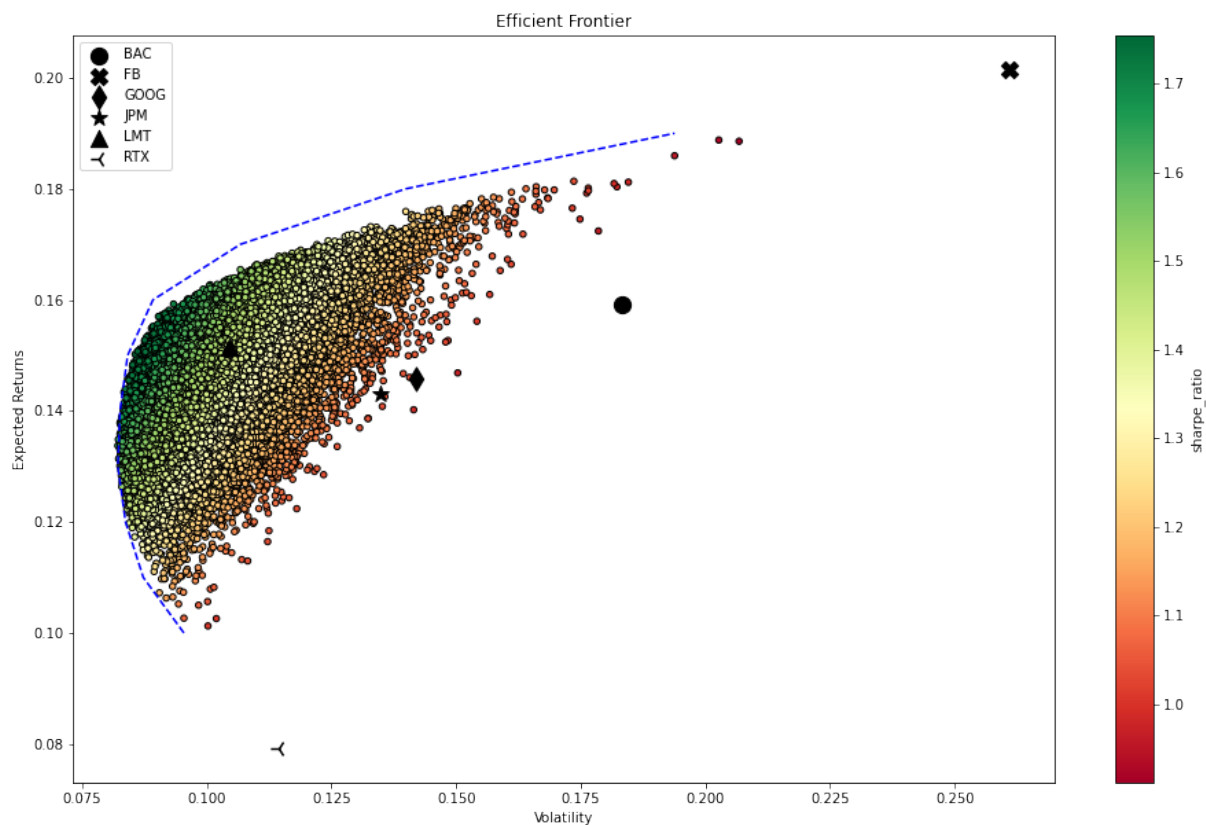


Figura 81: Grafico con tutti i portafogli generati da Monte Carlo usando dati di previsione

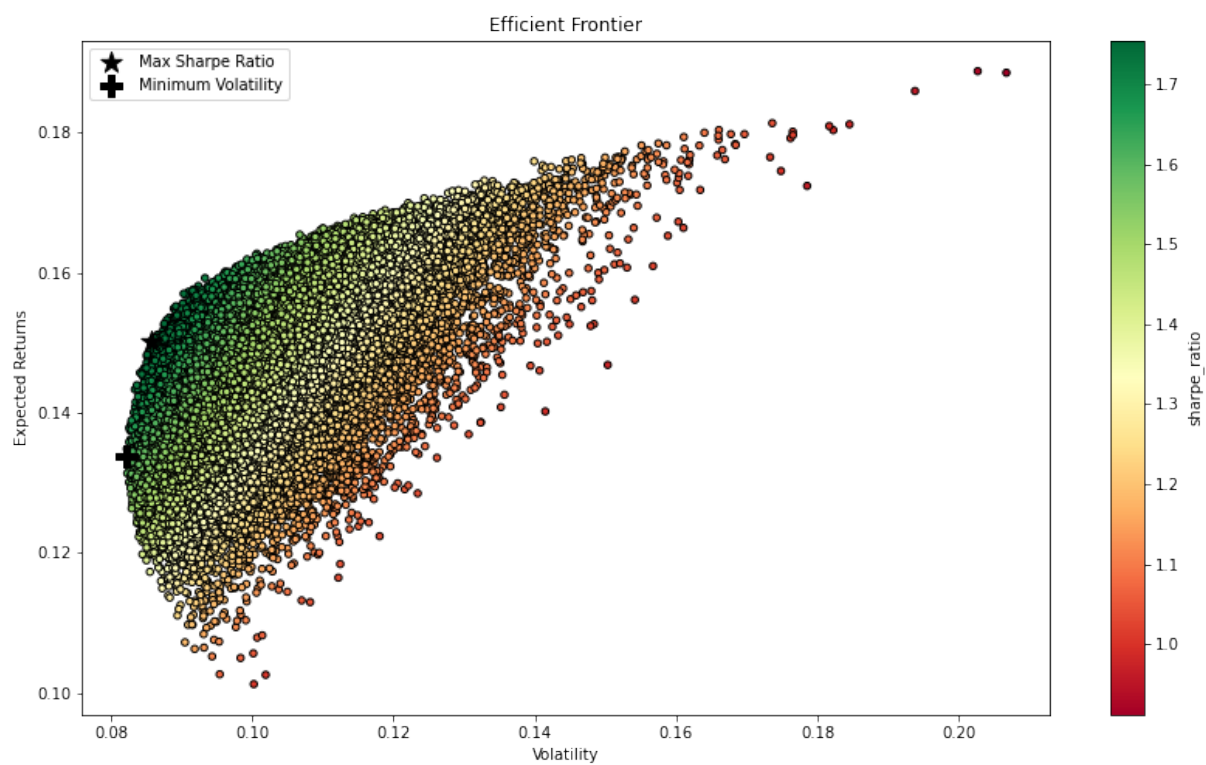


Figura 82: Composizione con pesi del portafoglio sub-ottimale con rendimento più alto usando i dati di previsione

La composizione del portafoglio sub-ottimale con il rendimento più alto costituita al mese 80 con dati di previsione per i successivi 27 mesi è la seguente (figura 83).

```
Maximum Sharpe ratio portfolio ----
Performance
returns: 15.02% volatility: 8.56% sharpe_ratio: 175.42%
Weights
BAC: 0.50% FB: 10.80% GOOG: 17.47% JPM: 28.70% LMT: 38.01% RTX: 4.52%
```

Figura 83: Composizione con pesi del portafoglio sub-ottimale con rendimento più alto usando i dati di previsione

Si può notare come l'utilizzo dei dati di regressione del modello **ARIMA** abbia reso il valore di volatilità notevolmente inferiore rispetto al portafoglio precedente a figura 80.

6.2 Beta dei portafogli ottimali

L'indice beta per i due portafogli (dati passati e dati di previsione) è stato calcolato usando come per il punto 5 l'indice S&P 500 (\sim GSPC).

6.2.1 beta del portfolio costituito solo con dati passati

Sono stati considerati tutti i dati passati con cadenza giornaliera e mediante il metodo della regressione è stato calcolato il beta (figura 84).

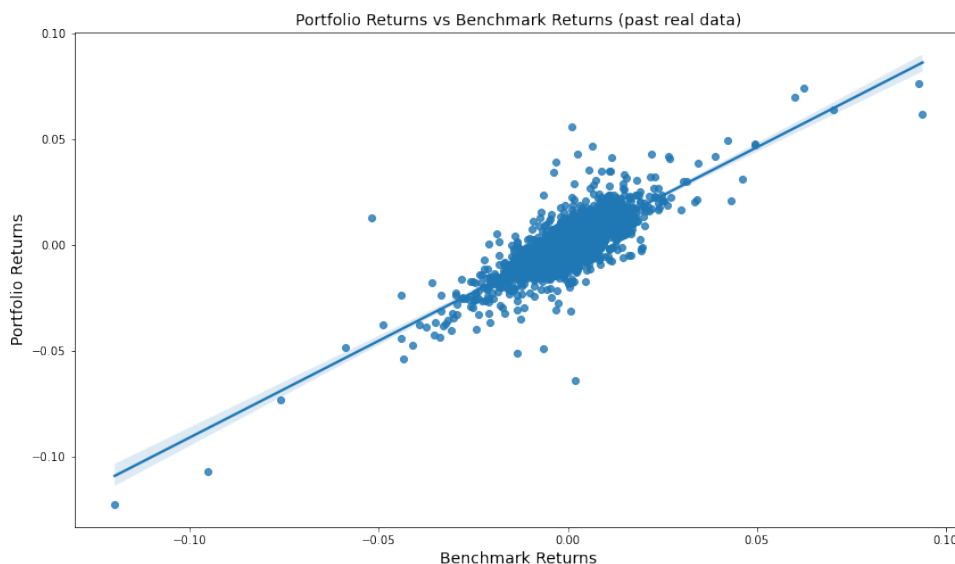


Figura 84: retta di regressione per calcolo del beta usando dati passati.

L'indice beta calcolato è pari a 0.9145.

6.2.2 beta del portfolio costituito con dati di previsione

In questo caso come per le sezioni precedenti per i 108 mesi sono stati usati sia i dati passati che quelli di previsione, Rispetto al calcolo di beta effettuato qui sopra è stata considerata una frequenza mensile invece di giornaliera in quanto il forecasting ha prodotto dati mensili (figura 85).

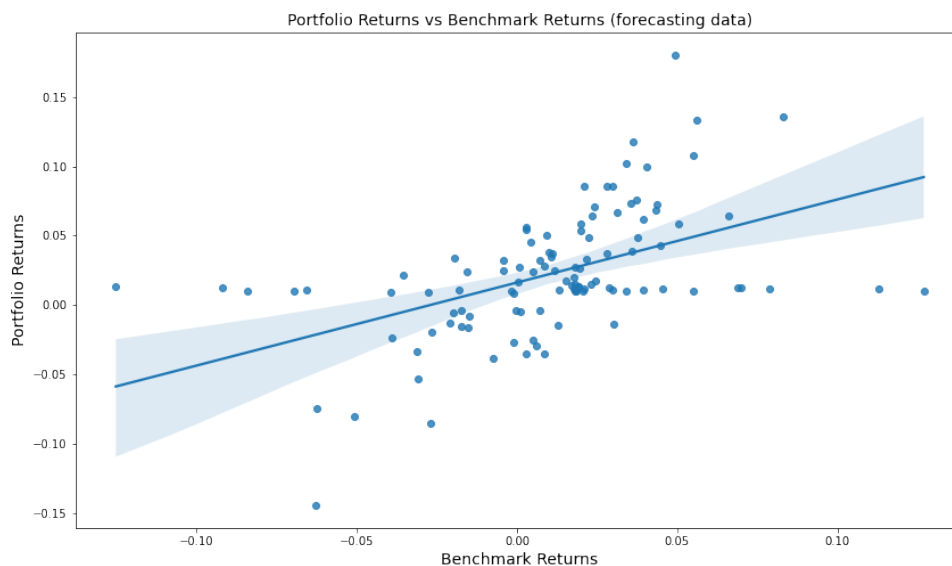


Figura 85: retta di regressione per calcolo del beta usando dati di previsione.

L'indice beta calcolato è pari a 0.5996.

6.3 Costruzione del portfolio effettivo e confronto

Per la costruzione del portfolio effettivo si può usare la tecnica di allocazione chiamata **1/n portfolio**, questa tecnica come da specifica del progetto alloca a tutti gli asset considerati peso uguale.

Nonostante il fatto che sembra una tecnica molto semplicistica è stato dimostrato che può essere complicato superare la performance di un **1/n portfolio** usando tecniche di asset allocation più complicate [DGU07].

Per effettuare tutti i benchmark e ottenere la valutazione del portfolio è stata utilizzata la libreria *pyfolio*.

6.3.1 Dati reali passati

Usando solo i dati reali passati si ottiene un tasso di rendimento annuale pari a 22.2% con volatilità annuale del 20.09%, mentre con la costruzione del portfolio usando Monte carlo abbiamo ottenuto un rendimento del 24.54% ottenendo tuttavia una volatilità inferiore pari a 19.48%.

6.3.2 Dati di previsione

In questo caso come nelle sezioni precedenti sono stati usati dati passati e dati di previsione in congiunzione, considerando una limitazione di *pyfolio* nei grafici di rendimento cumulativo è presente un effetto "gradino" per la mancanza di informazioni tra i mesi.

Il tasso di rendimento annuale identificato è pari al 20.1% mentre la volatilità 13.3%, con il metodo monte carlo si è identificato un portfolio con rendimento inferiore pari a 15.02% ma con volatilità inferiore pari a 8.56%.

7 Conclusioni

In questo progetto di Business Intelligence sono stati visti visti numerosi argomenti utili per ottenere informazioni e prendere decisioni critiche, a partire da numerose statistiche descrittive che hanno dimostrato importanti proprietà sui dati considerati come la elevata volatilità, predizione sul prezzo a partire dai dati utilizzando varie metodologie tra cui modelli statistici e deep learning, costruzione di un algoritmo per trading automatico e backtesting sulla strategia, stima sul rischio dei titoli considerati e in fine alla costruzione di un portfolio ottimizzato utilizzando anche le predizioni effettuate precedentemente.

Relativamente alle analisi di previsione è stato reso noto come la alta volatilità dei titoli abbia influito sulla previsione con modello **ARIMA**, il titolo **GOOG** è stato l'esempio più noto dove nel periodo di verifica il prezzo reale si è discostato oltre l'intervallo di confidenza.

Considerando la elevata volatilità dei titoli l'utilizzo di un algoritmo di deep learning avrebbe potuto ottenere dati di previsione più precisi, tuttavia dei test effettuati sul modello **MLP** hanno portato un livello di errore superiore a quello di naïve, segnalando la imprecisione del modello.

Nelle strategie di trading è stato molto interessante osservare come indicatori tecnici, come le Bande di Bollinger possano essere impiegate in un algoritmo per segnalare automaticamente quando effettuare l'acquisto o la vendita di un titolo, nel caso presentato nel capitolo 4 per esempio, durante backtesting si è riuscito nel periodo del 2020 ad avere un profitto, tuttavia anche se non presentato nel capitolo ci sono stati periodi come nel 2021 dove alla fine dell'anno il backtesting ha segnalato una perdita sul capitale investito inizialmente, ciò ha dimostrato che nonostante in alcuni casi l'algoritmo presentato funzioni, può comunque portare a una perdita del capitale investito, e per questo motivo è necessario il backtesting per cercare di creare algoritmi che mediante diverse strategie limitino il rischio di perdita del capitale.

Sui titoli considerati in questo progetto, nei capitoli 6 e 7 ne è stato calcolato il rischio utilizzando il **CAPM** per valutare il rischio e successivamente sono stati costruiti due portafogli "ottimizzati" utilizzando le simulazioni di monte carlo, il primo utilizzava unicamente i dati storici mentre il secondo usava sia i dati storici che quelli ottenuti tramite forecasting nel punto 3, la costruzione dei due portafogli è stata utile per evidenziare la dispersione notevolmente inferiore quando sono stati utilizzati i dati di autoregressione provenienti da **ARIMA**.

Per confrontare i portafogli ottimizzati sono stati costituiti successivamente altri due portafogli dove ogni titolo aveva lo stesso peso, sfruttando il concetto del **1/n portfolio**, in questo caso con solo i dati passati è stato evidenziato un rendimento inferiore con una volatilità più alta, mentre con i dati di previsione il portfolio **1/n** è riuscito ad avere una volatilità inferiore a discapito tuttavia del rendimento nettamente più basso.

Riferimenti bibliografici

- [Arr14] Argimiro Arratia. *Computational finance: An introductory course with R*. Atlantis Press, 2014.
- [DGU07] Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *The Review of Financial Studies*, 22(5):1915–1953, 12 2007.
- [Lew20] Eryk Lewinson. *Python for finance cookbook: Over 50 recipes for applying modern python libraries to financial data analysis*. Packt Publishing Ltd., 2020.
- [MDV21] Mieszko Mazur, Man Dang, and Miguel Vega. Covid-19 and the march 2020 stock market crash. evidence from s&p1500. *Finance Research Letters*, 38:101690, 2021.