

Progetto Business Intelligence per i Servizi Finanziari

Daniel Scalena, 844608

December/January, 2020/2021

Contents

1 Stocks Selection	3
1.1 Settore tecnologico	3
1.1.1 Apple (AAPL)	3
1.1.2 Nvidia (NVDA)	3
1.2 Settore alimentare	4
1.2.1 The Coca-Cola Company (KO)	4
1.2.2 Unilever UL	4
1.3 Settore bancario-finanziario	4
1.3.1 Bank of America Title (BAC)	4
1.3.2 American Express (AXP)	4
2 Descriptive analytics	5
2.1 Returns analysis and distribution	5
2.1.1 Settore tecnologico	5
2.1.2 Settore alimentare	8
2.1.3 Settore bancario-finanziario	11
2.2 Descriptive statistics	15
2.2.1 Settore tecnologico	16
2.2.2 Settore alimentare	18
2.2.3 Settore bancario-finanziario	19
2.3 Beta	21
2.4 Focus 2018-2020	25
3 Predictive analytics	35
3.1 SVM	35
3.2 ARIMA	39
4 Portfolio Management and Optimization	45
4.1 Portfolio investment using Markowitz model	45
4.2 Portfolio investment using CAPM	46
4.3 Portfolio investment using forecasted values	46
5 Web Applet	48
5.1 Descriptive Analytics	48
5.2 Portfolio Management and Optimization	48

Abstract

Il progetto prevede un'analisi esplorativa, riguardate un'ispezione visiva dei grafici, un'analisi dei valori statistici relativi ad essi, predittiva su delle azioni e contenente una parte di portfolio management. Viene anche utilizzata una webapplet per illustrare i principali risultati delle analisi e della parte di portfolio management

1 Stocks Selection

Le azioni scelte per questo progetto sono rappresentative di 3 grandi settori industriali e dei servizi.

Nello specifico vengono prese le azioni di Apple e Nvidia per quanto riguarda il settore tech. Questa scelta è dovuta al fatto che entrambe le aziende puntano molto sulla crescita e sul futuro, Apple nel campo dell'elettronica di consumo, Nvidia nell'intelligenza artificiale. Quest'ultima è infatti una dei pochi leader nel settore per ricerca e tra le principali aziende produttrici di hardware per l'addestramento di deep neural networks e lavori computazionalmente complessi.

Vengono poi considerate le azioni di Coca Cola e Unilever, aziende molto solide che operano nel settore alimentare. Queste due compagnie mantengono costante nel tempo la loro presenza sul mercato cercando di innovare ma senza assumeri rischi eccessivi.

Infine vengono considerate le azioni di Bank of America e American Express per il settore economico/finanziario. La prima è sicuramente una tra le banche statunitensi più note, mentre la seconda rappresenta uno tra i circuiti di credito più utilizzati per le transazioni economiche al momento.

Tutte queste azioni vengono scambiate nel New York Stock Exchange o nel mercato elettronico Nasdaq.

1.1 Settore tecnologico

Negli ultimi anni il settore tecnologico ha sicuramente catturato l'attenzione dei grandi investitori soprattutto per il suo generale andamento positivo nel mercato. Il successo di grandi aziende quali Google, Apple, Amazon e Microsoft, solo per fare qualche esempio, dimostra come, grazie alla tecnologia, si riescano a fondare multinazionali con solidissime basi capaci di durare a lungo nei decenni.

Le azioni scelte per questo settore sono:

1.1.1 Apple (AAPL)

Apple è sicuramente una delle aziende più iconiche a livello internazionale nel settore tecnologico. Si occupa principalmente della produzione di smartphone, tablet e computer, dello sviluppo di sistemi operativi e software dedicati, oltre che tanti altri dispositivi tecnologici di uso quotidiano. È attualmente l'azienda con capitalizzazione azionaria più alta al mondo e vanta numeri per fatturato e utile netto decisamente alti. Nel 2019 il fatturato si attesta a circa 260 miliardi di dollari con un utile netto di 55 miliardi di dollari

1.1.2 Nvidia (NVDA)

Nvidia è un'azienda produttrice di hardware per computer, più nello specifico di schede grafiche. È una nota azienda capace di innovarsi di volta in volta investendo nel futuro la maggior parte dei propri ricavi. Grazie a questi investimenti è attualmente una tra le aziende leader nel settore dell'intelligenza, soprattutto per quanto riguarda il deep learning, e nel settore grafico sia dal punto di vista della produzione professionale sia in quello intrattenitivo e videoludico. Recentemente è inoltre stata ufficializzata l'acquisizione di ARM per 40 miliardi, sempre in vista di un futuro successo di questa architettura. Nel 2019 l'azienda ha fatturato circa 11 miliardi e mezzo di dollari.

1.2 Settore alimentare

Caratteristica importante di questo settore è sicuramente la sua solidità e la sua efficacia nel mantere stabile nel tempo gli introiti senza alzare eccessivamente il rischio

1.2.1 The Coca-Cola Company (KO)

Coca Cola è probabilmente tra i brand più conosciuti di sempre in tutto il mondo. L'azienda ha da sempre mantenuto una fortissima presenza nel settore alimentare dimostrandosi leader nella produzione di bibite analcoliche. Oltre alla Coca Cola infatti l'azienda produce e vende anche Fanta, Sprite e altre bevande simili. Nel 2019 il suo fatturato è stato di circa 37 miliardi di dollari con un utile netto di circa 9 miliardi di dollari.

1.2.2 Unilever UL

Unilever è tra le più grandi aziende nel settore alimentare e non solo. È proprietaria di circa 400 brands e intrattiene accordi su produzione o distribuzione con i più importanti produttori di generi alimentari e per la cura personale. Tra questi 400 brands spiccano sicuramente Lipton, Algida, Calvè, Knoee, Findus e tante altre. Tra le collaborazioni invece notiamo quelle con Ferrero e Nestlè e altri ancora. Nel 2019 l'azienda ha fatturato quasi 52 miliardi di euro.

1.3 Settore bancario-finanziario

Il settore bancario-finanziario garantisce una buona solidità e rispecchia molto l'andamento e il benessere della società. I principali guadagni in questo settore provengono da investimenti, prestiti e mutui.

1.3.1 Bank of America Title (BAC)

Bank of America è la seconda banca più grande degli Stati Uniti dopo JP Morgan C. La sua alta capitalizzazione assicura una forte presenza nel settore mantenendo comunque un basso rischio. Il fatturato del 2019 di Bank of America si attesta a circa 23 miliardi di dollari

1.3.2 American Express (AXP)

American Express è una società principalmente operante nel settore dei servizi finanziari e dei viaggi. Insieme a Visa e Mastercard gestisce uno tra i principali circuiti di credito ma a differenza delle prime due emette direttamente le carte di credito. L'azienda è generalmente associata a una clientela di élite. Si occupa inoltre di investimenti, gestioni patrimoniali e assicurazioni. Nel 2019 il suo fatturato è stato di circa 43 miliardi di dollari, realizzando un utile netto di quasi 7 miliardi di dollari.

2 Descriptive analytics

In questo capitolo è contenuta un'analisi descrittiva prendendo in considerazione i dati delle azioni viste nel capitolo precedente. L'analisi sarà inoltre dimostrata visivamente attraverso l'uso di grafici o di dati grezzi presi grazie al servizio offerto da Yahoo Finance.

Il periodo delle analisi è stato ampliato per riuscire meglio ad analizzare il comportamento del mercato in un investimento a lungo termine che potrebbe comprendere quindi anche crisi di varia natura.

2.1 Returns analysis and distribution

I ritorni delle azioni sono espressi mensilmente e nei grafici è visualizzabile un periodo temporale che parte dal 2007 e raggiunge l'inizio del 2021. Per comodità di analisi e visualizzazione le azioni sono raggruppate per settore in modo da poter notare similitudini ed eventi che colpiscono maggiormente un settore rispetto ad un'altro.

2.1.1 Settore tecnologico

La prima analisi generale viene effettuata osservando i movimenti dei prezzi di chiusura nel tempo di entrambe le azioni. Nel caso di Apple nel grafico 1 notiamo un trend decisamente in crescita nel corso degli anni molto simile a quello di Nvidia rappresentato nel grafico 2. Nvidia, a differenza di Apple, ha una crescita decisamente più repentina dal 2016. Questa crescita è probabilmente dovuta a degli investimenti fatti nel 2014 che iniziavano a dare i propri frutti in settori di grafica, processori mobile e intelligenza artificiale. A fine 2018 però il prezzo scende nuovamente a causa del deprezzamento delle criptovalute e della conseguente sovrapproduzione di schede video da parte dell'azienda. Tra le altre similitudini evidenti troviamo una leggera flessione dei prezzi all'inizio del 2020 dovuta sicuramente alla crisi mondiale scatenata dalla pandemia causata dal virus Covid-19 analizzata nel dettaglio nel sottocapitolo dedicato al periodo 2018-2020. Nonostante la gravità di questa crisi il settore tech è sicuramente quello meno colpito grazie al forte aumento nell'utilizzo di dispositivi elettronici durante i periodi di quarantena. Come detto, vedremo più avanti un approfondimento su questo periodo e la reazione degli altri settori.

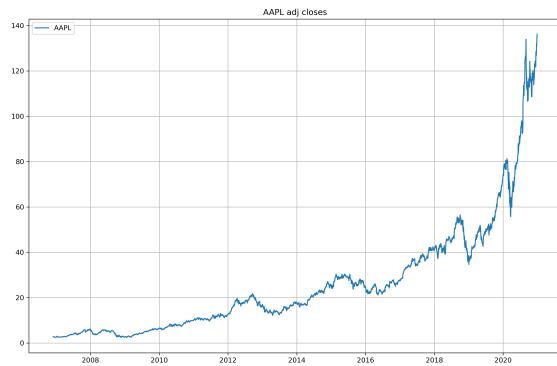


Figure 1: AAPL Adj Closes

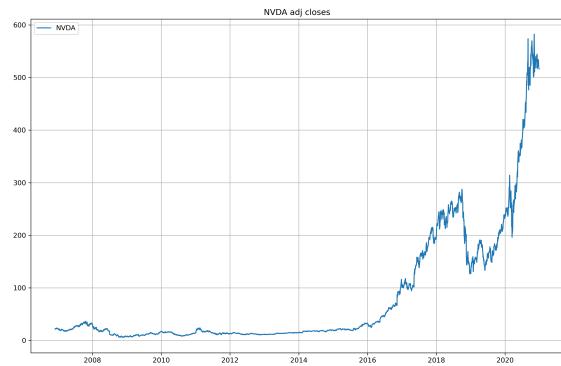


Figure 2: NVDA Adj Closes

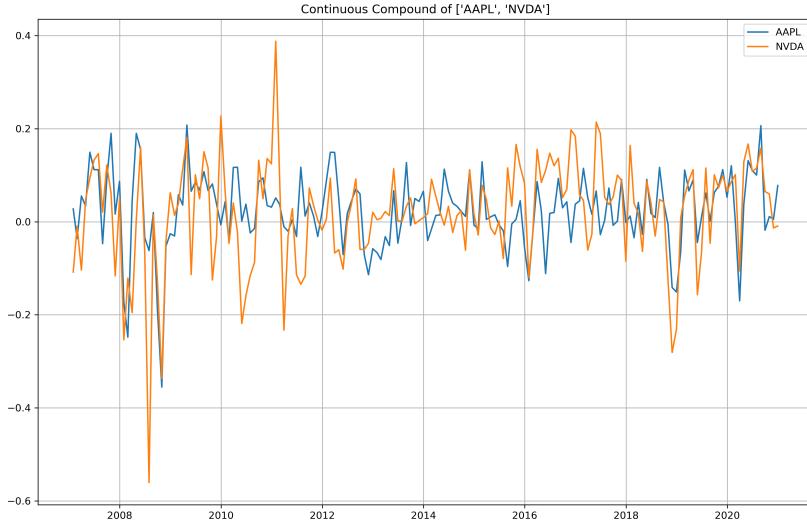


Figure 3: Continuous Compound of AAPL and NVDA

Per prima cosa è utile controllare il grafico 3 relativo ai ritorni di Apple e Nvidia. Entrambe le azioni, appartenendo allo stesso settore, sono correlate tra loro.

Grazie all'analisi sulla distribuzione dei ritorni mostrata nelle figure 4 e 5 è possibile notare come in entrambi i casi i ritorni sono generalmente positivi. Questo è osservabile dal picco della curva di densità della distribuzione leggermente spostata a destra dello 0. Questa è una caratteristica comune a tutte le azioni appartenente al settore tecnologico: grazie ai progressi negli ultimi anni le aziende tech hanno avuto una crescita enorme se comparata con azioni di altri settori longevi.

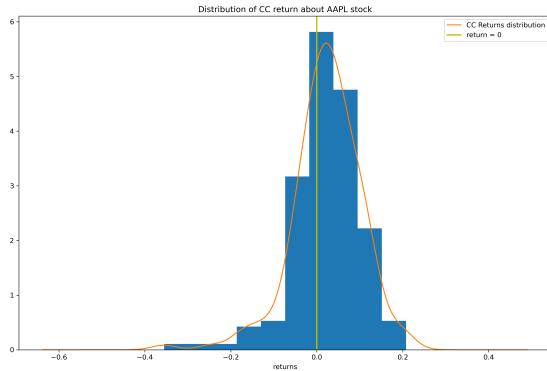


Figure 4: Distribution of CC return about AAPL

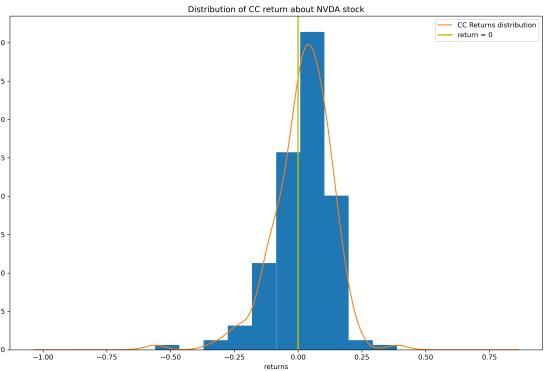


Figure 5: Distribution of CC return about NVDA

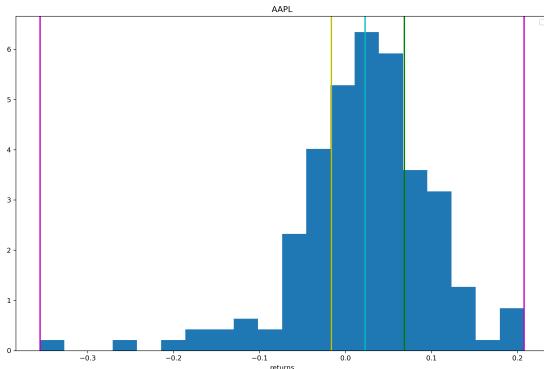


Figure 6: AAPL quantiles

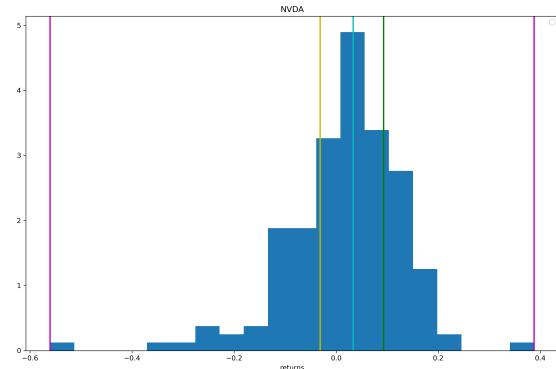


Figure 7: NVDA quantiles

È tuttavia evidente come il titolo di Nvidia sia decisamente più volatile rispetto a quello Apple. Quest'ultima affermazione è ancora più visibile osservando la distribuzione dei ritorni dei due titoli (figura 6 e 7) focalizzandosi sui valori dei quantili e sullo loro distanza. Nel caso di Apple la distribuzione dei ritorni è decisamente concentrata verso lo 0 se comparata con la distribuzione dei ritorni nel titolo di Nvidia.

Questo aspetto è osservabile anche utilizzando un grafico di tipo qqplot che permette di confrontare la distribuzione dei quantili dell'azione presa in considerazione con quella di una normale standard (rispettivamente pallini blu e linea rossa nel grafico 8 e 9). Come è possibile notare, soprattutto per quanto riguarda gli estremi del grafico, il titolo di Nvidia assume valori decisamente più lontani dallo 0 rispetto al titolo di Apple e in generale più lontani dalla retta della normale.

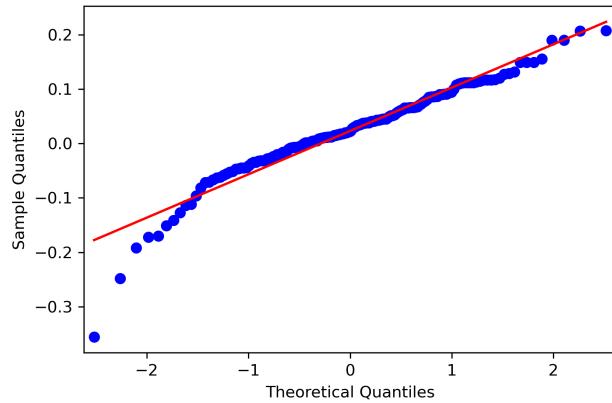


Figure 8: AAPL qqplot

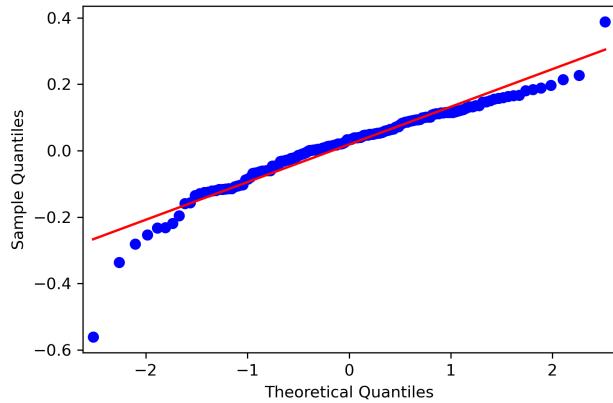


Figure 9: NVDA qqplot

Quanto affermato precedentemente riguardo la superiore volatilità del titolo di Nvidia trova ulteriore conferma nel boxplot 10: come si può notare il box di Nvidia risulta essere molto più ampio se comparato con quello di Apple così come lo sono anche gli outliers rappresentati dai palliini in figura.

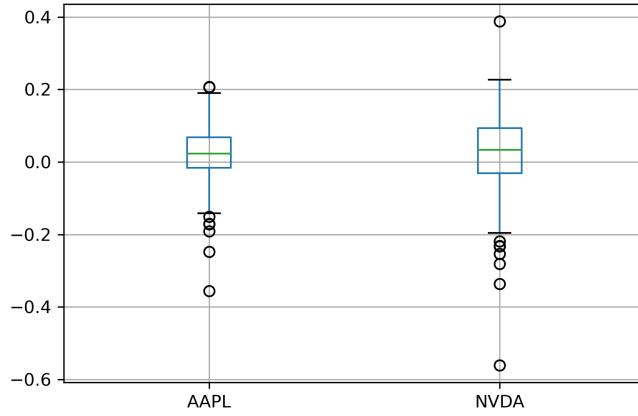


Figure 10: Boxplot APPL-NVDA

2.1.2 Settore alimentare

Il settore alimentare è sicuramente caratterizzato da una crescita lenta ma costante nel tempo come mostrato dai prezzi di chiusura mensili nel corso degli anni nei grafici 11 e 12. Ovviamente, vista la crescita lenta e costante, il rischio in questo settore è basso se comparato al rischio assunto sia nel settore tech analizzato precedentemente sia nel settore finanziario che analizzeremo successivamente. Oltre a un'evidente correlazione dell'andamento che verrà poi analizzata in seguito nel capitolo relativo ai dati statistici, possiamo notare come la crisi finanziaria, e successivamente economica, del 2008 influisca solamente per un piccolo periodo sulla crescita di questi due titoli non essendo un settore particolarmente esposto ai problemi di tipo finanziario. Molto simili, anche se leggermente peggiori, sono le conseguenze della pandemia del 2020. In questo cosa notiamo un calo netto del valore delle azioni che però continua progressivamente a riprendersi raggiungendo, come nel caso di Unilever, il suo valore pre-pandemia.

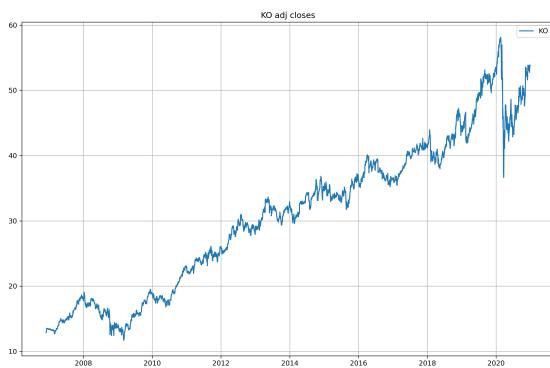


Figure 11: KO Adj Closes

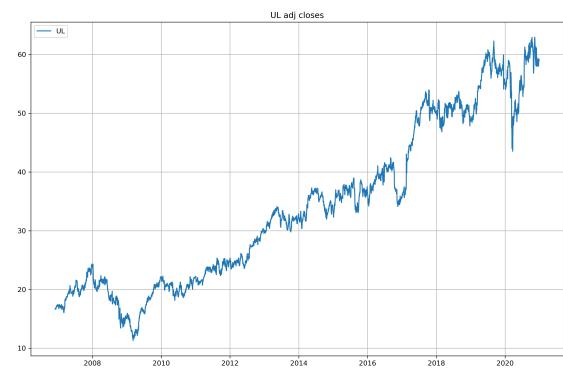


Figure 12: UL Adj Closes

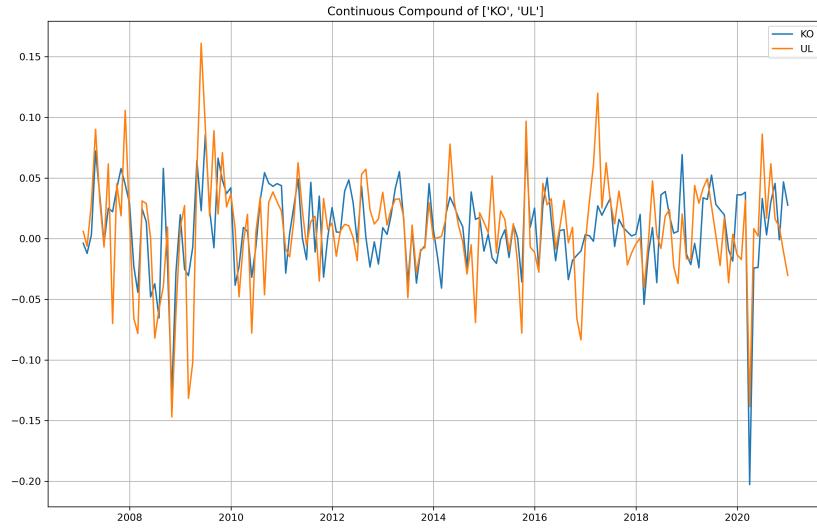


Figure 13: Continuous Compound of KO and UL

Analizzando il grafico 13 dei ritorni è sicuramente notevole come entrambi i titoli condividano lo stesso andamento per la maggior parte del tempo e, come visto precedentemente, i più forti cali nei ritorni si hanno a causa della crisi del 2008 e della pandemia nel 2020.

Entrambi i titoli non hanno mai ritorni eccessivamente alti o bassi e generalmente, come è possibile osservare dal picco della distribuzione illustrata nei grafici 14 e 15, si parla di ritorni lievemente positivi.

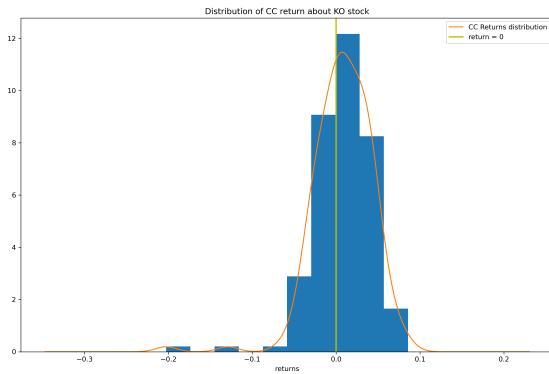


Figure 14: Distribution of CC return about KO

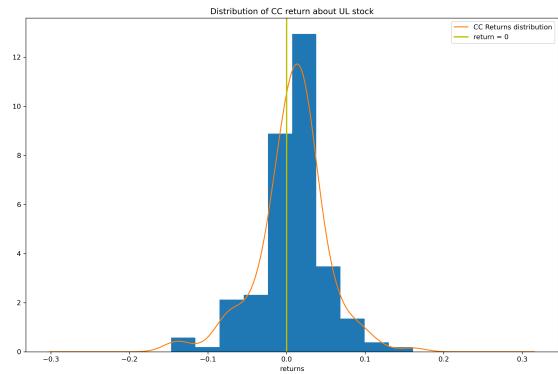


Figure 15: Distribution of CC return about UL

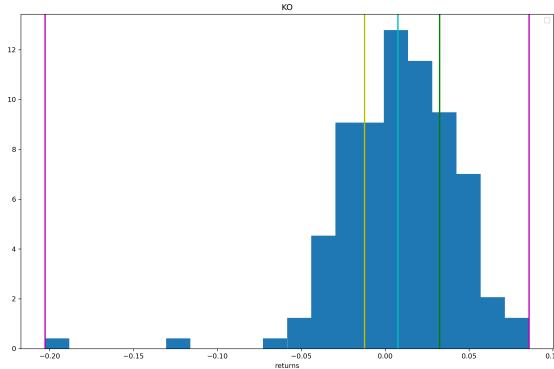


Figure 16: KO quantiles

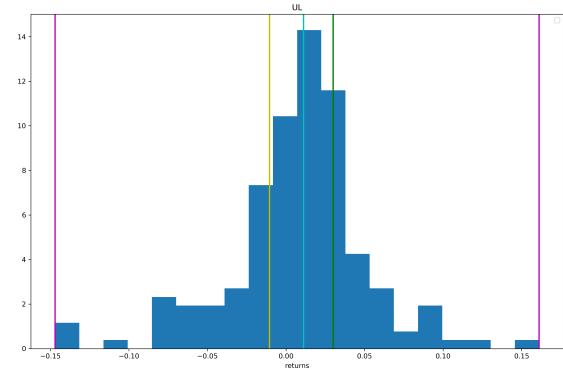


Figure 17: UL quantiles

Se si passa all’analisi dei quantili è possibile osservare come Coca Cola sia tendenzialmente più concentrata verso lo 0 e la sua distribuzione è più simile a una normale. È possibile osservare quanto detto nei grafici qqplot 18 e 19 in cui l’azione di Coca Cola ha quantili distribuiti più uniformemente sulla retta rossa rappresentante i quantili di una normale.

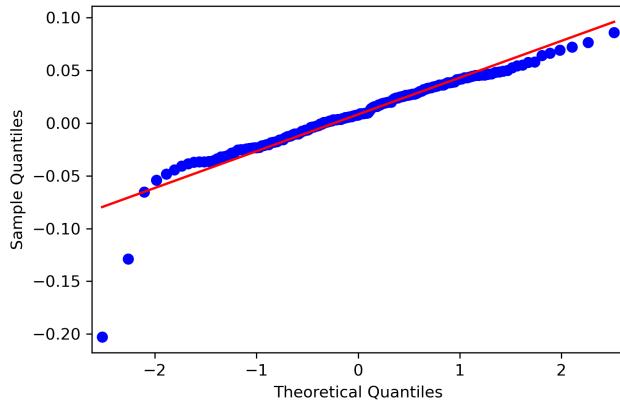


Figure 18: KO qqplot

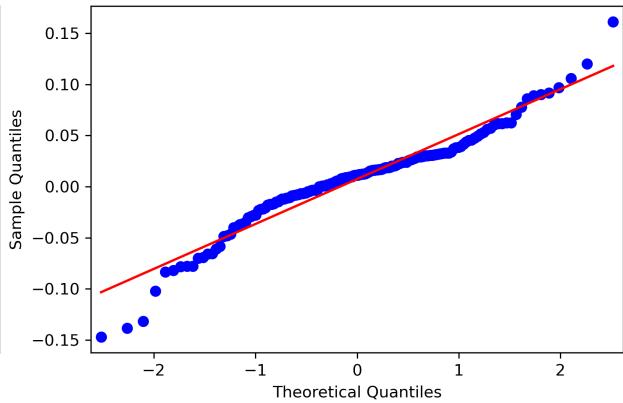


Figure 19: UL qqplot

Quanto affermato precedentemente trova ovviamente riscontro nei boxplot in figura 20. Unilever presenta sicuramente molti più outliers rispetto a Coca Cola determinando quindi un’attività sicuramente maggiore.

È facile quindi concludere che Unilever sia leggermente più volatile rispetto a Coca Cola, osservazione che sarà poi confermata dai dati statistici visti nel capitolo successivo.

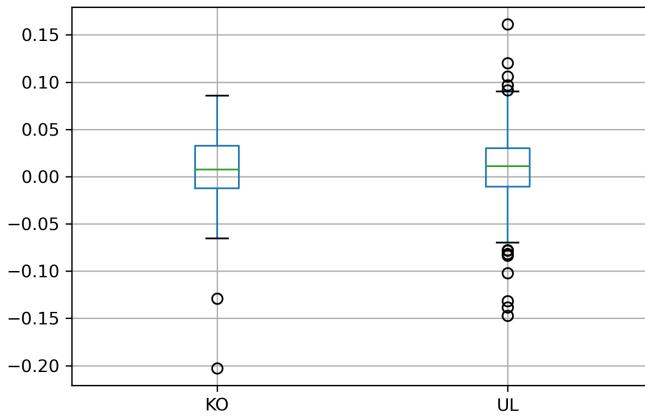


Figure 20: Boxplot KO-UL

È in ogni caso importante precisare che i valori assunti come volatili nell'analisi appena fatta riguardo il settore alimentare sono sicuramente più piccoli rispetto alla volatilità vista, ad esempio, per il settore tecnologico.

Questo settore non viene scelto per la sua capacità di offrire grandi ritorni quanto più per i suoi dividendi regolari e sicuri che permettono una crescita finanziaria quasi sicuramente positiva.

2.1.3 Settore bancario-finanziario

Questo settore è sicuramente il più colpito dalla crisi finanziaria ed economica del 2008 e degli anni successivi. Soprattutto per quanto riguarda Bank of America è notevole un grave calo del prezzo dal 2008 in poi nonostante la sua resilienza tanto da iniziare le trattative, anche se successivamente fallite, per il salvataggio e l'acquisizione di Lehman Brothers, un'importante banca di investimento duramente colpita dalla crisi. Osservando il grafico si nota un leggero rialzo verso il 2009 probabilmente dovuto a un'iniezione di liquidità da parte del governo americano per un programma di salvataggio delle banche e quindi alla conseguente crescita della fiducia, se pur sempre moderta, degli investitori nel sistema bancario. Allo stesso modo America Express risente della crisi mondiale seppur in minore intensità rispetto a Bank of America. Entrambe le azioni e più in generale l'intero settore negli anni successivi è caratterizzato comunque da una crescita abbastanza costante con però delle difficoltà per entrambi all'inizio del 2020 sempre a causa della pandemia da Corona Virus.

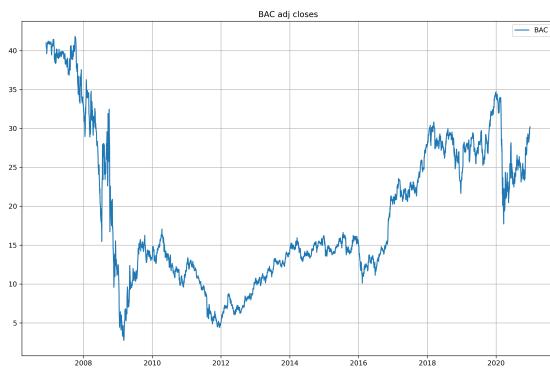


Figure 21: BAC Adj Closes

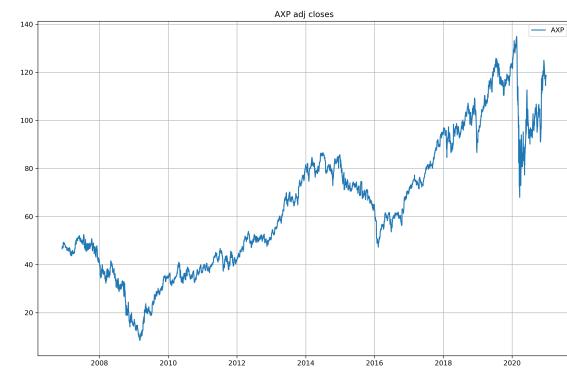


Figure 22: AXP Adj Closes

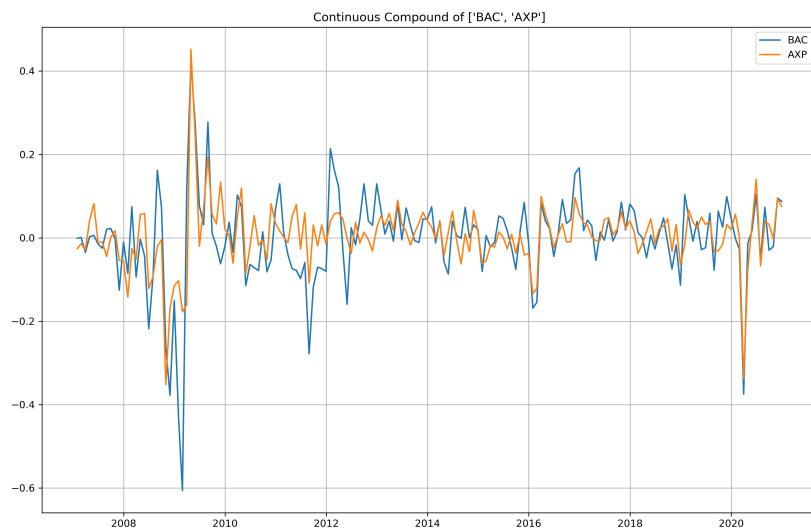


Figure 23: Continuous Compound of BAC and AXP

Guardando al grafico dei CC returns 23 è possibile notare come generalmente le due azioni siano correlate tra loro e assumono prevalentemente valori vicino allo 0 (guardando la distribuzione dei ritorni nei grafici 24 e 25)

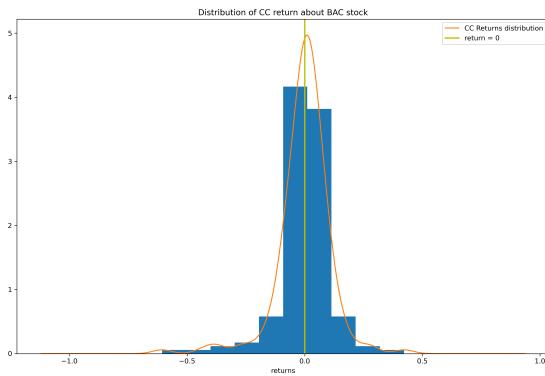


Figure 24: Distribution of CC return about BAC

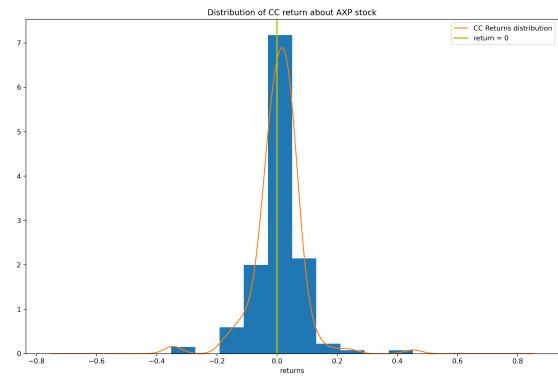


Figure 25: Distribution of CC return about AXP

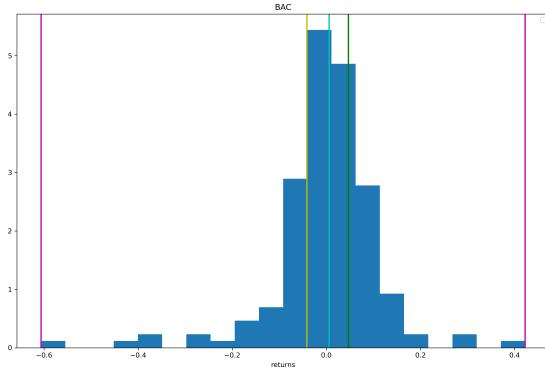


Figure 26: BAC quantiles

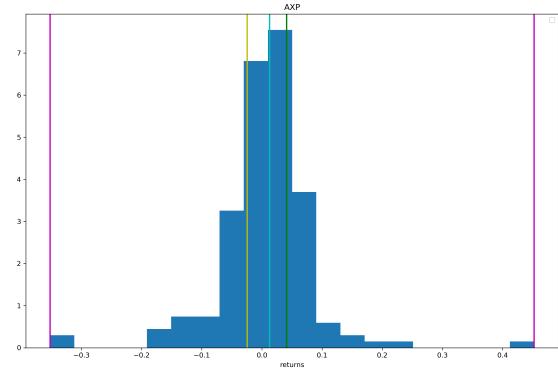


Figure 27: AXP quantiles

Grazie agli istogrammi che evidenziano i quantili (figura 26 e 27) e ai qqplot che li confrontano con i relativi quantili di una normale (figura 28 e 29) è possibile notare come le varie crisi che ci sono state influenzino molto sia la distribuzione dei ritorni che la presenza di outliers. Questione evidenziabile anche nel boxplot 30.

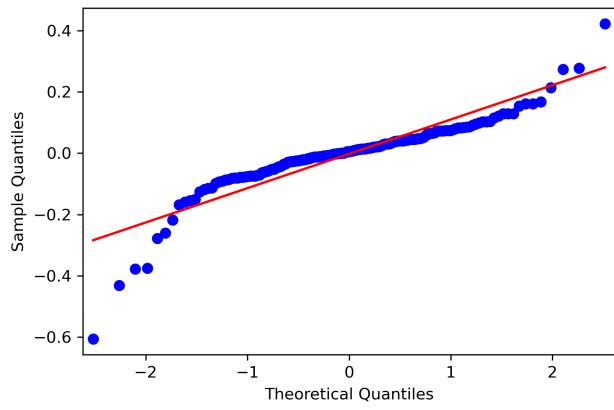


Figure 28: BAC qqplot

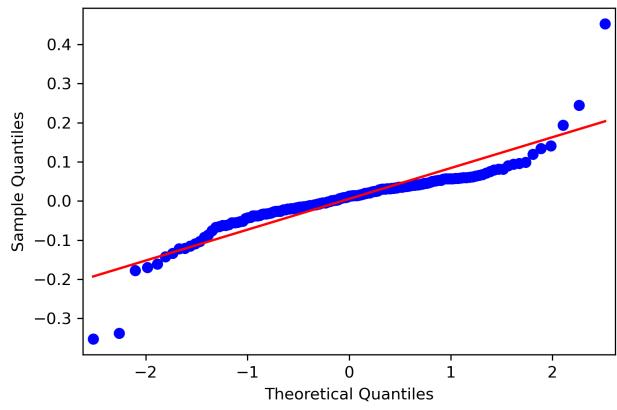


Figure 29: AXP qqplot

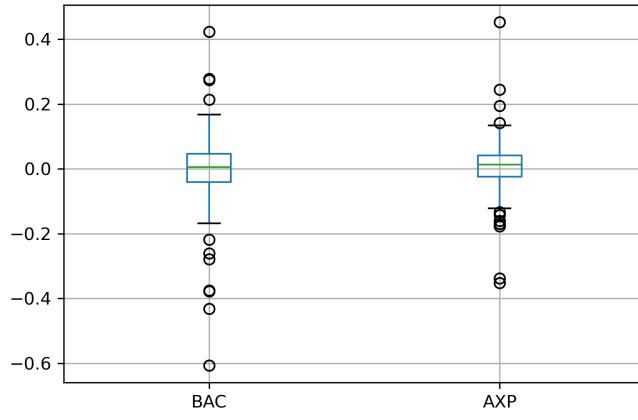


Figure 30: Boxplot BAC-AXP

Per valutare in maniera più accurata Bank of America quando non si trova in gravi periodi di crisi è possibile **escludere i dati relativi alla crisi del 2008 e degli anni successivi** facendo partire il calcolo dei CC returns dal 2010 in poi. In questo caso si ottiene un qqplot come in figura 31. Da questo grafico è possibile notare come i valori anomali siano decisamente di meno rispetto al qqplot di figura 28 riducendo quindi la varianza e dando un'idea più precisa del comportamento del titolo in condizioni di "tranquillità" dei mercati.

Anche dopo quest'analisi, escludendo gli anni dal 2008 al 2010, è comunque possibile notare come American Express riesca a mantenere una volatilità minore rispetto al titolo bancario rendendolo quindi meno rishioso.

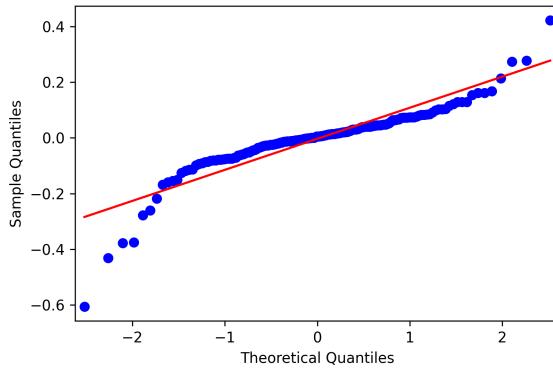


Figure 31: BAC qqplot starting from 2010

Nel passato questo settore più di altri ha accusato i danni peggiori delle crisi di ogni genere ma come dimostrato sembra comunque tenere una solidità nel tempo grazie anche all'importanza che ha nella vita economica di ogni persona

2.2 Descriptive statistics

In questo capitolo sono presentati dati descrittivi e numerici riguardo quanto affermato nel capitolo precedente. Sono inoltre stati riportati gli scatter plot e analisi di covarianza e correlazione incrociando tutte le diverse azioni scelte.

Anche in questa parte di statistica descrittiva i dati delle azioni vengono divisi per settore per comodità di visualizzazione e chiarezza del contenuto.

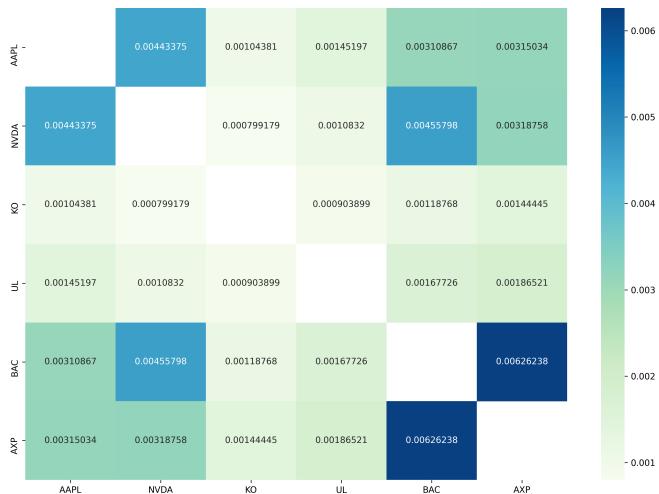


Figure 32: Covariance Matrix

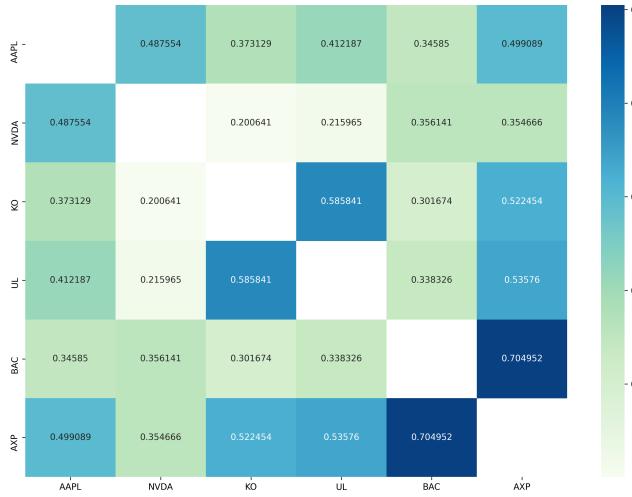


Figure 33: Correlation Matrix

2.2.1 Settore tecnologico

	Stat	Value
1	mean	0.02297
2	var	0.00638
3	std	0.07989
4	min	-0.35532
5	max	0.20776
6	skew	-0.94525
7	kurt	3.23605
8	0.25	-0.01650
9	0.50	0.02279
10	0.75	0.06840

Table 1: AAPL CC return stats

	Stat	Value
1	mean	0.01877
2	var	0.01295
3	std	0.11379
4	min	-0.56069
5	max	0.38797
6	skew	-1.07740
7	kurt	4.15838
8	0.25	-0.03126
9	0.50	0.03345
10	0.75	0.09303

Table 2: NVDA CC return stats

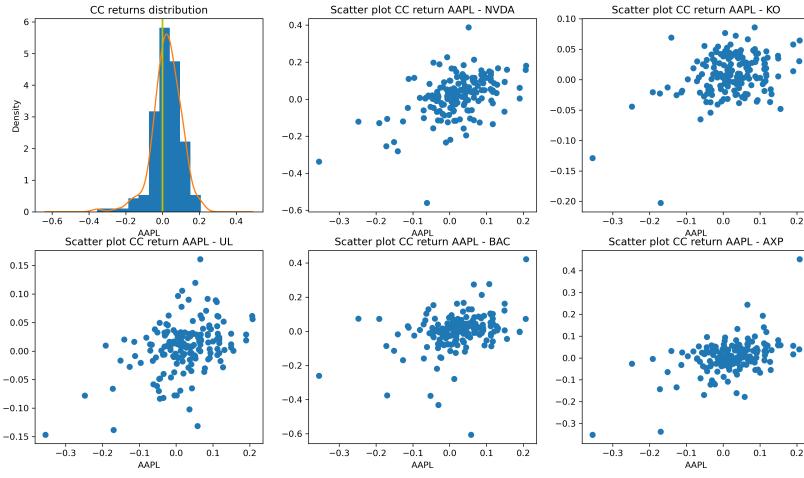


Figure 34: Scatter plot AAPL

Proprio come anticipato e visto nei diagnostic plots precedenti entrambi i titoli tecnologici hanno una media di ritorni superiore rispetto agli altri titoli presi in considerazione. Nvidia si conferma più volatile e quindi rischiosa rispetto a Apple guardando alla sua varianza decisamente più alta. Un valore importante inoltre è quello di skewness (asimmetria) della distribuzione di densità dei ritorni. Per Apple, e ancor di più per Nvidia, questa assume un valore negativo indicando come la coda della distribuzione (in arancione nel grafico 34 e 35) abbia una lunghezza maggiore rispetto a una curva di una distribuzione gaussiana. Il valore di kurtosis invece risulta essere positivo in entrambi i casi ma maggiormente nel titolo di Nvidia: questo indica che la curva è leptocurtica, ovvero più appuntita rispetto a una normale.

Analizzando invece i grafici di dispersione 34 e 35 notiamo come la correlazione sia sempre positiva tra tutti gli asset. Nello specifico Apple risulta avere una correlazione prevedibile con Nvidia, confermando quanto detto nell'analisi dei grafici diagnostici nel capitolo precedente. Oltre che con Nvidia è possibile notare una correlazione anche con America Express confermato dal valore visto nella figura 33.

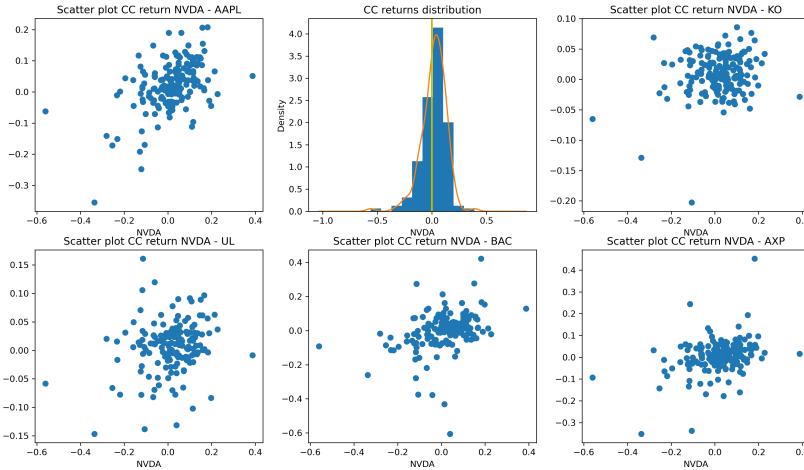


Figure 35: Scatter plot NVDA

2.2.2 Settore alimentare

	Stat	Value
1	mean	0.00820
2	var	0.00122
3	std	0.03500
4	min	-0.20274
5	max	0.08584
6	skew	-1.55945
7	kurt	7.85750
8	0.25	-0.01223
9	0.50	0.00763
10	0.75	0.03251

Table 3: KO CC return stats

	Stat	Value
1	mean	0.00737
2	var	0.00194
3	std	0.04409
4	min	-0.14698
5	max	0.16113
6	skew	-0.44239
7	kurt	2.27185
8	0.25	-0.01048
9	0.50	0.01120
10	0.75	0.03001

Table 4: UL CC return stats

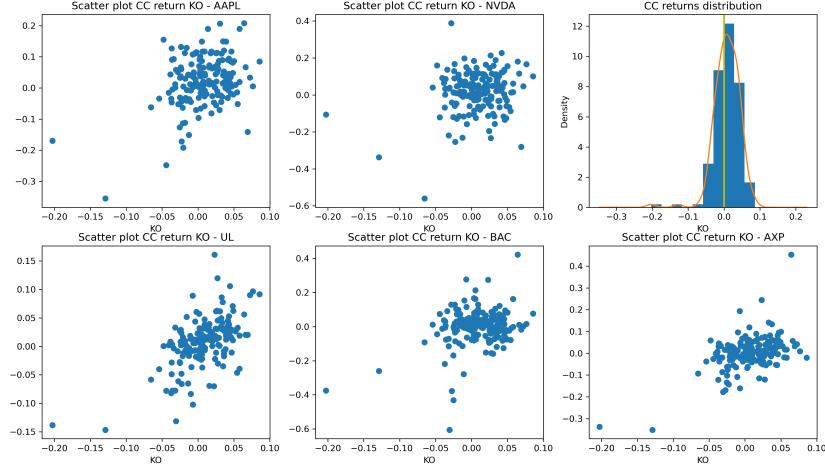


Figure 36: Scatter plot KO

Come anticipato per il settore alimentare la media dei ritorni risulta essere decisamente più bassa rispetto agli altri titoli ma comunque positiva. Questo però viene compensato sicuramente dalla varianza molto bassa che si attesta a un valore quasi 6 volte inferiore rispetto al titolo di Apple e 10 volte inferiore rispetto a quello Nvidia. Grazie a questi valori come anticipato nel primo capitolo le azioni di questo settore si confermano come le più solide e stabili tra tutte quelle selezionate.

Il valore di asimmetria della distribuzione dei ritorni invece risulta essere in entrambi i casi negativo, soprattutto nel caso di Coca Cola, indicando come la coda della distribuzione sia più lunga rispetto alla testa mostrando come i ritorni negativi siano più lontani dal centro o maggiori in quantità rispetto a quelli positivi. Anche nel caso del valore di curtosi Coca Cola spicca indicando come in entrambi i casi si ha una distribuzione sicuramente più slanciata e appuntita rispetto a una normale.

Osservando infine gli scatter plot 36 e 37 si conferma la correlazione tra le due azioni appartenenti allo stesso settore ma anche una correlazione con American Express mostrata anche dal valore riportato in figura 33.

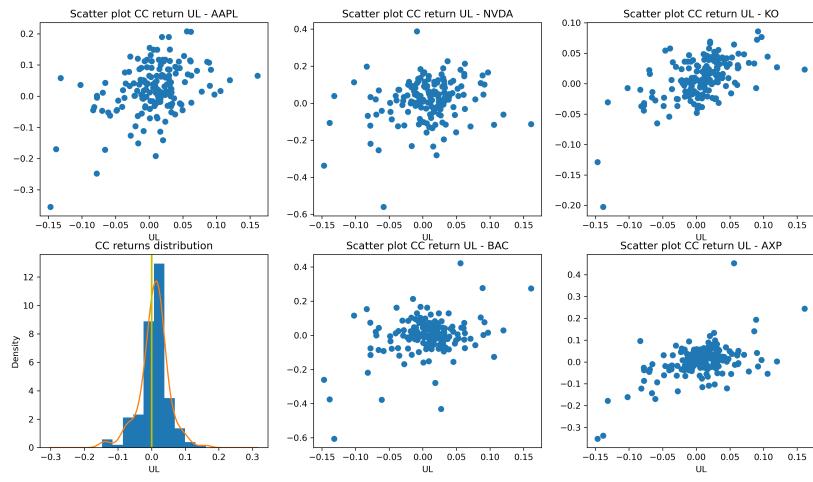


Figure 37: Scatter plot UL

2.2.3 Settore bancario-finanziario

	Stat	Value
1	mean	-0.00201
2	var	0.01265
3	std	0.11246
4	min	-0.60630
5	max	0.42253
6	skew	-1.32180
7	kurt	7.62129
8	0.25	-0.04116
9	0.50	0.00585
10	0.75	0.04693

Table 5: BAC CC return stats

	Stat	Value
1	mean	0.00543
2	var	0.00624
3	std	0.07899
4	min	-0.35215
5	max	0.45246
6	skew	0.01920
7	kurt	9.89422
8	0.25	-0.02455
9	0.50	0.01267
10	0.75	0.04097

Table 6: AXP CC return stats

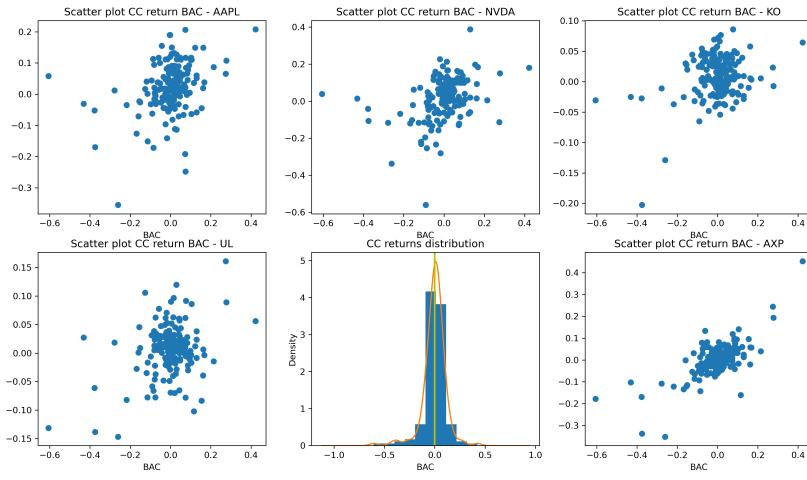


Figure 38: Scatter plot BAC

Anche il settore bancario e finanziario dimostra di avere valori bassi soprattutto per quanto riguarda Bank of America che ha una media dei ritorni di poco negativa. Questo, come anticipato, è sicuramente dovuto alla crisi degli anni successivi al 2008. Come visto precedentemente infatti se si considerano soltanto i valori di BAC relativi al periodo temporale 2010-2020 la sua media dei ritorni ritorna positiva attestandosi a un valore di 0.00810.

La varianza per BAC è sicuramente più alta rispetto a quella di American Express quasi come visto per i titoli di Nvidia e Apple rispettivamente (ottenendo però ritorni sicuramente più alti nel caso delle compagnie tecnologiche).

In American Express notiamo un valore di skewness leggermente maggiore di zero mostra come la testa della distribuzione dei ritorni sia leggermente più lunga della coda a differenza di Bank of America. Come facilmente intuibile invece dal grafico 24 e 25 riguardante le distribuzioni dei CC returns di BAC e AXP il valore di kurtosis è sicuramente fortemente positivo in entrambi i casi.

Questi due titoli presentano la correlazione più forte tra loro di tutti quelli analizzati precedentemente come è possibile osservare dagli scatterplot 38 e 39. American Express sembra inoltre essere correlata agli altri titoli (tranne per Nvidia dove il valore è leggermente più basso) probabilmente perchè il business centrale di questi settori è la vendita di beni materiali al pubblico.

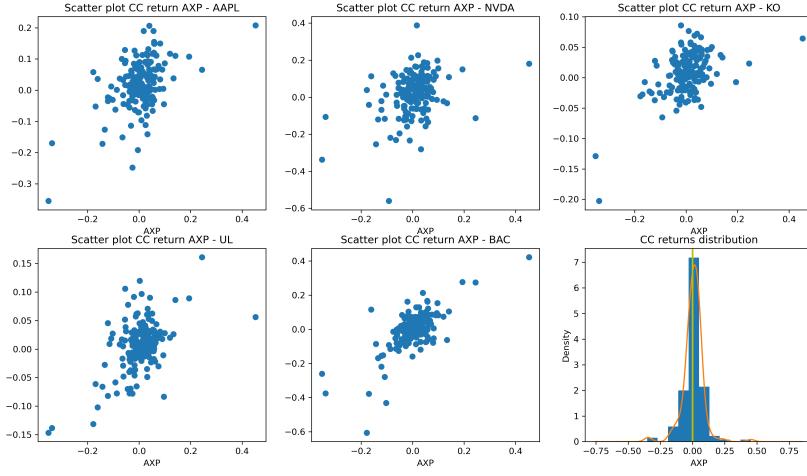


Figure 39: Scatter plot AXP

2.3 Beta

Il beta è un valore che misura come un titolo si muove strettamente insieme al mercato o un indice di riferimento. Questo valore è usato per misurare il rischio sistematico di un'azione che non può quindi essere eliminato con la diversificazione. È calcolato come segue:

$$\beta = \frac{Cov(R_i, R_m)}{Var(R_m)} \quad (1)$$

Può assumere un valore superiore a 1 quando un'azione amplifica il movimento del mercato o dell'indice di riferimento: questo ovviamente rende l'azione più rischiosa. Se il valore di β assume un valore compreso tra 0 e 1 l'azione tende a muoversi come il mercato mantenendo un rischio uguale o più basso. Infine, in caso di valore minore di 0, l'azienda si muove inversamente all'andamento del mercato.

Come indice di mercato è stato scelto l'indice Standard and Poor 500 che meglio rappresenta l'andamento della borsa americana dalla quale tutte queste azioni provengono. È composto da 500 società prese dal NYSE e Nasdaq che hanno un peso direttamente proporzionale al loro valore di mercato. Al suo interno trovano luogo aziende per la maggior parte tecnologiche (grazie al loro alto valore) come Apple, Microsoft, Amazon, Facebook, Alphabet (Google) ma anche Berkshire Hathaway, Johnson and Johnson, JPMorgan Chase and Co. e Visa.

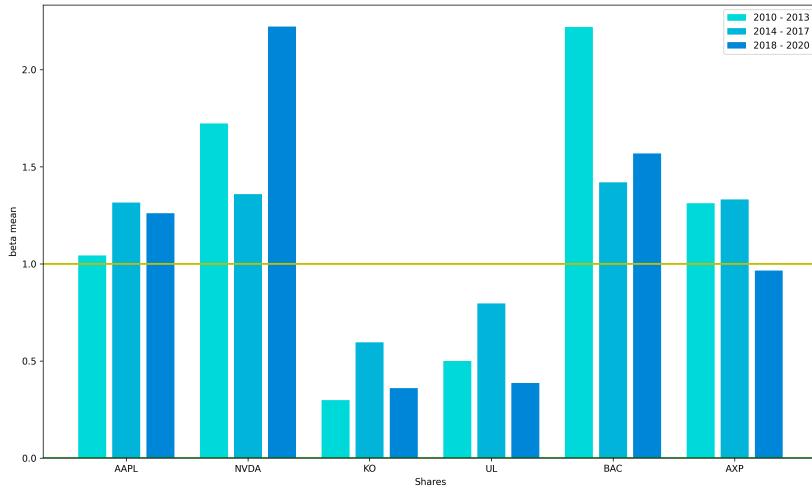


Figure 40: Beta from 2010 to 2020

In figura 40 è sono riportati i beta di tutte le azioni prese in considerazione divise in 3 periodi partendo dal 2010. A prima vista è possibile generalizzare guardando al mercato alimentare come il meno rischioso e al mercato tecnologico come il più esposto.

Questa generalizzazione trova corrispondenza anche nella realtà osservando il modello di imprenditoria delle principali aziende. Apple, e ancor di più Nvidia, adottano politiche di investimento rischiose investendo molto sul futuro e su progetti più a rischio: il loro obiettivo è conquistare nuove fette di mercato, accrescere la loro influenza e cercare di espandersi in più settori possibili. Al contrario compagnie come Coca Cola e Unilever mantengono un assetto più stabile puntando a rimanere solide nella posizione di mercato che già si sono ritagliate non facendo quindi nessun investimento particolarmente rischioso "tranquillizzando" di conseguenza il loro comportamento in borsa.

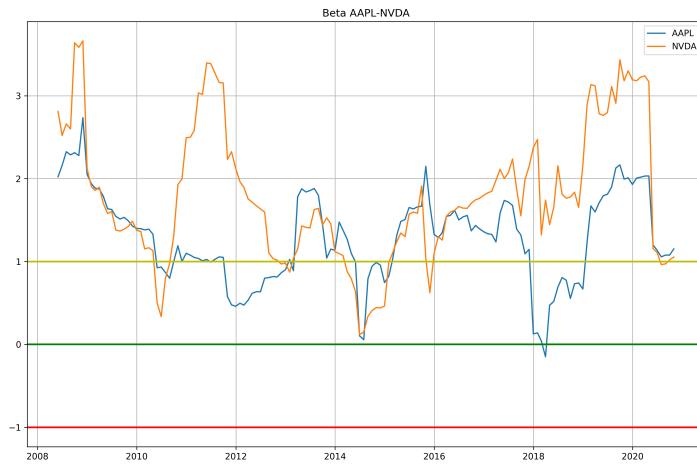


Figure 41: Beta APPL NVDA

Come anticipato entrambe le aziende tecnologiche prese in esame hanno una volere β che supera nella

maggior parte dei casi l'unità. Questo indica un comportamento più volatile del mercato di riferimento che viene evidenziato soprattutto nel caso di Nvidia. Quest'ultima azienda è stata protagonista di molti settori diversi negli ultimi anni, e continua nell'investire e collaborare attivamente in nuove tecnologie scommettendo sul loro successo (si prenda per esempio gli studi in materia di deep learning o la recente acquisizione di ARM).



Figure 42: Beta KO UL

Nel settore alimentare invece troviamo un valore di β che è sicuramente più rassicurante in termini di rischio. Soprattutto negli ultimi anni sia Coca Cola che Unilever seguono l'andamento del mercato senza mai alzare il rischio. Il loro valore di β infatti rimane nella maggior parte dei casi tra 0 e 1.



Figure 43: Beta BAC AXP

Il settore bancario-finanziario invece risente sicuramente della crisi originatasi nel 2008 ma riesce comunque a mantenersi stabile seguendo il mercato e non eccedendo di molto dal valore 1 come nel mercato tecnologico.

2.4 Focus 2018-2020

In questa sezione sono presenti le analisi relative solamente al periodo ottobre 2018 - ottobre 2020. Molte delle considerazioni restano simili a quanto visto precedentemente negli anni passati: per questa ragione ci si soffermerà solo sulle differenze sostanziali.

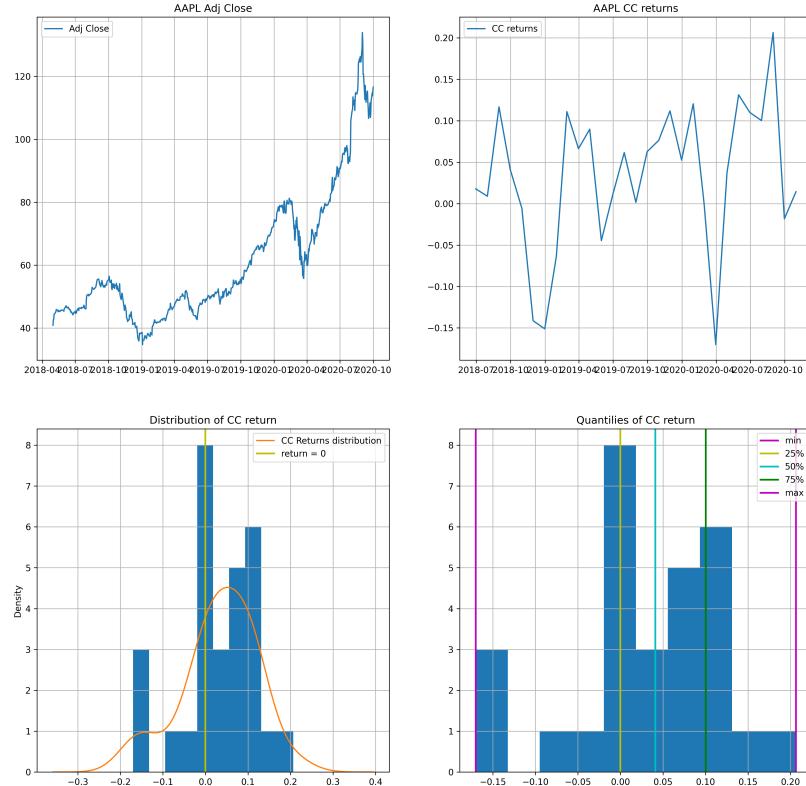


Figure 44: Recap Apple 2018-2020

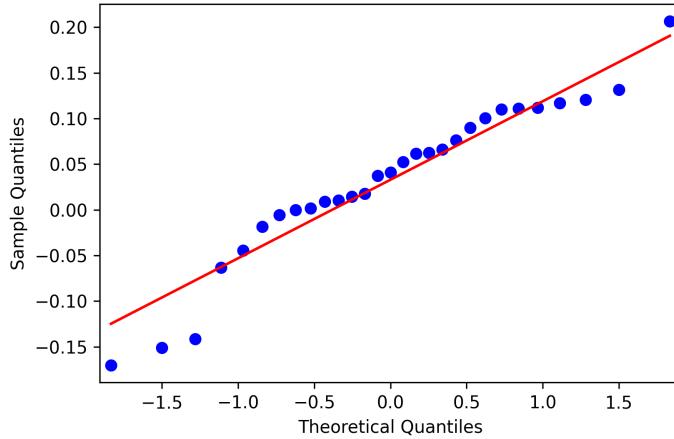


Figure 45: qqplot AAPL

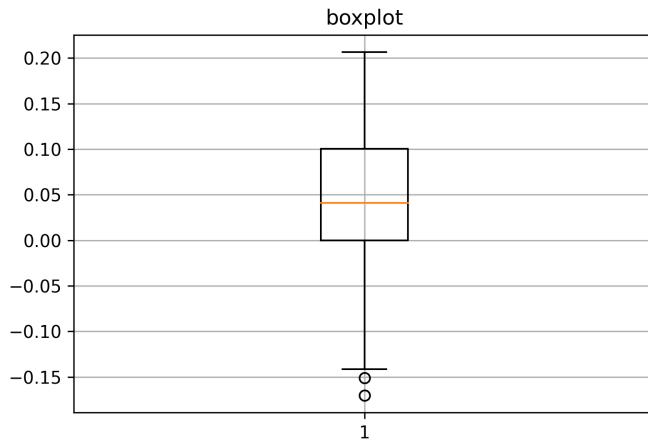


Figure 46: boxplot AAPL

Apple e Nvidia in questo periodo hanno sicuramente ottenuto un ritorno positivo. La loro media dei ritorni è rispettivamente 0.032 e 0.027 come dimostrabile dal picco della curva di densità in zona positiva. In entrambi i casi come accennato nei capitoli precedenti si nota il calo di prezzo repentino a marzo 2020. Questo è sicuramente causato dalla pandemia ma, nonostante questo, il settore tecnologico è stato capace di sopravvivere alla quarantena alzando addirittura il suo valore, come spiegato precedentemente, grazie all'aumento del consumo dei prodotti digitali e di elettronica per smartworking e intrattenimento.

Data la forte oscillazione del prezzo in entrambi i casi la varianza non è tra le più basse: rispettivamente 0.0076 e 0.0129 confermando quindi Nvidia come il titolo più volatile tra i due. Quest'ultima affermazione può essere notata anche confrontando i valori rappresentati nei due relativi boxplot: il box di Nvidia è più alto rispetto a quello di Apple e anche gli outliers assumono valori più bassi sempre rispetto a quelli Apple.

Analizzando infine il qqplot notiamo come i valori dei quantili del titolo di Nvidia si discostano di più dalla normale rispetto ai valori assunti dai quantili del titolo di Apple.

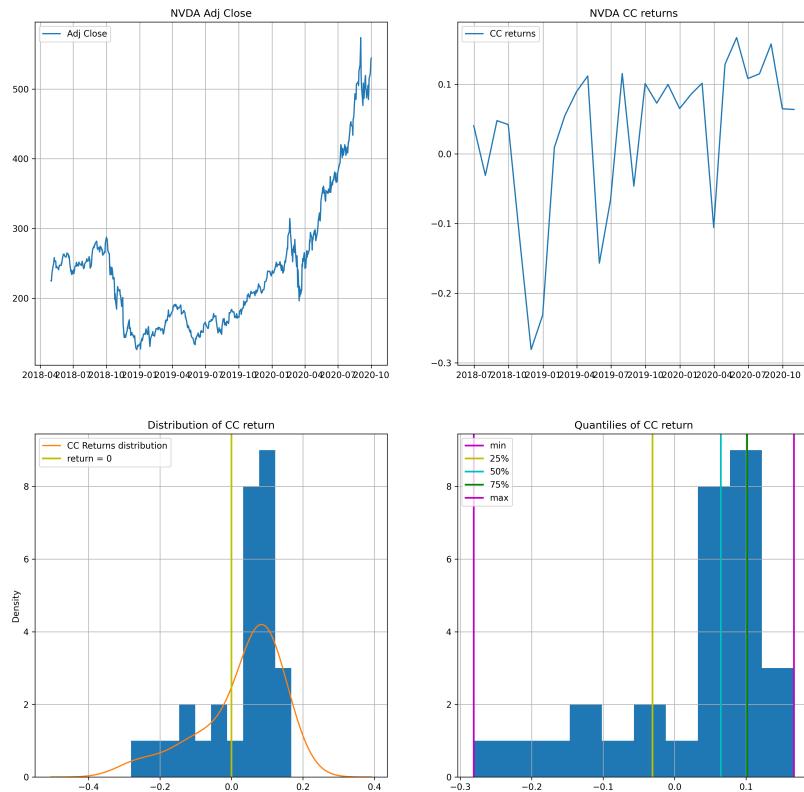


Figure 47: Recap Nvidia 2018-2020

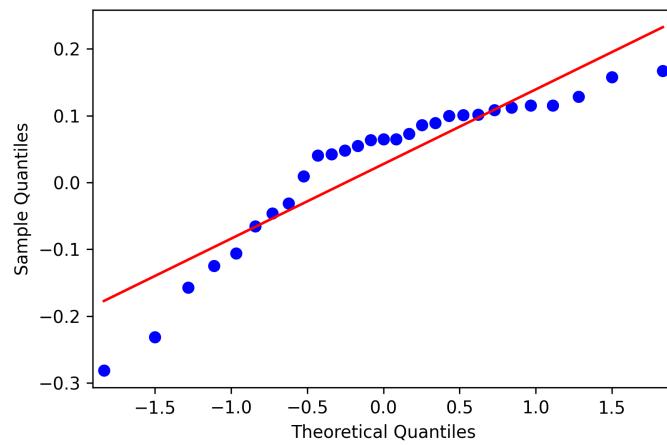


Figure 48: qqplot NVDA

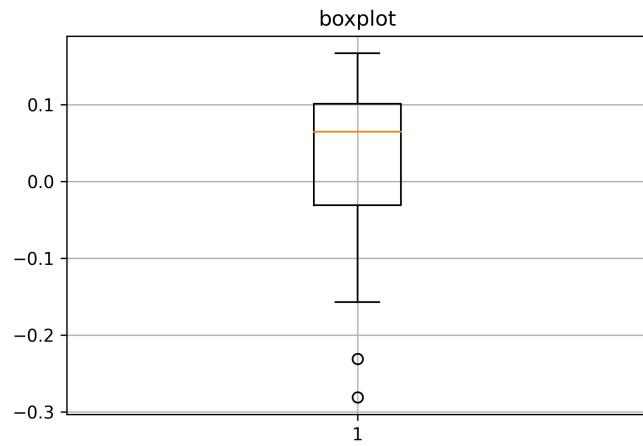


Figure 49: boxplot NVDA

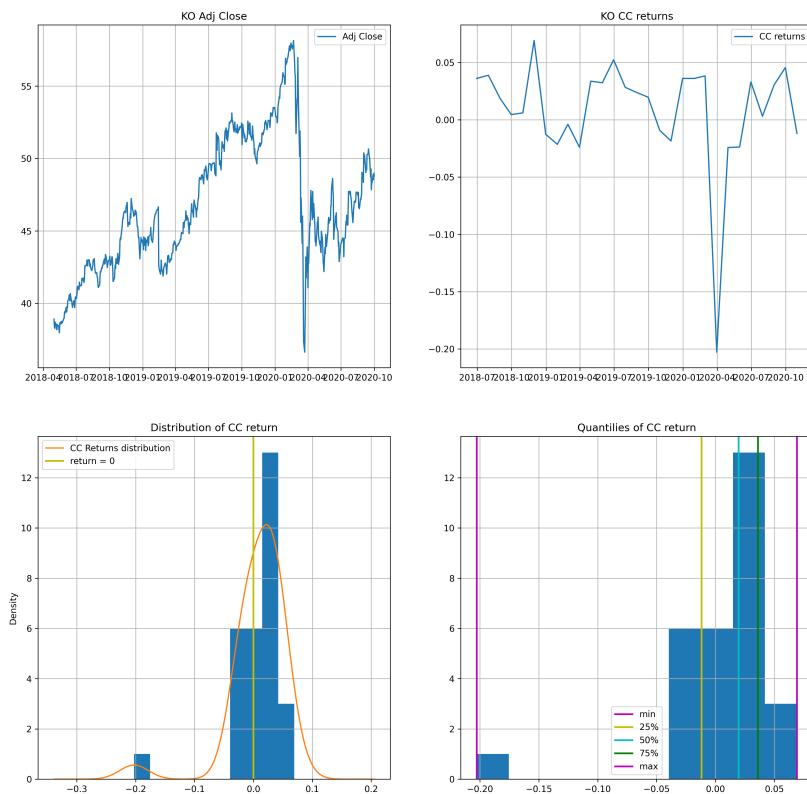


Figure 50: Recap The Coca Cola Company 2018-2020

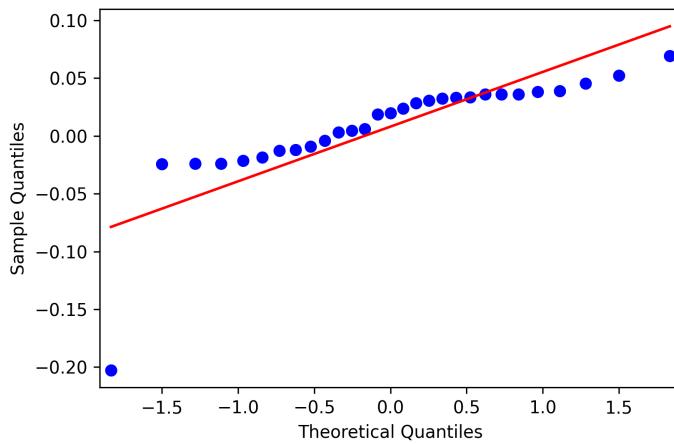


Figure 51: qqplot KO

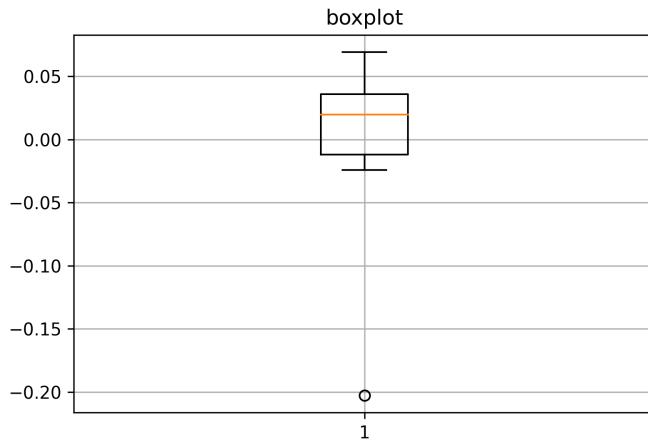


Figure 52: boxplot KO

Il settore alimentare nel periodo 2018-2020 sembra confermarsi come il più stabile tra tutti. Nonostante questo risulta gravemente colpito dal periodo di quarantena molto più di Apple e Nvidia viste precedentemente. Nello specifico Coca Cola, più di Unilever, ha difficoltà nella ripresa.

In entrambi i casi però abbiamo una media dei ritorni positiva anche se di poco. Coca Cola riesce a portare la media a 0.0080 mentre Unilever ha una media di 0.0063. La varianza rispettivamente a 0.0023 e 0.0016 è sicuramente infuenzata dai pochi, seppur presenti, valori anomali dei ritorni assunti durante il crollo dovuto al virus Covid-19. Questi sono sicuramente visibili sia nel qqplot che nel boxplot di riferimento visto che assumono valori molto lontani dalla distribuzione dei ritorni dei rispettivi titoli.

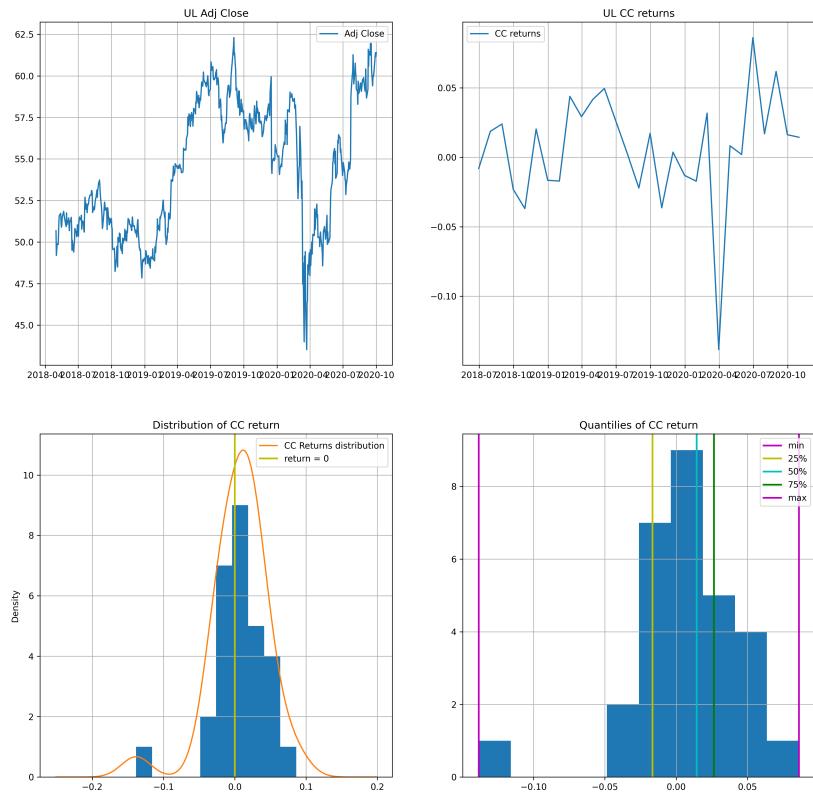


Figure 53: Recap Unilever 2018-2020

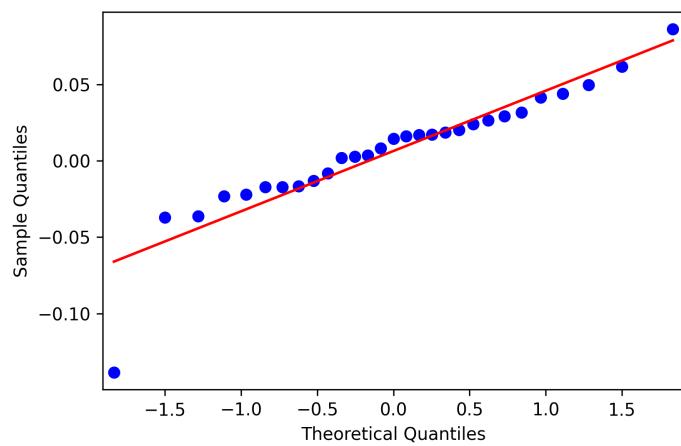


Figure 54: qqplot UL

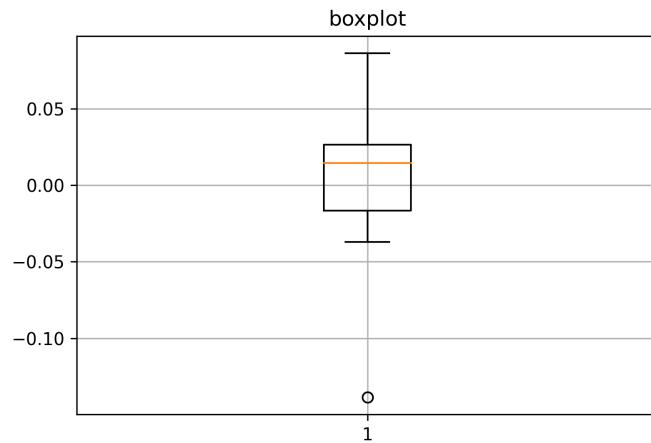


Figure 55: boxplot UL

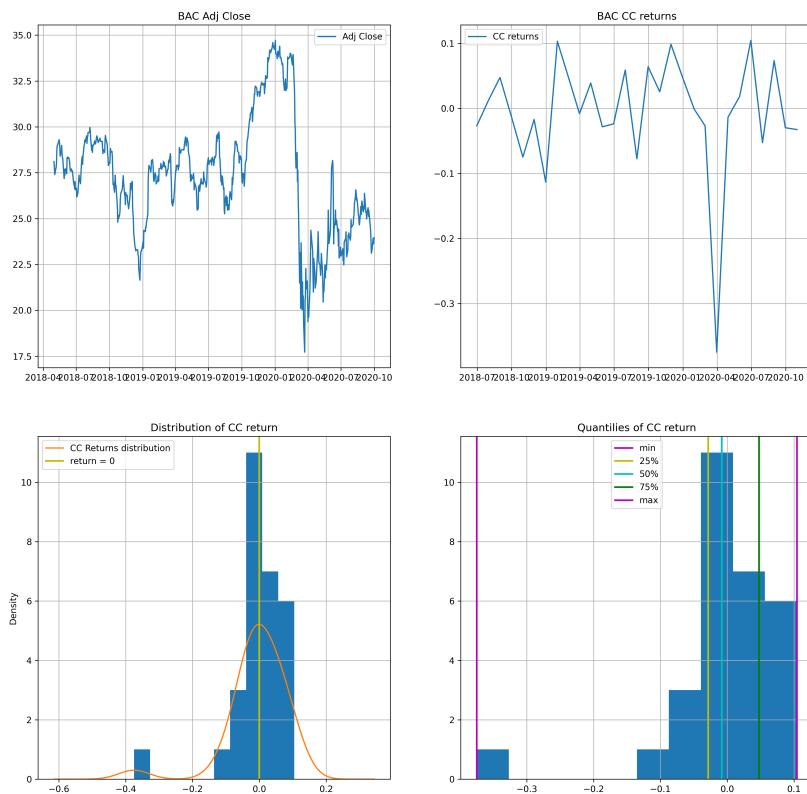


Figure 56: Recap Bank of America 2018-2020

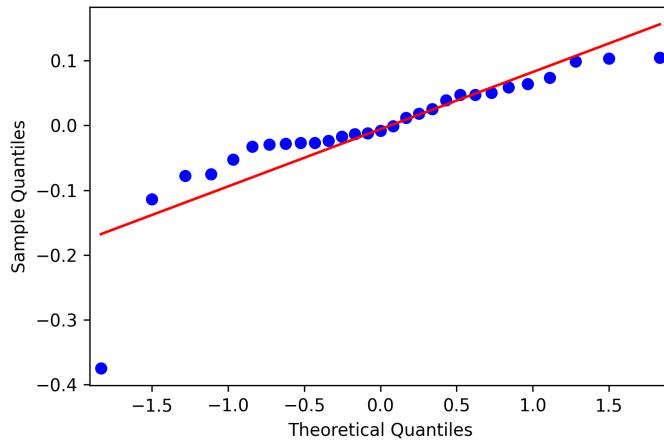


Figure 57: qqplot BAC

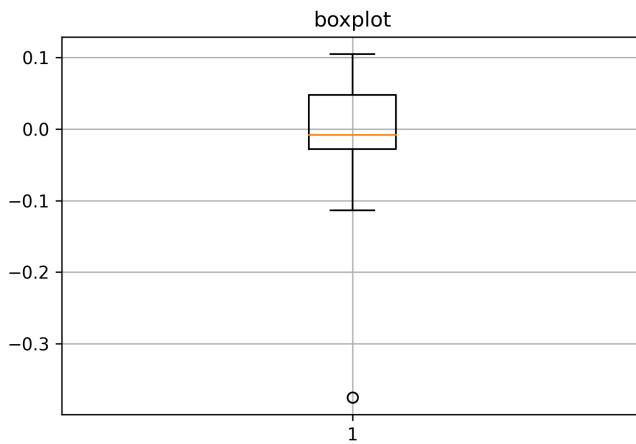


Figure 58: boxplot BAC

Concludiamo infine con il settore bancario-finanziario dove ancora una volta Bank of America esce da una crisi con un ritorno medio negativo. Il valore si attesta a -0.0058 mentre per quanto riguarda American Express si ha un ritorno medio di poco superiore allo zero: 0.0016. La varianza è rispettivamente di 0.0080 e 0.0063 confermando quindi la grande variazione dei prezzi, osservabile anche a occhio nei grafici, avuta in questo periodo. Anche i rispettivi qqplot e boxplot mostrano i valori anormali dovuti alla crisi ovviamente decisamente negativi.

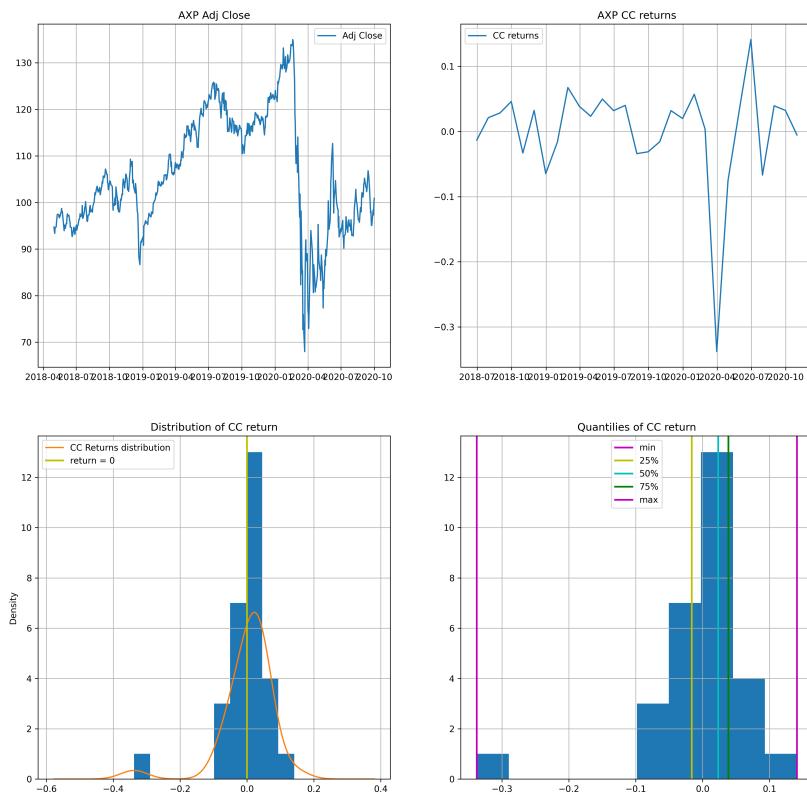


Figure 59: Recap American Express 2018-2020

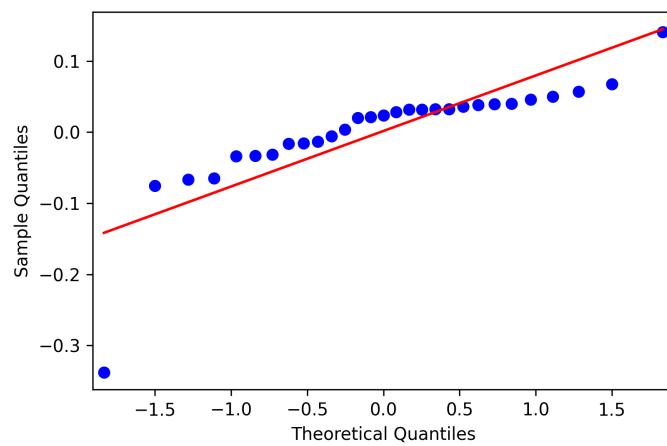


Figure 60: qqplot AXP

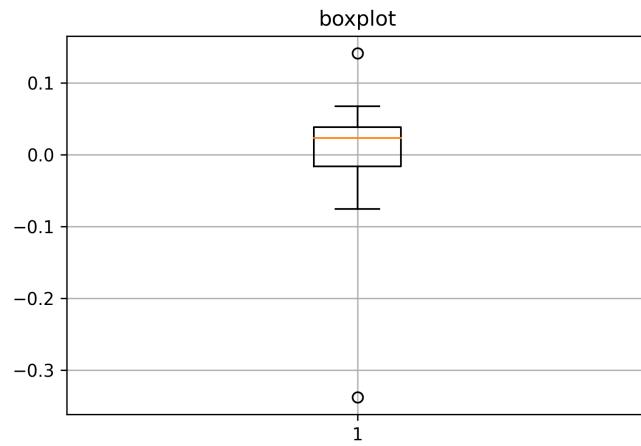


Figure 61: boxplot AXP

In generale questo periodo è sicuramente caratterizzato dalle difficoltà dei titoli azionari rispetto la pandemia globale. È comunque possibile notare però come la maggior parte di queste aziende reagisce in modo prevedibile secondo le analisi svolte nei capitoli precedenti. Nello specifico Nvidia conferma la sua alta volatilità, e più in generale il settore tecnologico si dimostra in forte crescita quando non direttamente colpito da nessuna crisi (come dimostra la sua media dei ritorni superiore agli altri settori analizzati). Il settore alimentare invece si conferma come il più solido tra tutti riuscendo a mantenere basso il più possibile il suo rischio nonostante tutto.

3 Predictive analytics

In questo capitolo proveremo a prevedere l'andamento del mercato attraverso l'utilizzo di tecniche quale ARIMA e SVM. Questo compito è ovviamente semplificato rispetto alle reali metodologie di previsioni dato che risulta difficile prevedere il prezzo di un'azione nel tempo: considerare solo i ritorni passati può risultare riduttivo visto che il prezzo è continuamente influenzato da altre variabili quali il management aziendale, la politica, il comportamento generale del mercato, il rapporto e i gusti dei consumatori e così via. Inoltre analisi più approfondite richiederebbero uno sforzo computazionale notevole che non è a tutti disponibile.

Come anticipato le previsioni saranno effettuate sui ritorni e non sui prezzi di chiusura aggiustati. Questo accade perché i ritorni hanno la caratteristica di essere una serie temporale stazionaria, ovvero la loro distribuzione non cambia ma viene semplicemente traslata nel tempo.

3.1 SVM

Le macchine a vettori di supporto (supported vector machines da cui deriva l'abbreviazione SVM) sono modelli ad apprendimento supervisionato che analizzano i dati attraverso analisi di classificazione e regressione. Oltre a una classificazione lineare le SVMs possono anche utilizzare una classificazione non lineare usando dei kernel trick (trucchi nel kernel) che permettono di mappare l'input in spazi multi-dimensionali.

La previsione con l'utilizzo è così strutturata: vengono utilizzati 120 mesi totali (10 anni). 80 di questi sono dedicati al training del modello, 30 sono per il test del modello e infine gli ultimi 10 mesi sono utilizzati per la previsione del modello per poi successivamente confrontare i dati ottenuti con quelli reali del mercato.

Per prima cosa viene creato un dataset composto da quattro colonne: le prime tre contengono il valore di 3 periodi temporali precedenti alla quarta colonna. Se ad esempio volessi prevedere il valore del mese di aprile (valore posto nella quarta colonna) andrò a guardare i valori nella prima, seconda e terza colonna ed effettuare la previsione in funzione di questi (rispettivamente il mese di gennaio, febbraio e marzo).

Viene poi spezzato il dataset di 120 mesi in uno da 110 mesi dedicati per trainig e test e in uno da 10 mesi finali dedicati per le previsioni finali.

Viene poi importato il modulo `svm` facendo uso della libreria `sklearn` e di questo vengono ottimizzati gli hyperparameters C , γ ed ϵ . Viene effettuata cross validation sul dataset di test (30 mesi) e viene minimizzato il più possibile l'errore quadratico medio.

Una volta trovati gli hyperparametri ideali viene effettuata la previsione sui dati reali, ovvero i 10 mesi finali riservati all'inizio. Di seguito sono riportati i risultati numerici (includendo anche l'errore quadratico medio per una successiva analisi) che vengono ulteriormente rappresentati attraverso un grafico per ogni azione considerata nel progetto:

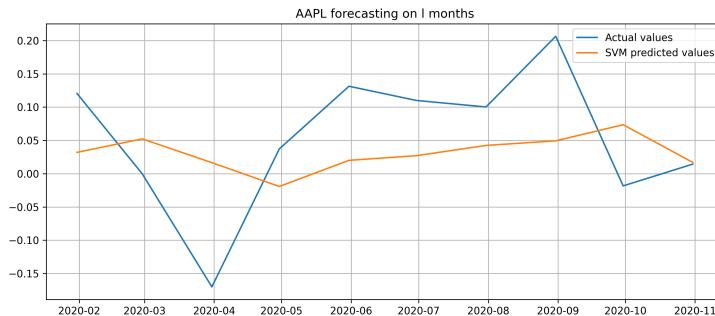


Figure 62: Forecast result AAPL

Date	Actual values	SVM predicted values
2020-01-31	0.120431	0.032004
2020-02-29	-0.000212	0.052295
2020-03-31	-0.170114	0.016539
2020-04-30	0.037183	-0.019236
2020-05-31	0.131362	0.020019
2020-06-30	0.110000	0.027029
2020-07-31	0.100235	0.042448
2020-08-31	0.206391	0.049313
2020-09-30	-0.018389	0.073574
2020-10-31	0.014357	0.016918

Mse between actual values and predicted ones: 0.01043565550277948

Figure 63: Forecast numeric result AAPL

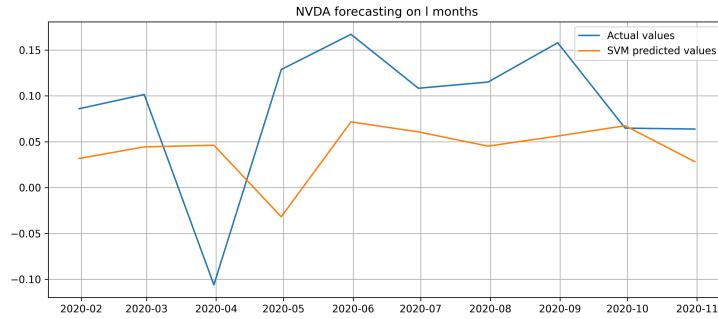


Figure 64: Forecast result NVDA

Date	Actual values	SVM predicted values
2020-01-31	0.085942	0.031712
2020-02-29	0.101437	0.044333
2020-03-31	-0.106000	0.046190
2020-04-30	0.128672	-0.031826
2020-05-31	0.167032	0.071605
2020-06-30	0.108242	0.060691
2020-07-31	0.115042	0.045191
2020-08-31	0.157841	0.056169
2020-09-30	0.064802	0.067441
2020-10-31	0.063806	0.028220

Mse between actual values and predicted ones: 0.008298019593245463

Figure 65: Forecast numeric result NVDA

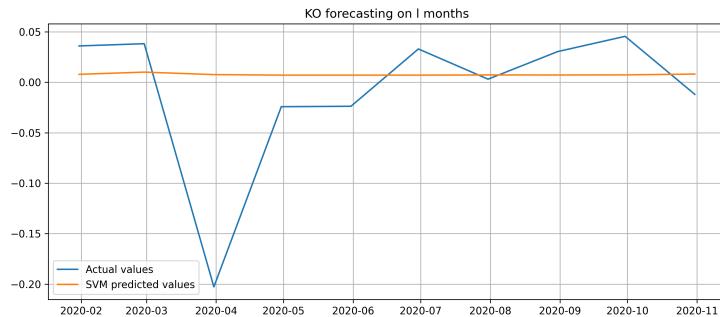


Figure 66: Forecast result KO

Date	Actual values	SVM predicted values
2020-01-31	0.036081	0.007917
2020-02-29	0.038269	0.010111
2020-03-31	-0.202736	0.007649
2020-04-30	-0.024159	0.007115
2020-05-31	-0.023748	0.007115
2020-06-30	0.033059	0.007115
2020-07-31	0.003121	0.007310
2020-08-31	0.030477	0.007243
2020-09-30	0.045589	0.007377
2020-10-31	-0.011911	0.008182

Mse between actual values and predicted ones: 0.005087302809548045

Figure 67: Forecast numeric result KO

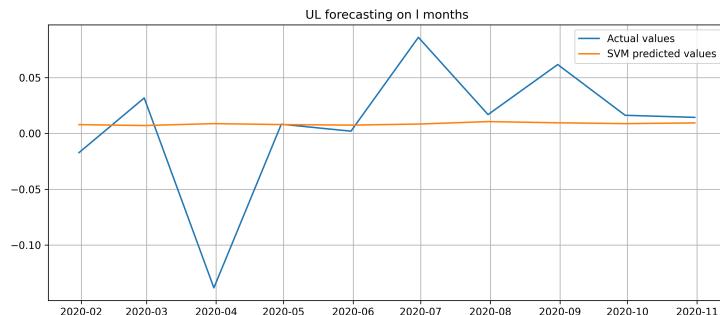


Figure 68: Forecast result UL

Date	Actual values	SVM predicted values
2020-01-31	-0.017225	0.007873
2020-02-29	0.031768	0.007066
2020-03-31	-0.138449	0.008779
2020-04-30	0.008231	0.007915
2020-05-31	0.002004	0.007442
2020-06-30	0.086126	0.008370
2020-07-31	0.016871	0.010602
2020-08-31	0.061743	0.009523
2020-09-30	0.016235	0.008768
2020-10-31	0.014368	0.009398

Mse between actual values and predicted ones: 0.003183863773013345

Figure 69: Forecast numeric result UL

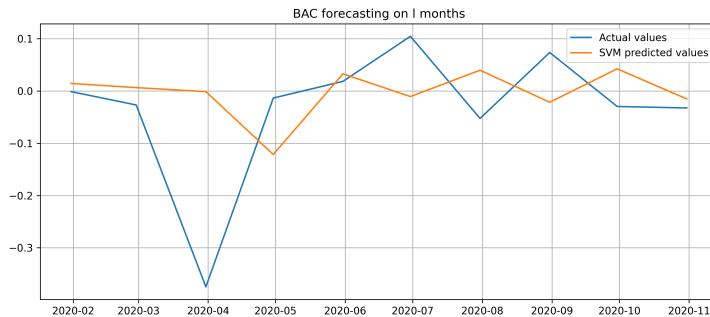


Figure 70: Forecast result BAC

Date	Actual values	SVM predicted values
2020-01-31	-0.001262	0.014366
2020-02-29	-0.026835	0.006372
2020-03-31	-0.374818	-0.001257
2020-04-30	-0.013567	-0.121519
2020-05-31	0.018348	0.032786
2020-06-30	0.104527	-0.010711
2020-07-31	-0.052524	0.039706
2020-08-31	0.073563	-0.021618
2020-09-30	-0.029690	0.042311
2020-10-31	-0.032545	-0.015272

Mse between actual values and predicted ones: 0.01890855715299826

Figure 71: Forecast numeric result BAC

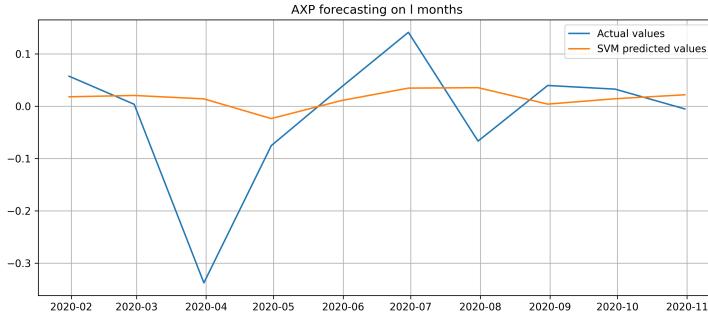


Figure 72: Forecast result AXP

Date	Actual values	SVM predicted values
2020-01-31	0.057201	0.017846
2020-02-29	0.003516	0.020380
2020-03-31	-0.337853	0.013855
2020-04-30	-0.075495	-0.023786
2020-05-31	0.035784	0.010447
2020-06-30	0.141077	0.034413
2020-07-31	-0.066868	0.035287
2020-08-31	0.039466	0.003847
2020-09-30	0.032466	0.014118
2020-10-31	-0.005352	0.021681

Mse between actual values and predicted ones: 0.015299670470027829

Figure 73: Forecast numeric result AXP

Come precisato all'inizio di questo capitolo le analisi sono sommarie e difficilmente riescono a prevedere con esattezza l'andamento del mercato. Non per questo però sono totalmente inutili: è possibile notare come nel caso delle azioni nel mercato alimentare (figure 67 e 69) il modello, addestrato sugli anni passati, sia stato conservativo sui movimenti dell'azione prevedendo un'oscillazione minima. Questo ovviamente non è accaduto, non perché il modello fosse sbagliato, ma perché è praticamente impossibile prevedere una crisi mondiale dovuta a una pandemia totalmente inaspettata.

La stessa dinamica può essere notata anche sui titoli tecnologici (figure 63 e 65): il modello aveva certamente previsto una certa volatilità basandosi sugli anni precedenti ma come è possibile notare durante il periodo di marzo anche in questo caso non è riuscito a prevedere il crollo improvviso della borsa dovuto alle notizie su Corona Virus. Successivamente al periodo di marzo per il modello non è stato certamente semplice prevedere una forte salita dei titoli tecnologici grazie proprio alla stessa pandemia che a Marzo aveva causato loro una grave perdita.

3.2 ARIMA

Il modello Autoregressive integrated moving average (o in italiano Modello autoregressivo integrato a media mobile) permette di indagare serie temporali.

La parte auto regressiva (AR) utilizza la relazione di dipendenza tra un'osservazione e un certo numero di osservazioni ritardate (gestita dal parametro p), la parte di media mobile (MA) utilizza la dipendenza tra un'osservazione e un errore residuo da un modello a media mobile applicato alle osservazioni ritardate (gestita dal parametro q) e infine la parte integrata che unisce i due modelli e permette di definire la differenza di un'osservazione da un'osservazione nella fase temporale precedente (parametro d).

Per trovare il valore migliore tra questi parametri elencati è stato implementato un algoritmo che, provando ogni combinazione possibile tra i valori dati, sceglie la combinazione di parametri migliori basandosi sul minimo RMSE (root mean square error) trovato. L'algoritmo è illustrato in figura 74

```
In [49]: # finding best parameters:
best_found = (0, 0, 0)
min_rmse = np.inf

import warnings
from statsmodels.tools.sm_exceptions import ConvergenceWarning
warnings.simplefilter('ignore', ConvergenceWarning)

import time
start = time.time()
# misuro il tempo di esecuzione

for p in range(0, 10):
    for d in range(0, 3):
        for q in range(0, 8):
            # order = (number of autoregressor, _, numbers of moving averages)
            model = sm.tsa.statespace.SARIMAX(ccrs[id_stocks], order=(p, d, q),
                                              enforce_stationarity=False,
                                              enforce_invertibility=False)
            results = model.fit()

            # trying to predict the present values
            # let the data start from gen 2018 (last day of 2017)
            # dynamic parameter: True if I want to predict basing my guessing on real data, False otherwise
            pred = results.get_prediction(start=pd.to_datetime('2017-12-31'), dynamic=False)
            # calculating performance using Root Mean Squared Error and Mean Squared Error
            # closer to 0 -> better
            rmse = np.sqrt(((pred.predicted_mean - ccrs[id_stocks])**2).mean())
            if(rmse < min_rmse):
                min_rmse = rmse
                best_found = (p, d, q)

display(results.summary())
print(f'Best found = {best_found}')
print(f'rmse = {min_rmse}')
end = time.time()
print(f'Execution time: {end - start}'')
```

Figure 74: Algorithm for searching optimal hyperparameters of an ARIMA model

Nelle figure 75, 77, 79, 81, 83, 85 vengono mostrati i risultati ottenuti con i parametri trovati, mentre nelle figure 76, 78, 80, 82, 84, 86 vengono mostrati i risultati delle previsioni sui ritorni futuri.

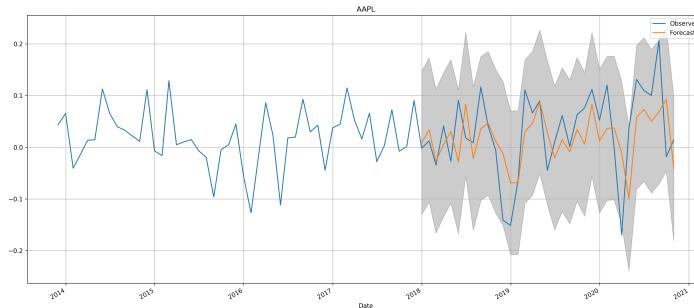


Figure 75: Test result AAPL

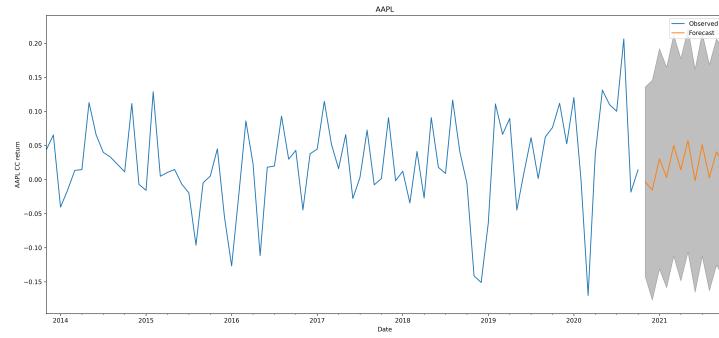


Figure 76: Forecasting APPL

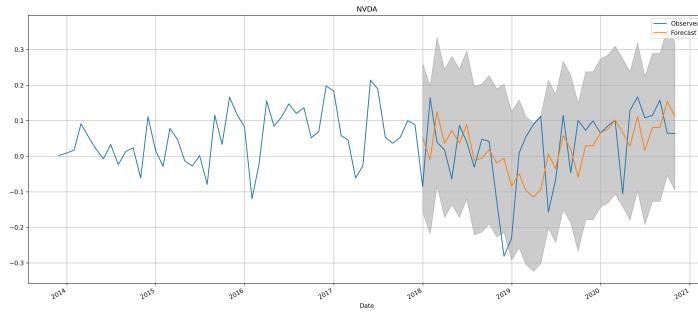


Figure 77: Test result NVDA

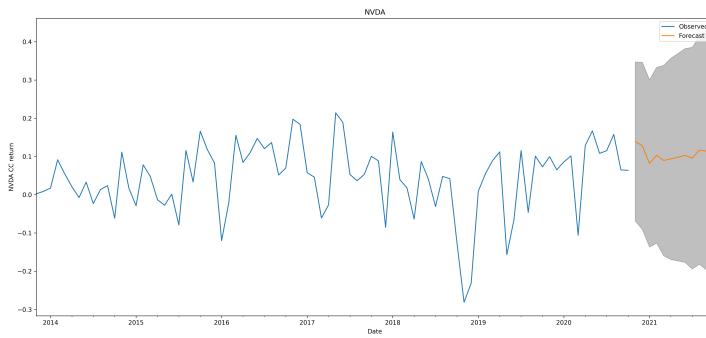


Figure 78: Forecasting NVDA

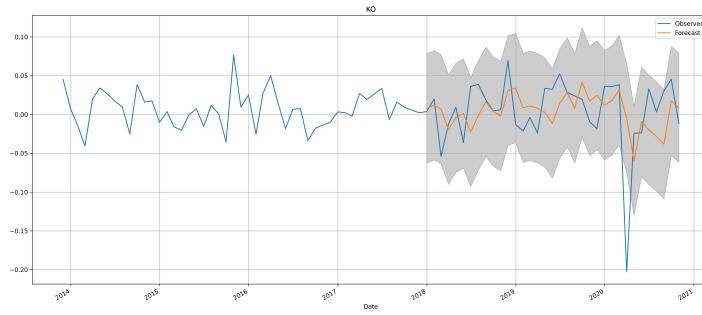


Figure 79: Test KO

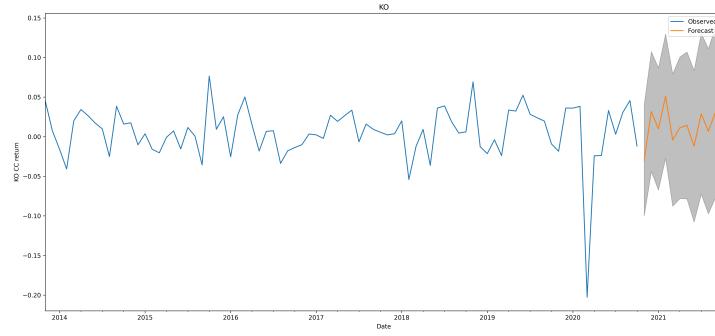


Figure 80: Forecasting KO

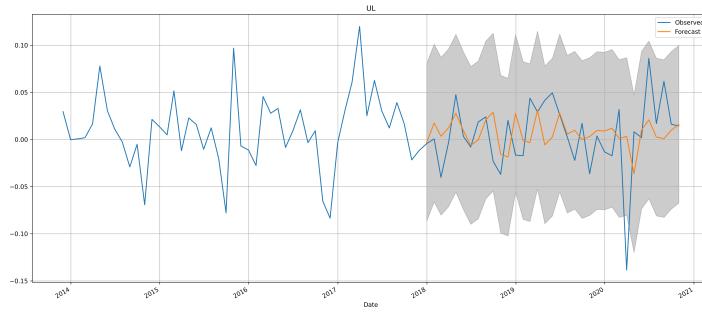


Figure 81: Test UL

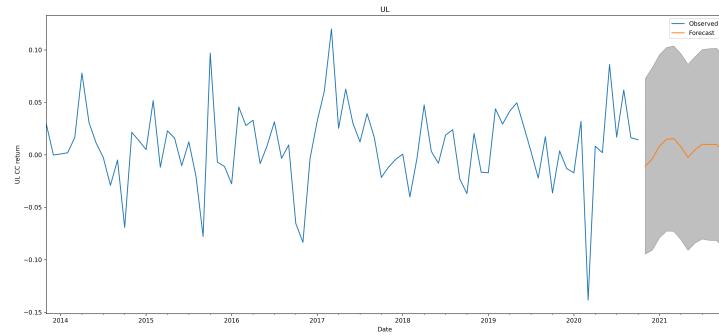


Figure 82: Forecasting UL

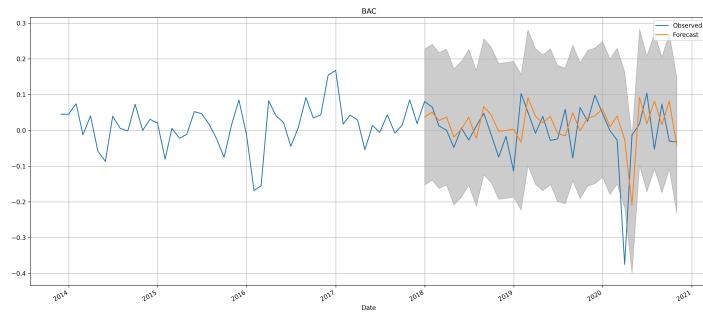


Figure 83: Test BAC

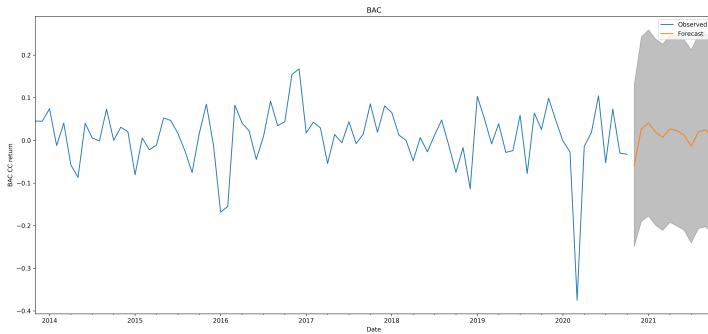


Figure 84: Forecasting BAC

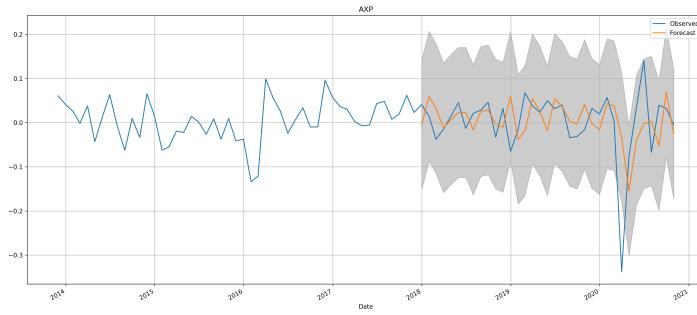


Figure 85: Test result AXP

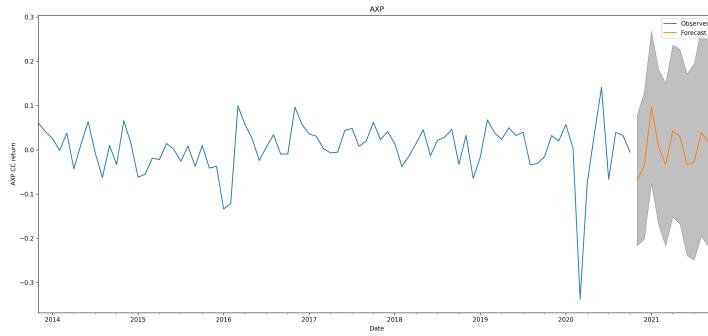


Figure 86: Forecasting AXP

I risultati illustrati nella tabella di seguito mostrano come il titolo di Nvidia, e subito dopo quello di Bank of America, siano i più difficili da prevedere data la loro volatilità. I titoli con un RMSE invece si confermano KO e UL che, come visto, hanno una volatilità molto minore

Share	RMSE	Parameters
AAPL	0.068013	(8, 1, 5)
NVDA	0.093638	(5, 0, 5)
KO	0.041243	(9, 0, 6)
UL	0.036704	(5, 1, 6)
BAC	0.082900	(0, 0, 0)
AXP	0.066467	(2, 0, 4)

4 Portfolio Management and Optimization

4.1 Portfolio investment using Markowitz model

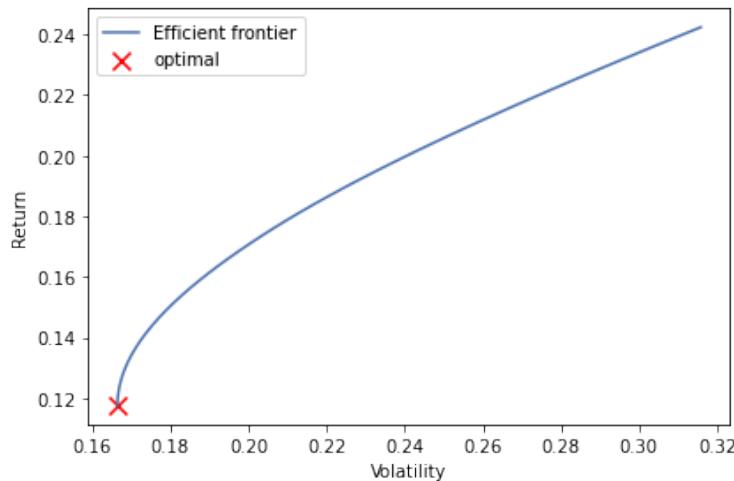
Per la creazione del portfolio ci baseremo sui tre principi di Harry Markowitz economista premio Nobel del 1952. Il primo tra questi afferma che per costruire un portfolio un investitore si basa su due dati fondamentali: rendimento medio atteso e rischio medio atteso. Il secondo principio porta l'investitore a considerare il periodo dell'investimento come uniperiodale: i movimenti del portfolio durante il periodo dell'investimento non vengono quindi considerati. L'ultimo principio, ma non ultimo per importanza, afferma che l'investitore deve essere in ogni modo avverso al rischio cercando di minimizzarlo il più possibile.

Da questi principi è possibile formulare un problema matematico nello specifico un problema di ottimizzazione quadratica:

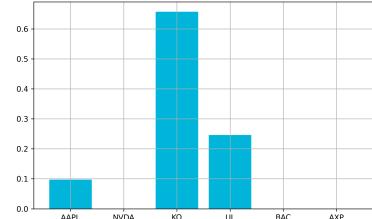
$$\begin{aligned} \min_w \quad & w' C w \\ \text{s.t.} \quad & w' \mu = r^* \\ & \sum_{i=1}^N w_i = 1 \end{aligned} \tag{2}$$

Bisogna quindi trovare i pesi $w = (w_1, w_2, \dots, w_n)$ tali che, dato un tasso di rendimento r^* , il ritorno atteso del portfolio determinato da w è r^* e la sua varianza è minima.

Possiamo quindi variare il tasso di rendimento r^* ottenendo una frontiera efficiente:



	Stock	Weight
1	AAPL	0.09707
2	NVDA	0.00000
3	KO	0.65730
4	UL	0.24564
5	BAC	0.00000
6	AXP	0.00000



Ottenuti i pesi del portfolio si può procedere nella costituzione dello stesso calcolando quante azioni comprare con il budget massimo considerato. Prima di questo però è possibile fare qualche considerazione sui pesi trovati. Come è semplice notare vengono solamente selezionate quote di Apple, Coca Cola Company

e Unilever. Questo comportamento prevedibile è sicuramente dovuto al buon rapporto tra ritorni e volatilità di Apple e soprattutto alla bassa varianza delle due aziende nel settore alimentare.

Nei calcoli successivi vengono considerati dei costi di transazione all'1,5% come riportati in tabella 87.

	Num of shares	Price for each share \$	Transaction costs \$	Purchase cost \$
AAPL	24	38.52	0.57780	938.35
NVDA	0	132.76	1.99140	0.00
KO	145	44.38	0.66570	6531.63
UL	49	49.04	0.73560	2439.00
BAC	0	23.46	0.35190	0.00
AXP	0	92.29	1.38435	0.00

The total spent on the investment is 9908.98\$
 You have 91.02\$ left

Figure 87: M. optimal portfolio composition

Vengono infine calcolati i ritorni attesi e, grazie al dataset anche i ritorni reali del portfolio. Nell'esempio riportato in figura 87 l'investimento iniziale è pari a 9908.98 dollari, il valore del portfolio alla fine del periodo di investimento è invece di 12365.90 dollari pari al 26.83% di ritorno reale rispetto all'11.76% di ritorno atteso.

4.2 Portfolio investment using CAPM

È inoltre possibile calcolare i ritorni attesi di un portfolio attraverso l'utilizzo del valore Beta di ogni azione. In questo caso viene utilizzata la formula $r_i = \beta_i * (r_m - r_f) + r_m$ dove r_i e β_i sono rispettivamente il ritorno atteso della security scelta, r_f è il risk free rate (impostato in questo esempio a un valore simbolico di 0.01) e infine r_m è il ritorno del market index (rappresentato in questo esempio dall'indice Standard and Poor 500).

Il risultato ottenuto usando quindi la formula del Capital Asset Pricing Model e gli stessi pesi trovati precedentemente è un ritorno atteso minore rispetto a quello ottenuto dal portfolio di Markowitz. Il suo valore si attesta a circa 5.83%

4.3 Portfolio investment using forecasted values

Infine viene effettuata una simulazione di investimento basando il calcolo dei ritorni sui valori predetti nel capitolo precedente. In questo caso ci aspettiamo valori decisamente meno accurati perché, come precedentemente spiegato, le previsioni effettuate hanno unicamente un fine dimostrativo e non sono affidabili in alcun modo per un reale investimento.

In questo caso la simulazione verrà effettuata sugli stessi l mesi usati per il forecasting effettuato con SVM nel capitolo precedentemente. Per comodità sono riportati nella tabella di seguito i ritorni reali e quelli basati sui valori predetti dal modello SVM.

	Stock	Actual return	Forecasted return
1	AAPL	0.5312	0.3130
2	NVDA	0.8868	0.4197
3	KO	-0.0759	0.0785
4	UL	0.0816	0.0856
5	BAC	-0.3348	-0.0348
6	AXP	-0.1760	0.1480

Ipotizzando un budget iniziale di 10000\$ e usando i pesi trovati nella soluzione al problema di ottimizzazione di Markowitz vista precedentemente

	Stock	Weight
1	AAPL	9.7%
2	KO	65.7%
3	UL	24.6%

otteniamo un ritorno lordo di 0.20239 se si considerano i ritorni reali e di 0.10299 se si considerano i ritorni previsti dal modello SVM. I seguenti calcoli vengono effettuati considerando una fee per il costo di transazione pari all'1.5%

	Value
Gross initial investment	10000.00\$
Net initial investment	9850.00\$
Real investment gross value	12023.90\$
Real investment net value	11843.54\$
Forecasted investment gross value	11029.90\$
Forecasted investment net value	10864.45\$

5 Web Applet

Le analisi descrittive condotte nel secondo capitolo e le operazioni di portfolio optimization usando il modello di Markowitz vengono mostrate attraverso l'utilizzo di una web applet. Di seguito viene descritto brevemente il funzionamento di entrambe:

5.1 Descriptive Analytics

Lo script *main_descr.py* fa partire un web server in localhost sulla porta 8050 che mostrerà un'interfaccia dove saranno presenti diverse informazioni riguardanti l'analisi descrittiva di una specifica azione.

Una volta selezionata l'azione di nostro interesse dal menu dropdown, verranno mostrati i prezzi di chiusura aggiustati, i continuous compound return, un istogramma sulla distribuzione dei ritorni e un box plot sempre rappresentati i ritorni visti precedentemente. In una tabella finale riassuntiva invece vengono mostrati diversi valori descrittivi riguardo i ritorni dell'azione come il conteggio dei dati (raggruppati mensilmente), la loro media, deviazione standard e i diversi quantili.

5.2 Portfolio Management and Optimization

Lo script *main_portf.py*, come il precedente, fa partire un web server in localhost sulla porta 8050 che mette a disposizione uno strumento per calcolare le diverse soluzioni al modello di Markowitz per la costituzione di un portfolio. Nello specifico una volta selezionato il budget attraverso lo slider e un range di date riguardo l'inizio e la fine del proprio investimento sarà possibile visualizzare, cliccando sull'apposito pulsante refresh, i valori calcolati dal modello. Vengono anche automaticamente applicati dei costi di transazione sugli acquisti pari all' 1,5%.

I dati rappresentati riguardano il ritorno atteso e la volatilità del portfolio ottimo trovato. In una tabella vengono presentate il numero di azioni per ogni titolo, il loro costo unitario, il costo della transazione e il prezzo totale per l'acquisto. Viene successivamente indicato il costo totale dell'investimento, l'eventuale resto dal budget iniziale e infine il ritorno reale dell'investimento basato sui dati reali.

References

- [1] AmEx shares fall as 2016 earnings outlook disappoints (Reuters),
<http://reut.rs/3hx5vUr>