

Progetto finale BISF

Con Python

Emanuele Ceglia | Mat. 837255 | Business Intelligence per i Servizi Finanziari

INDICE

* Data summary
* Descriptive analysis
* Predictive analysis
* Portfolio management
* Conclusione

**Data Summary**

Lo scopo di questa relazione, a cui è stata affiancata la creazione di un’applicazione web, consiste in un’analisi dettagliata delle serie storiche dei prezzi azionari di sei società attraverso le metodologie analitiche apprese durante le lezioni pratiche e teoriche del corso.

Le sei società in questione devono appartenere a tre settori economici differenti (due società per ogni settore), personalmente come settori ho selezionato: retail, trasporti ed energetico in quanto vengono utilizzati praticamente da chiunque e quindi di fondamentale importanza per il progresso economico dei paesi e per lo svolgimento delle attività di qualsiasi persona.

**Retail**

Per quanto riguarda il settore retail ho selezionato le più grosse società per capitalizzazione nel corrispettivo segmento di mercato, con ciò sottolineo il fatto che la prima opera nella vendita al dettaglio attraverso negozi fisici mentre la seconda attraverso l’e-commerce, le due società sono:

**Walmart (WMT)**

Descrizione: la Walmart Stores Inc è una multinazionale statunitense, proprietaria dell'omonima catena di negozi al dettaglio Walmart, fondata da Sam Walton nel 1962. È la più grande catena al mondo nel canale della grande distribuzione organizzata con, al 31 luglio 2020, 11.496 negozi e club in 27 paesi.

Quotata in borsa il 30 gennaio del 1981, capitalizza oggi circa 414 miliardi di dollari, nel 2020 ha fatturato circa 524 miliardi di dollari con un utile pari a 20,5 miliardi.

**Amazon (AMZN)**

Descrizione: Amazon.com, Inc. è un'azienda di commercio elettronico statunitense, con sede a Seattle nello stato di Washington. È la più grande Internet company al mondo.

Quotata in borsa il 16 maggio 1997, capitalizza oggi circa 1,65 triliardi di dollari, al 30 settembre 2020 ha fatturato circa 260 miliardi di dollari con un utile pari a 14 miliardi.

**Trasporti**

Il settore dei trasporti è stata una scelta un po' azzardata in quanto sta soffrendo in maniera particolarmente evidente la crisi causata dall’emergenza sanitaria Covid-19, questo perché in un’ottica di pandemia gli spostamenti sono stati ridotti al minimo necessario, riducendo notevolmente le entrate di tali società, le quali devono sostenere costi aziendali molto importanti.

Per questo motivo ho selezionato le due società principali per quanto riguarda la produzione di aeromobili e crociere, inoltre in uno scenario futuro ottimistico possiamo credere che tale settore potrà ritornare ai regimi pre-emergenza. Le due società che ho selezionato sono:

**Airbus (AIR.PA)**

Descrizione: Airbus è un costruttore multinazionale europeo di aeromobili con sede in Francia. È il primo produttore di aerei civili al mondo per numero di consegne avendo superato la rivale americana Boeing nella prima metà del 2019.

Quotata in borsa il 2 luglio 1999, capitalizza oggi circa 67,17 miliardi di euro, al 29 settembre 2020 ha fatturato circa 30 miliardi di euro con una perdita di esercizio di circa 1 miliardo.

**Carnival Corporation (CCL)**

Descrizione: Carnival Corporation & plc, società anglo-statunitense, è il più grande operatore al mondo nel settore delle crociere.

Quotata in borsa il 24 luglio 1987, capitalizza ad oggi 21,27 miliardi di dollari, il 30 settembre 2020 il suo fatturato ammontava a circa 5,5 miliardi di dollari con una perdita di esercizio pari a circa 2 miliardi di dollari.

**Energetico**

Infine, il settore energetico, ho selezionato due società operanti nell’energia rinnovabile in quanto è innegabile che in uno scenario futuro bisognerà convertirsi il più possibile all’utilizzo di fonti energetiche sostenibili per l’ambiente, dando alle aziende operanti oggi in questo settore grosse opportunità di crescita. Le due società che ho selezionato sono:

**Canadian Solar (CSIQ)**

Descrizione: Canadian Solar Inc. è una società quotata in borsa che produce moduli fotovoltaici solari e gestisce progetti solari su larga scala.

Quotata in borsa il 10 novembre 2006, capitalizza ad oggi 3,74 miliardi di dollari, il 29 settembre 2020 il suo fatturato ammontava a circa 2,4 miliardi di dollari con un utile di esercizio di circa 218 milioni di dollari.

**Iberdrola (IBDRY)**

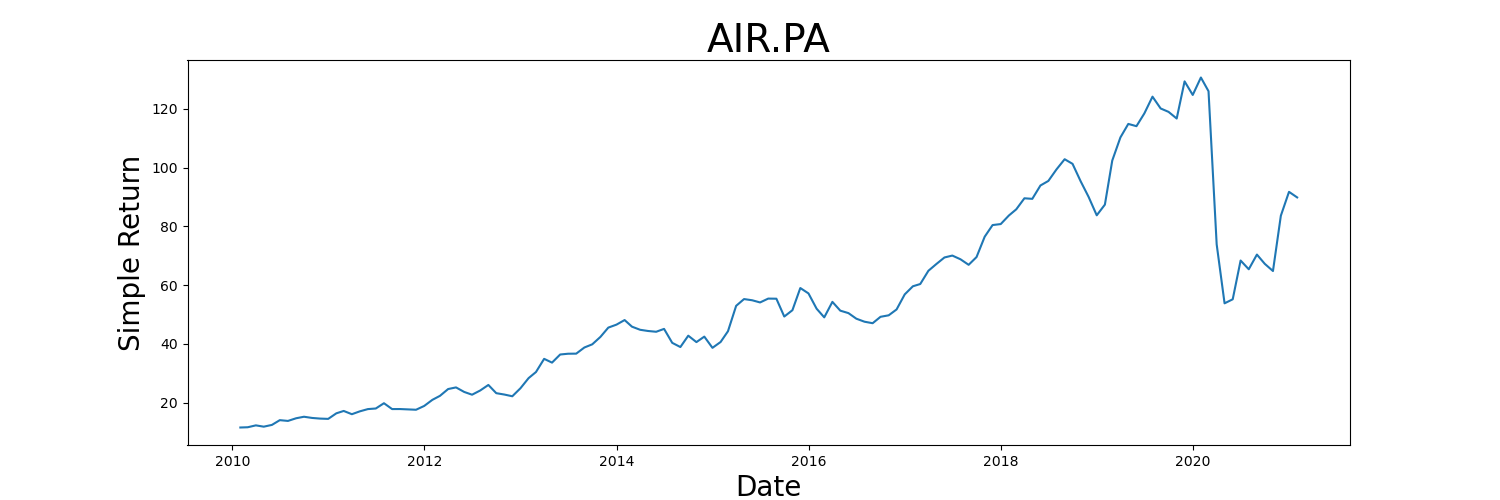
Descrizione: Iberdrola è un'azienda spagnola specializzata nella produzione, distribuzione e commercializzazione di energia elettrica e gas naturale. La direzione generale della società è situata a Bilbao, in Spagna.

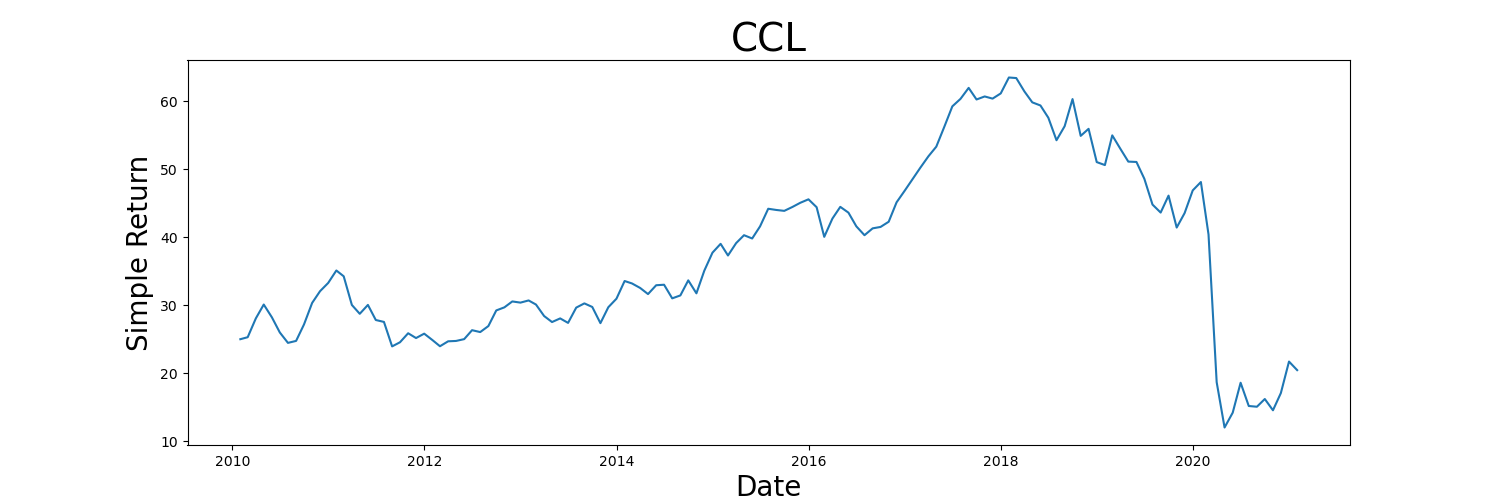
Quotata il 4 gennaio 2002, capitalizza ad oggi circa 73 miliardi di euro, il 29 settembre 2020 il suo fatturato ammontava a circa 24 miliardi di euro con un utile di circa 2,7 miliardi.

**Descriptive Analysis**

Osserviamo come primo step i ritorni semplici degli ultimi 10 anni con granularità mensile delle società che abbiamo scelto:

**Settore trasporti**



****

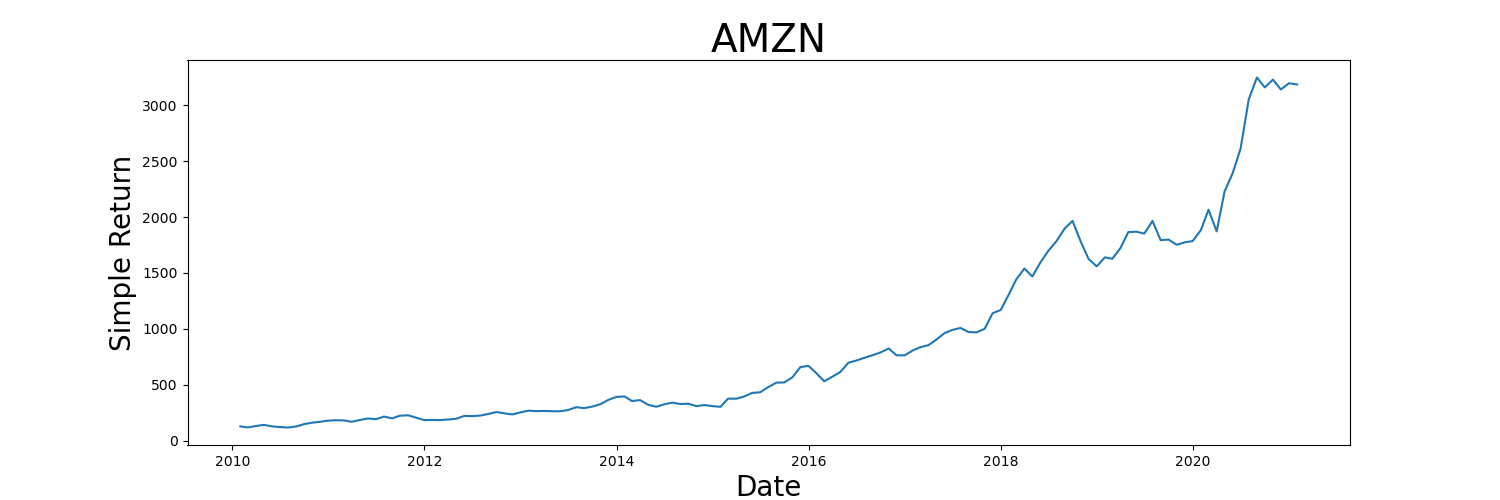
A prima vista possiamo subito osservare come il trend di Airbus sia più crescente rispetto a Carnival Corporation, entrambe le società tuttavia non sono riuscite a ritornare ai livelli pre-crash dei mercati di metà marzo, Carnival Corporation, rispetto all’ 01-01-2010, ha un ritorno addirittura negativo in quanto si può osservare che il prezzo ad oggi si trova sotto al livello iniziale.

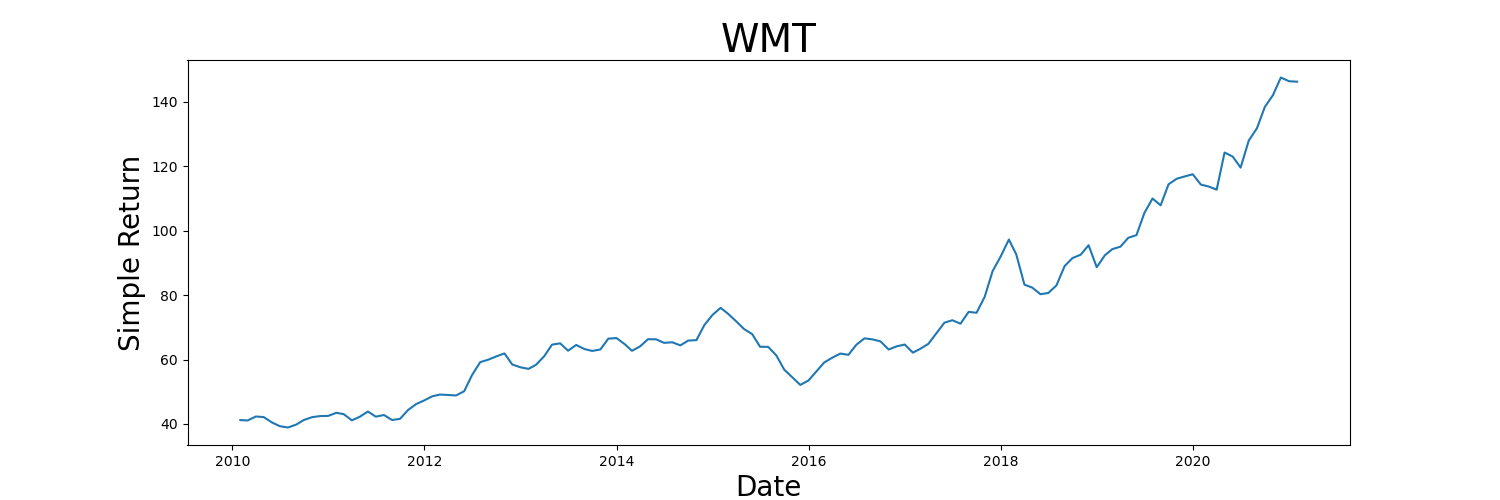
In generale i due grafici sembrano abbastanza correlati tra di loro, l’unica differenza sostanziale possiamo notarla dal 2018 al 2020, infatti in questo arco temporale di circa due anni Airbus ha avuto un trend crescente mentre Carnival ha avuto un andamento decisamente decrescente.

Entrambi i grafici presentano un piccolo spike a novembre 2020 causato dall’annuncio dei vaccini di Pfizer e Moderna, inoltre sempre in quel periodo l’annuncio della vittoria da parte di Biden delle elezioni USA è stato scontato positivamente dai mercati finanziari.

Per quanto riguarda la volatilità non si riesce a trarre una conclusione comparando i due grafici ad occhio.

**Settore retail**





Entrambe le società nell’arco di dieci anni hanno avuto un trend molto positivo, Amazon sembrerebbe avere una crescita addirittura esponenziale, ciò potrebbe far supporre che esse siano anche fortemente correlate tra di loro.

È curioso osservare come il crash dei mercati di marzo 2020 sia praticamente “invisibile” rispetto alle due società legate al settore dei trasporti, questo fenomeno possiamo cercare di motivarlo focalizzandoci sul business delle due società:

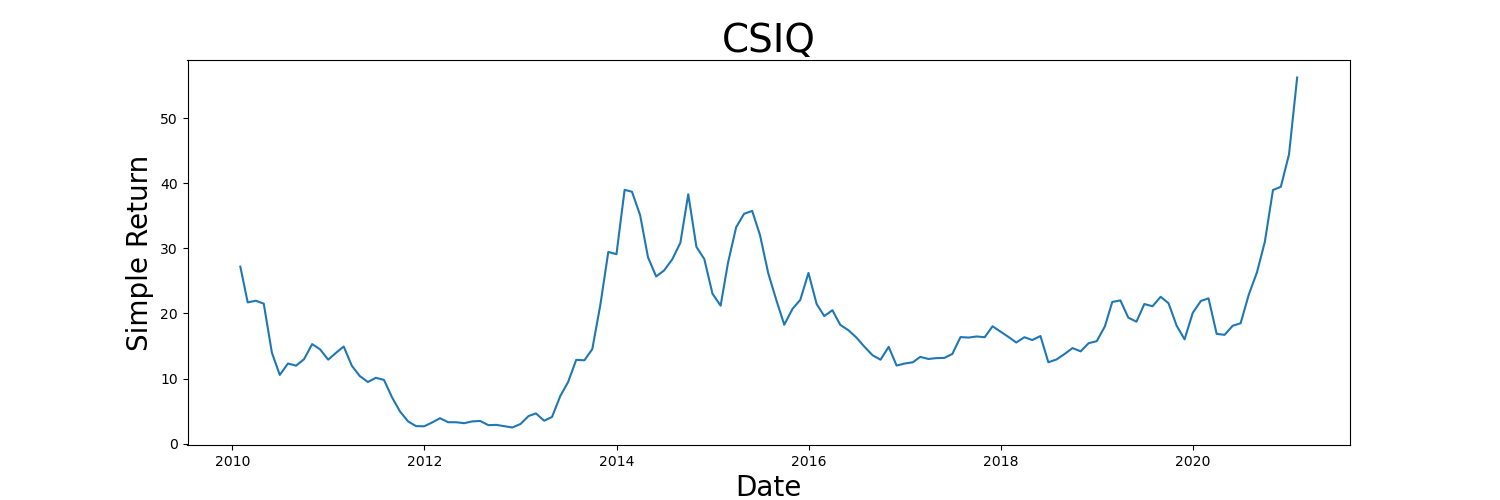
Il settore retail, infatti, al posto di subire danni a causa dei vari lockdown, ha visto un aumento dei propri fatturati i quali sono stati immediatamente scontati dal mercato (basti notare come si sia impennato il prezzo di Amazon e Walmart durante il 2020).

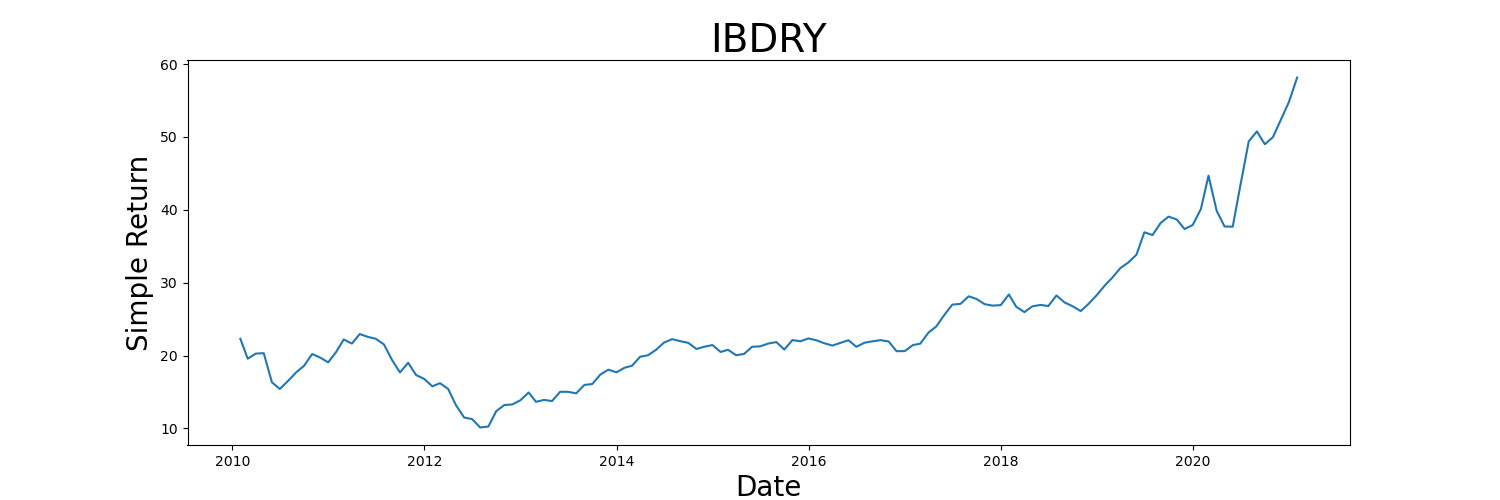
L’aumento del fatturato di Walmart trova risposta nella reazione delle persone (non sono negli USA, ma anche nel resto del mondo) all’annuncio del lockdown e delle varie restrizioni causando dei veri propri “supermarkets-run” alimentati dal panico delle persone.

Per quanto riguarda Amazon invece il suo punto forte è stato proprio l’e-commerce, infatti, in una situazione di distanziamento sociale forzato e paura di contagio il poter acquistare beni sul web al posto di dover uscire di casa ha favorito società operanti nel business online.

**Settore energetico**





Anche in questo caso le due società non hanno subito danni dal crash dei mercati di marzo 2020, se Iberdrola ha avuto un trend sostanzialmente crescente nell’arco dei dieci anni passati, Canadian Solar è stata invece molto più volatile nel corso di questo periodo. Tuttavia, entrambe le società si trovano ora ai massimi livelli di capitalizzazione.

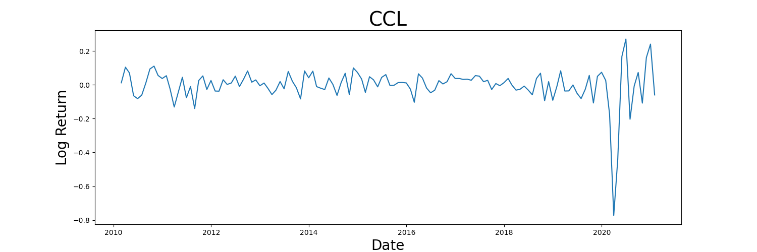
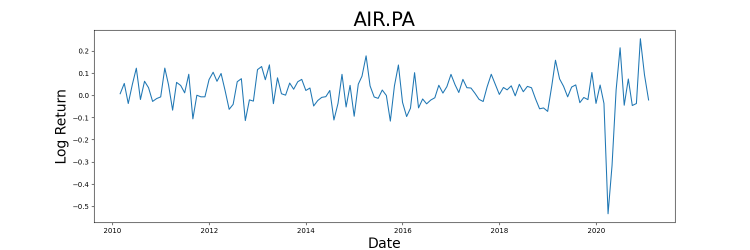
Per quanto riguarda le correlazioni tra i due grafici possiamo chiaramente osservare che essi sembrano essere non correlati tra di loro, mentre per quanto riguarda la volatilità non si riesce a trarre un giudizio accurato.

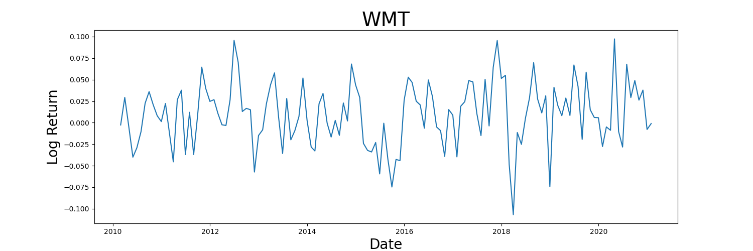
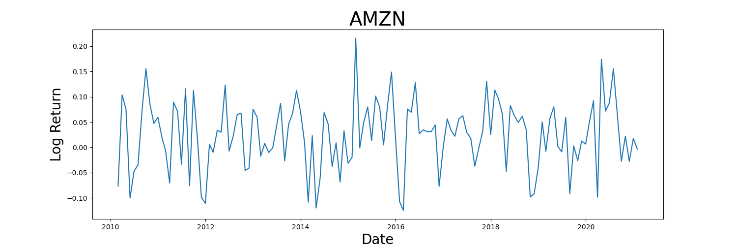
I grafici dei ritorni semplici ci consentono di osservare il trend delle azioni e se esse sono in profitto o meno da inizio periodo, per quanto riguarda la correlazione e la volatilità possiamo solamente fare supposizioni.

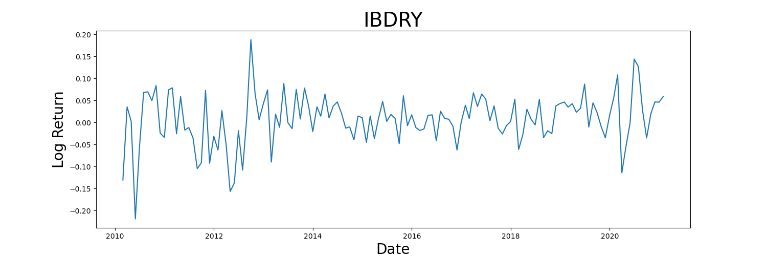
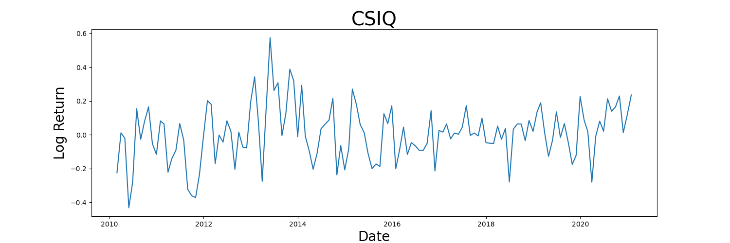
In risposta a questo problema si ricorre infatti ai Log-ritorni, questa tecnica ci permette di confrontare le serie storiche di due società le quali hanno livelli di prezzi completamente diverse, per esempio Amazon e Walmart, entrambe ai massimi storici e con un trend similare, ma la prima quota circa 3000 dollari mentre la seconda circa 140, se “plottassimo” le due serie nello stesso grafico potremmo osservare solamente la serie di Amazon, mentre la serie di Walmart risulterebbe schiacciata in quanto il suo prezzo rispetto alla prima è troppo basso.

Per calcolare i log ritorni abbiamo bisogno del prezzo al tempo t (Pt) e del prezzo al tempo precedente (Pt-1) e calcolare mediante il logaritmo naturale: ln(Pt / Pt-1) oppure ln(Pt) – ln(Pt-1)

Osserviamo ora i Log-ritorni delle sei società, con granularità mensile:

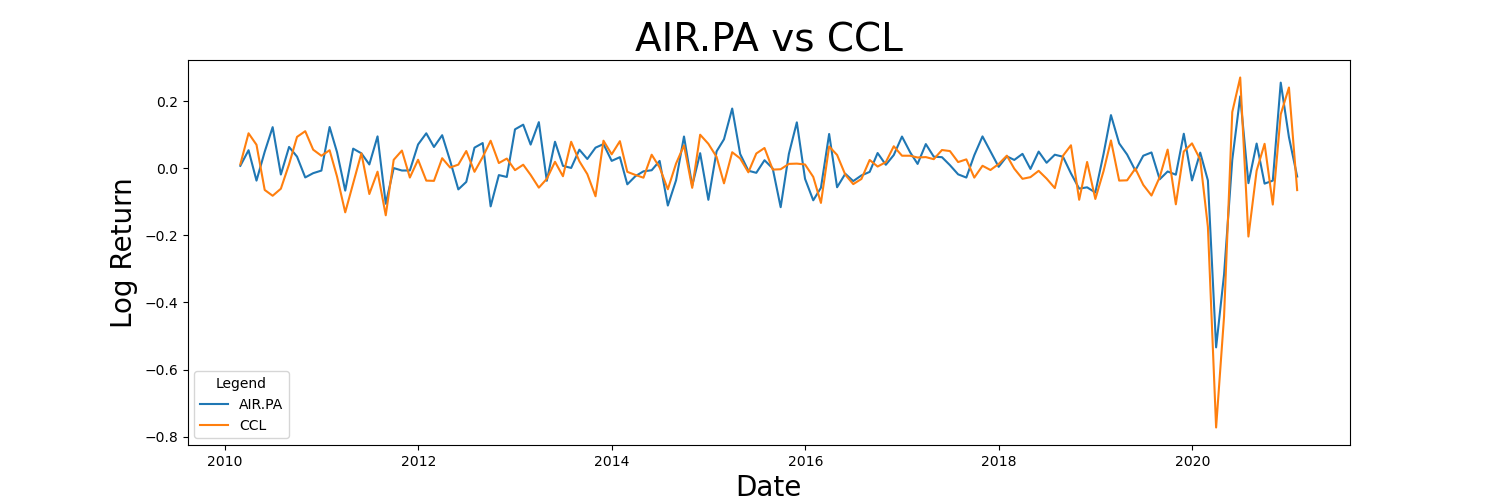






Per trarre giudizi più accurati risulta utile fare il plot dei grafici delle società appartenenti allo stesso settore nel medesimo box:

**Settore trasporti**:

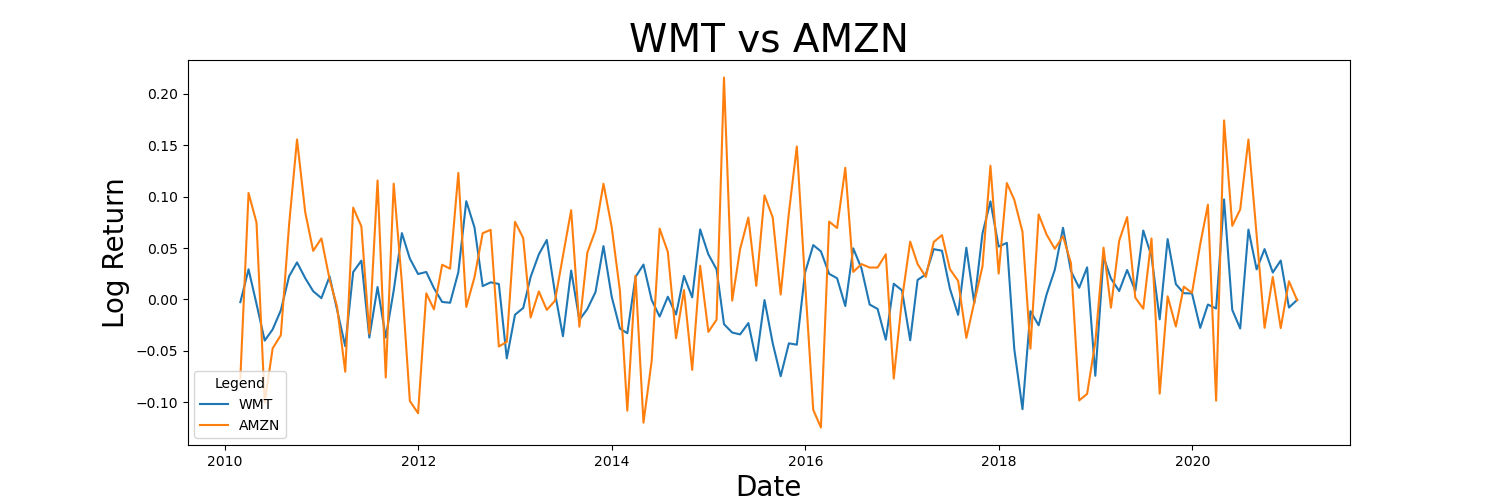


Possiamo affermare ora la presenza di una correlazione positiva tra i due titoli, tendenzialmente i log ritorni si muovono nella medesima direzione tranne in alcuni casi sporadici. Nella finestra temporale che va dal 2018 a inizio 2020, come abbiamo osservato nel grafico dei ritorni semplici, Carnival ha avuto un rendimento negativo mentre Airbus positivo, difatti se osserviamo il grafico dei log ritorni essi risultano essere correlati negativamente, ma col drop dei mercati a causa del Covid-19 la correlazione è tornata positiva.

Infine, possiamo osservare anche come i ritorni negativi di Carnival rispetto ai ritorni negativi di Airbus siano sempre più volatili e ciò ci fa dedurre anche una più alta volatilità da parte della società di crociere.

Tuttavia, possiamo osservare come i due titoli sono sempre stati poco volatili tranne nel 2020 dove il crash dei mercati ha aumentato drasticamente la volatilità di entrambe le società.

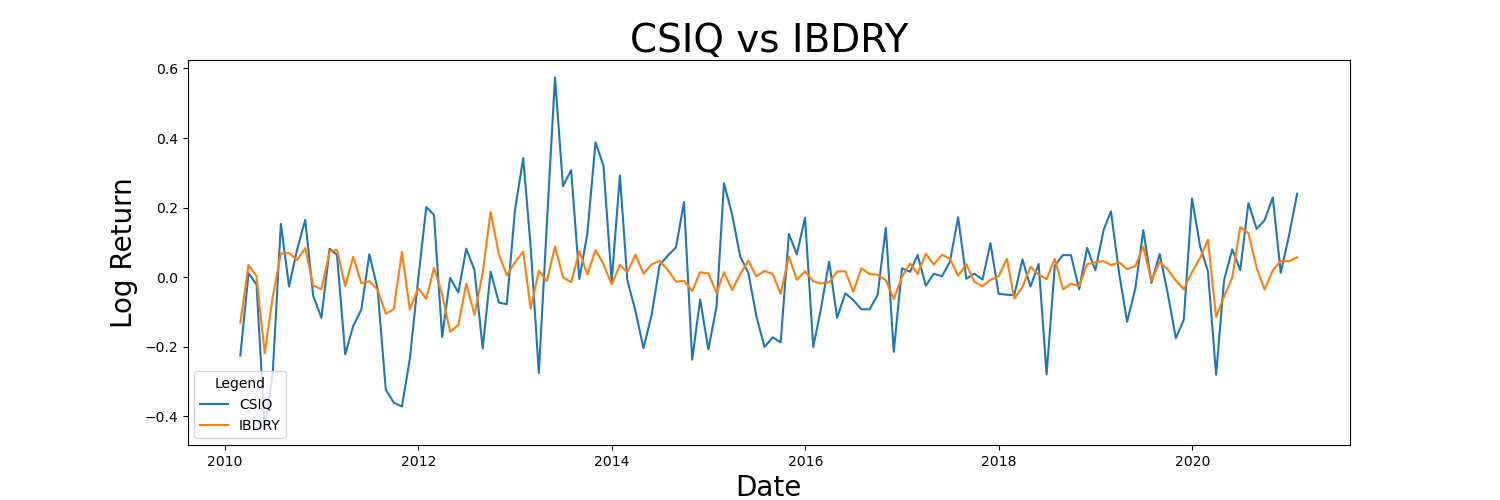
**Settore retail:**



Osservando il grafico dei log ritorni di Amazon e Walmart possiamo subito affermare che i due titoli sembrano essere non correlati tra di loro, nonostante le due società abbiamo avuto rendimenti molto positivi nell’arco dei dieci anni e un andamento similare dei ritorni semplici.

Successivamente possiamo notare la differenza di volatilità, Amazon risulta infatti essere decisamente più volatile rispetto a Walmart sia nei ritorni positivi e sia nei ritorni negativi, questa caratteristica è curiosa in quanto Amazon capitalizza circa quattro volte tanto Walmart ma, nonostante ciò, è soggetta di più a oscillazioni di mercato.

**Settore energetico:**

****

Concludendo col settore energetico, possiamo notare per prima cosa la grossa differenza di volatilità tra i due titoli, questo dettaglio che non si riusciva a cogliere nei grafici dei ritorni semplici ora è chiaramente visibile: Canadian Solar è decisamente più volatile rispetto a Iberdrola.

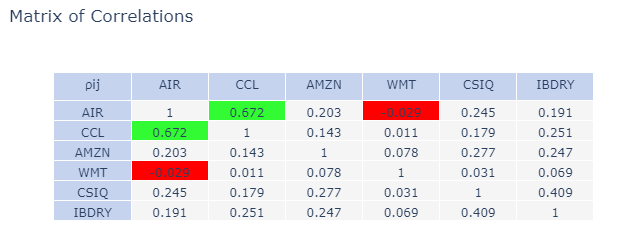
La grossa differenza di volatilità rende però complicato verificare o meno la presenza di correlazione tra i due titoli.

Per risolvere questo problema possiamo calcolare matematicamente la correlazione tra le varie azioni utilizzando la statistica multivariata.

La formula per calcolare la correlazione tra due titoli x e y è: Cov(x,y) / (Std Dev(x)\*Std Dev(y))

Dove Std Dev è la deviazione standard, ovvero la radice quadrata della varianza.

Mentre Cov è la covarianza.



Questa immagine (scaricata direttamente dalla web application) ci mostra in termini numerici ciò che prima abbiamo provato a spiegare a parole osservando i grafici dei log ritorni.

In verde possiamo vedere le coppie di azioni con la correlazione positiva più forte, ovvero Airbus e Carnival Corporation mentre in rosso la coppia con la correlazione meno forte, ovvero Walmart e Carnival Corporation.

Si confermano quindi le società appartenenti al **settore dei trasporti** le più correlate tra di loro, con una correlazione positiva pari a **0.672,** per quanto riguarda gli altri settori possiamo osservare:

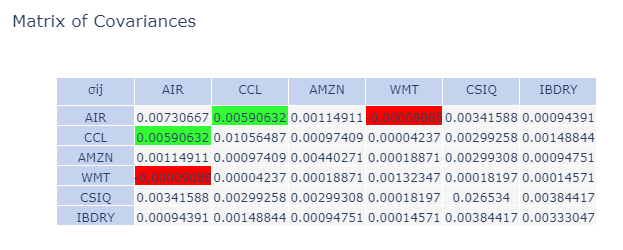
**Settore retail: 0.078**

**Settore energetico: 0.409**

Il settore retail contiene la coppia di titoli meno correlata, a conferma di ciò che abbiamo osservato prima nei grafici dei log ritorni.

Per quanto riguarda Canadian Solar e Iberdrola, possiamo affermare che la correlazione tra i due titoli (che prima non riuscivamo a cogliere a causa della grossa differenza di volatilità) è abbastanza debole.

A conferma delle affermazioni fatte sin ora osserviamo anche la matrice delle covarianze, la quale ci illustra circa la medesima informazione della matrice delle correlazioni:



In parole povere, entrambi i termini misurano la relazione e la dipendenza tra due variabili.

“Covarianza” = misura la direzione della relazione lineare tra le variabili.

“Correlazione” = misura sia la forza che la direzione della relazione lineare tra due variabili.

La correlazione è una funzione della covarianza. Ciò che le distingue è il fatto che i valori di correlazione sono standardizzati, mentre i valori di covarianza non lo sono.

Abbiamo visto quindi come i log ritorni ci tornano utili per:

* Osservare attraverso i grafici la volatilità e la correlazione tra le varie società
* Confermare le osservazioni fatte attraverso il calcolo della matrice delle correlazioni e covarianze.

Tuttavia, nell’ambito della statistica univariata e multivariata i log-ritorni ci servono per effettuare un’analisi ancora più approfondita delle nostre società e poter trarre quindi conclusioni più accurate.

In particolare, la finanza quantitativa e la statistica ci insegnano che più la distribuzione dei log-ritorni di una società risulta essere normalmente distribuita più essa si può ritenere un investimento sicuro (poco rischioso).

Per effettuare questa analisi la programmazione ci viene incontro con varie soluzioni grafiche: istogrammi, qqplot, i boxplot e le distribuzioni di densità.

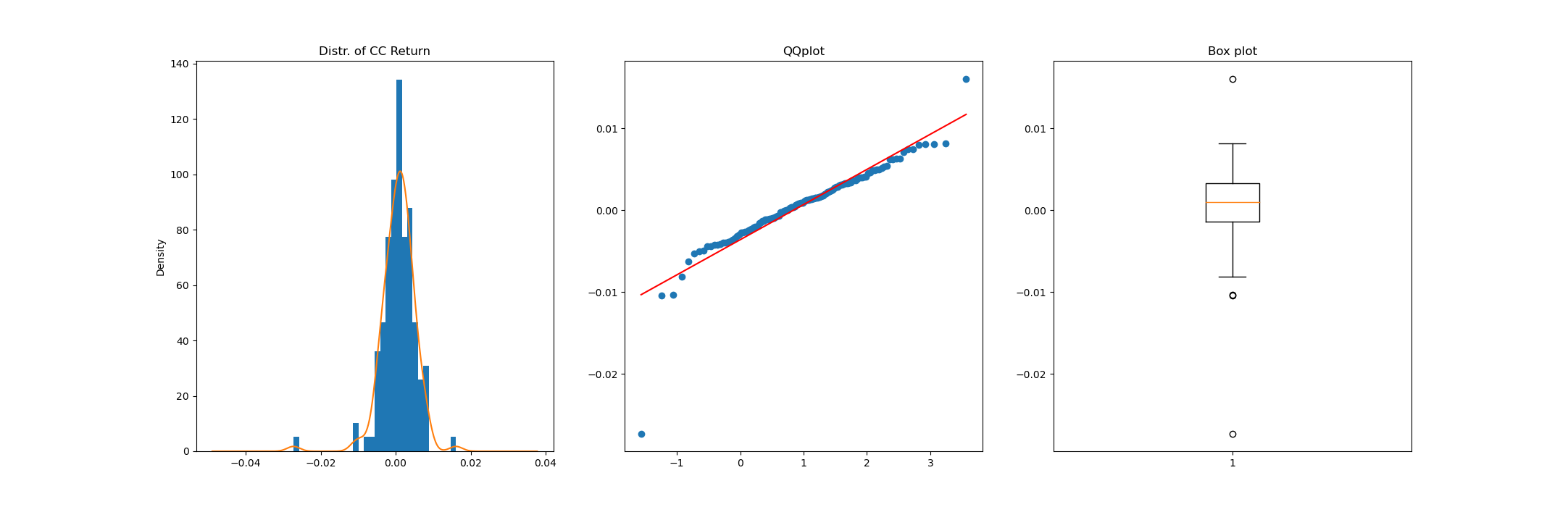
* Istogramma: l’istogramma ci mostra con un grafico a barre la distribuzione dei log-ritorni, ad esso abbino la distribuzione di densità anche sotto forma di linea così da rendere più chiara l’informazione. Da questo grafico riusciamo a cogliere le informazioni di skewness e kurtosis e possiamo osservare se la distribuzione assume la forma simile a una normale.
* QQplot: detto anche quantile-quantile plot, mi consente di fare un confronto tra i quantili teorici della distribuzione (bisettrice, ovvero i quantili della normale) con i quantili empirici (pallini). Possiamo osservare che ai margini (soprattutto a sinistra) del segmento i pallini si scostano tanto, questo dovuto da una kurtosis diversa da quella della normale.
* Boxplot: Il quadrato contiene il 50% della distribuzione (ovvero tra il 25% e il 75%) la linea rossa è la mediana (50%), il minimo e il massimo sono i pallini (detti outliers) i quali si discostano pesantemente dalla distribuzione, le linee orizzontali sono i baffi del boxplot e possono essere visti come l'upperbound e il lowerbound della distribuzione, oltre di essi ci sono gli outliers.

Successivamente possiamo trarre delle conclusioni più accurate attraverso il calcolo di vari indici statistici quali: media, varianza, deviazione standard, quantili, skewness e kurtosis.

* Media: la media dei log-ritorni tende a centrarsi in zero e in finanza tale valore viene interpretato come il rendimento atteso della società.
* Varianza e Deviazione standard: entrambe ci forniscono una misura della volatilità del titolo che stiamo analizzando, più in particolare la deviazione standard viene usata come indice di rischiosità di un titolo, quindi più essa è alta -> più il titolo è volatile -> più è rischioso.
* Quantili: ci consentono di osservare la distribuzione di densità negli intervalli: 0-0.25, 0.25-0.5, 0.5-0.75, 0.75-1
* Skewness: misura la simmetria della distribuzione attorno alla media, più il valore fornito è vicino a zero e più la distribuzione è simile alla normale. Se il valore fornito è maggiore di zero allora la distribuzione avrà la coda di destra più lunga, se minore di zero allora la distribuzione avrà la coda di sinistra più lunga. Tendenzialmente la skewness calcolata sui ritorni azionari ha sempre valore negativo.
* Kurtosis: la curtosi misura lo spessore delle code della nostra distribuzione, anche in questo caso più il valore fornito è vicino a zero e più la distribuzione risulterà normalmente distribuita, se il valore fornito è maggiore di zero significa che le code sono più spesse rispetto alla normale, se minore di zero viceversa.

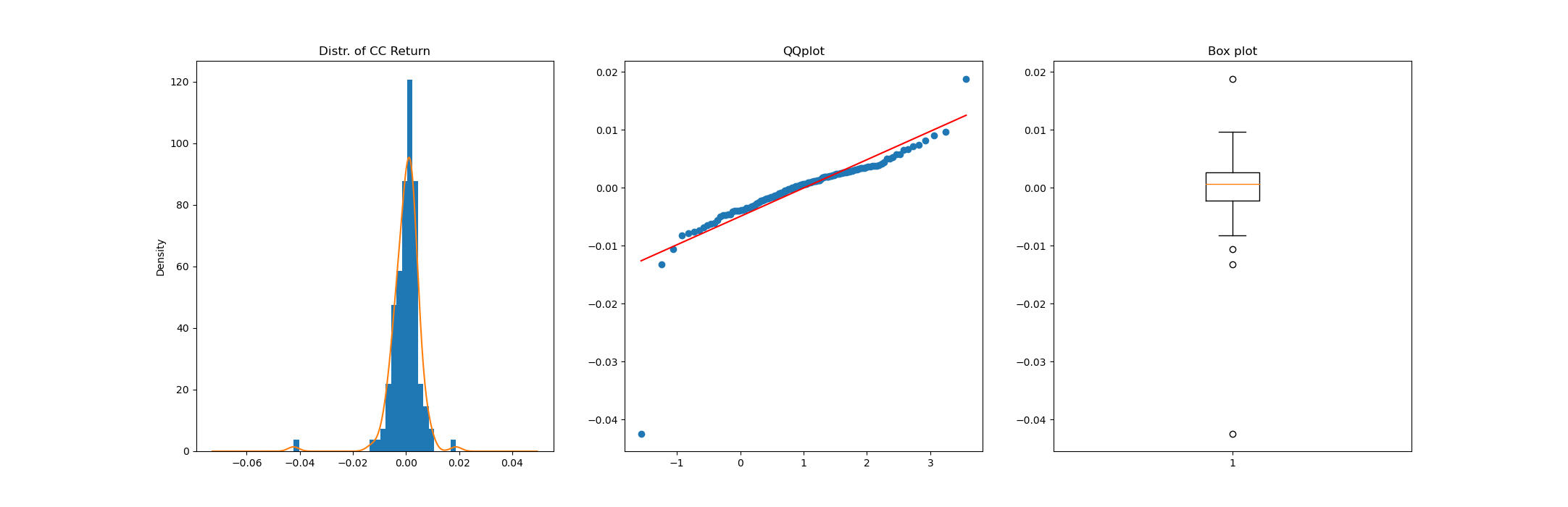
Osserviamo ora per ogni società i grafici e gli indici statistici che abbiamo appena menzionato:

AIRBUS



|  |  |
| --- | --- |
| Mean: 0.01483  Variance: 0.007307  Std Dev: 0.085479  Skewness: -2.132129  Kurtosis: 13.581299 | Min: -0.533801  25%: -0.02597  Median: 0.021929  75%: 0.055321  Max: 0.255094 |

CARNIVAL CORPORATION



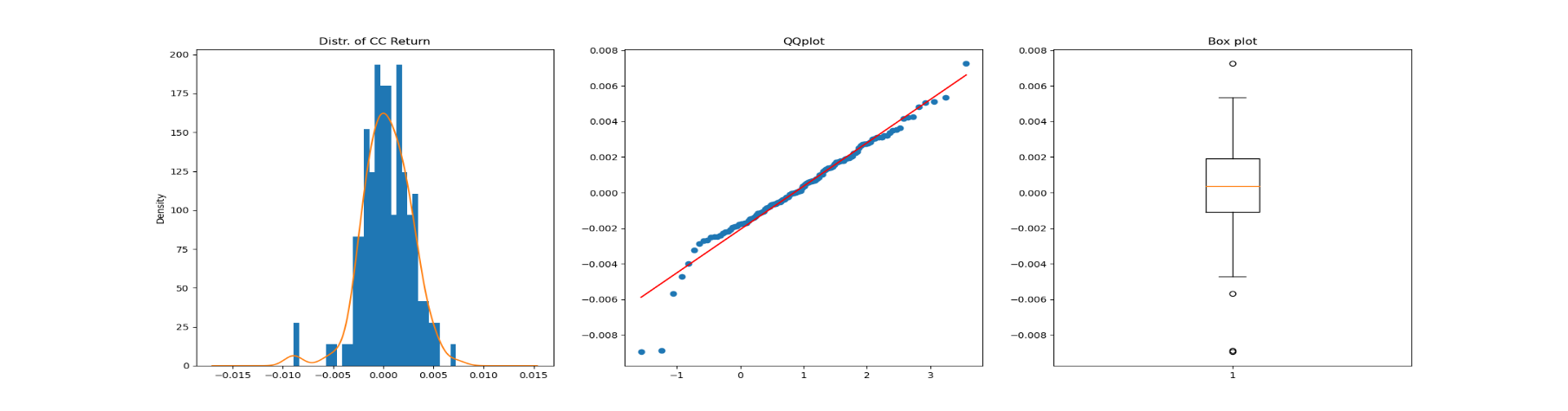
|  |  |
| --- | --- |
| Mean: -0.002138 Variance: 0.010557 Std Dev: 0.102748 Skewness: -3.637397 Kurtosis: 25.766301 | Min: -0.77258  25%: -0.031528  Median: 0.011207  75%: 0.041515  Max: 0.27041 |

Confrontando le due società possiamo notare ciò che avevamo notato osservando i grafici dei ritorni semplici, il ritorno atteso (mean) di Carnival è negativo, inoltre risulta essere anche più rischiosa rispetto ad airbus in quanto presenta una deviazione standard più elevata e, per quanto riguarda skewness e kurtosis possiamo osservare che Airbus risulta essere più normale di Carnival.

Infine, a sostegno di quanto detto sin ora Airbus presenta solo 3 outliers (uno positivo e due negativi) mentre Carnival ne presenta 4 (uno positivo e tre negativi).

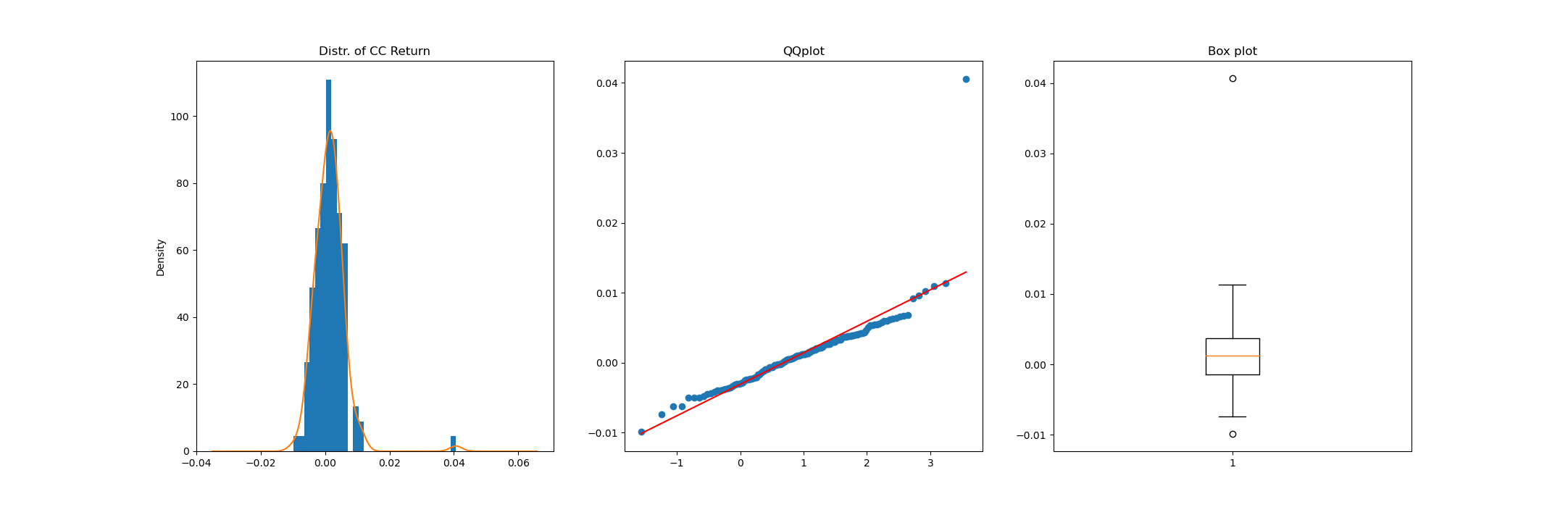
Possiamo concludere quindi che Airbus è sotto tutti i punti di vista un investimento migliore rispetto a Carnival Corporation.

WALMART



|  |  |
| --- | --- |
| Mean: 0.009159  Variance: 0.001328  Std Dev: 0.036437  Skewness: -0.155519  Kurtosis: 0.226673 | Min: -0.106858  25%: -0.011418  Median: 0.008822  75%: 0.029319  Max: 0.09737 |

AMAZON



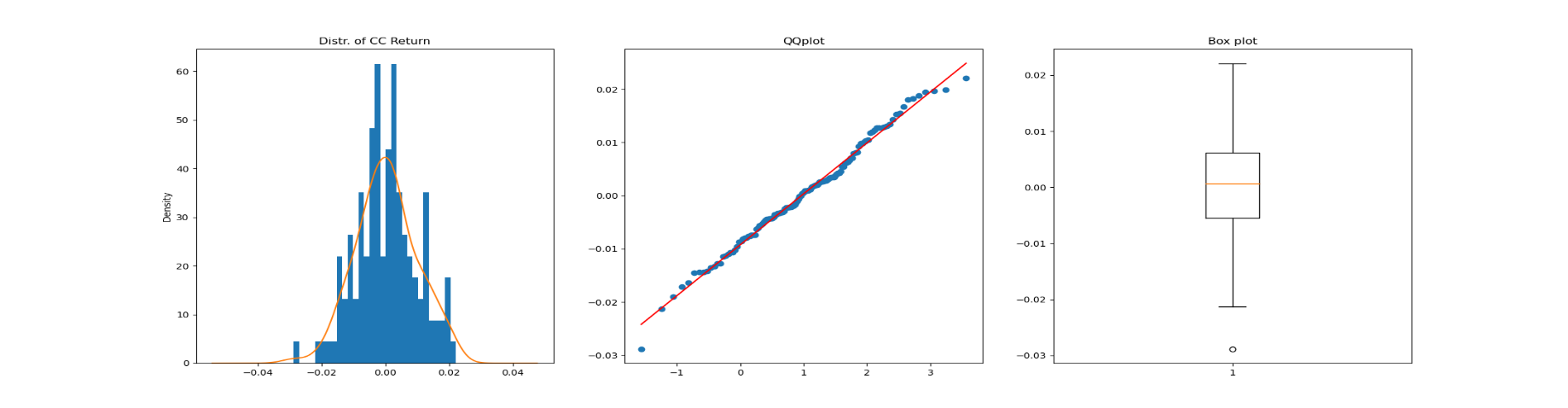
|  |  |
| --- | --- |
| Mean: 0.024564  Variance: 0.004405  Std Dev: 0.066368  Skewness: -0.134116  Kurtosis: -0.002587 | Min: -0.124651  25%: -0.010159  Median: 0.029924  75%: 0.068931  Max: 0.216012 |

Possiamo ora confermare al 100% che Amazon è più volatile (quindi più rischiosa) rispetto a Walmart, tuttavia tutti gli altri indici sono a favore della società di e-commerce, infatti possiamo osservare come il ritorno atteso di Walmart sia inferiore, entrambe le società presentano skewness negativa ma la più prossima a zero appartiene ad Amazon, anche per quanto riguarda la kurtosis possiamo notare come il valore di Amazon risulta essere decisamente prossimo a zero.

Infine, Walmart presenta tre outliers (uno positivo e due negativi) mentre Amazon ne presenta solo due (uno molto positivo e uno leggermente negativo).

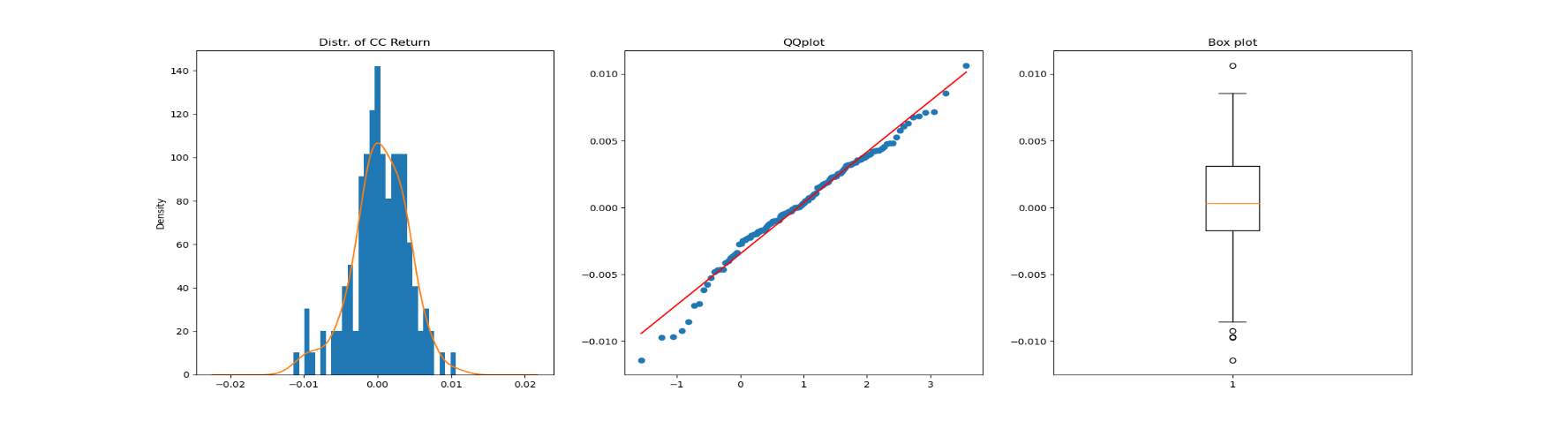
Possiamo quindi constatare che Amazon risulta essere un investimento migliore rispetto a Walmart.

CANADIAN SOLAR



|  |  |
| --- | --- |
| Mean: 0.005414  Variance: 0.02652  Std Dev: 0.162851  Skewness: 0.100631  Kurtosis: 0.759301 | Min: -0.431773  25%: -0.090855  Median: 0.010261  75%: 0.087221  Max: 0.573944 |

IBERDROLA



|  |  |
| --- | --- |
| Mean: 0.006718  Variance: 0.003331  Std Dev: 0.057713  Skewness: -0.57765  Kurtosis: 2.107748 | Min: -0.218969  25%: -0.020652  Median: 0.009359  75%: 0.042958  Max: 0.187926 |

Infine, analizzando le due società energetiche, di cui eravamo riusciti a cogliere già abbastanza informazioni, possiamo notare come il ritorno atteso di Iberdrola sia più alto rispetto a Canadian Solar, risulta essere anche meno rischiosa (come avevamo già dedotto) tuttavia la distribuzione di Canadian Solar è decisamente più normale rispetto a Iberdrola, anche per quanto riguarda la presenza di outliers, Canadian Solar ne possiede solo uno negativo mentre Iberdrola ne possiede quattro: uno positivo e tre negativi.

In questo caso trarre un giudizio su quale delle due possa essere l’investimento migliore risulta essere più complicato. Dal punto di vista di “vicinanza alla normale” dovremmo scegliere la prima, tuttavia, la seconda risulta essere statisticamente meno rischiosa e più remunerativa.

Personalmente le riterrei entrambe un buon investimento.

Osserviamo ora un riassunto di quanto abbiamo appena analizzato.

Nota bene: negli indici più bassi skewness e kurtosis ho inserito le società che presentano la maggior vicinanza a zero in valore assoluto.

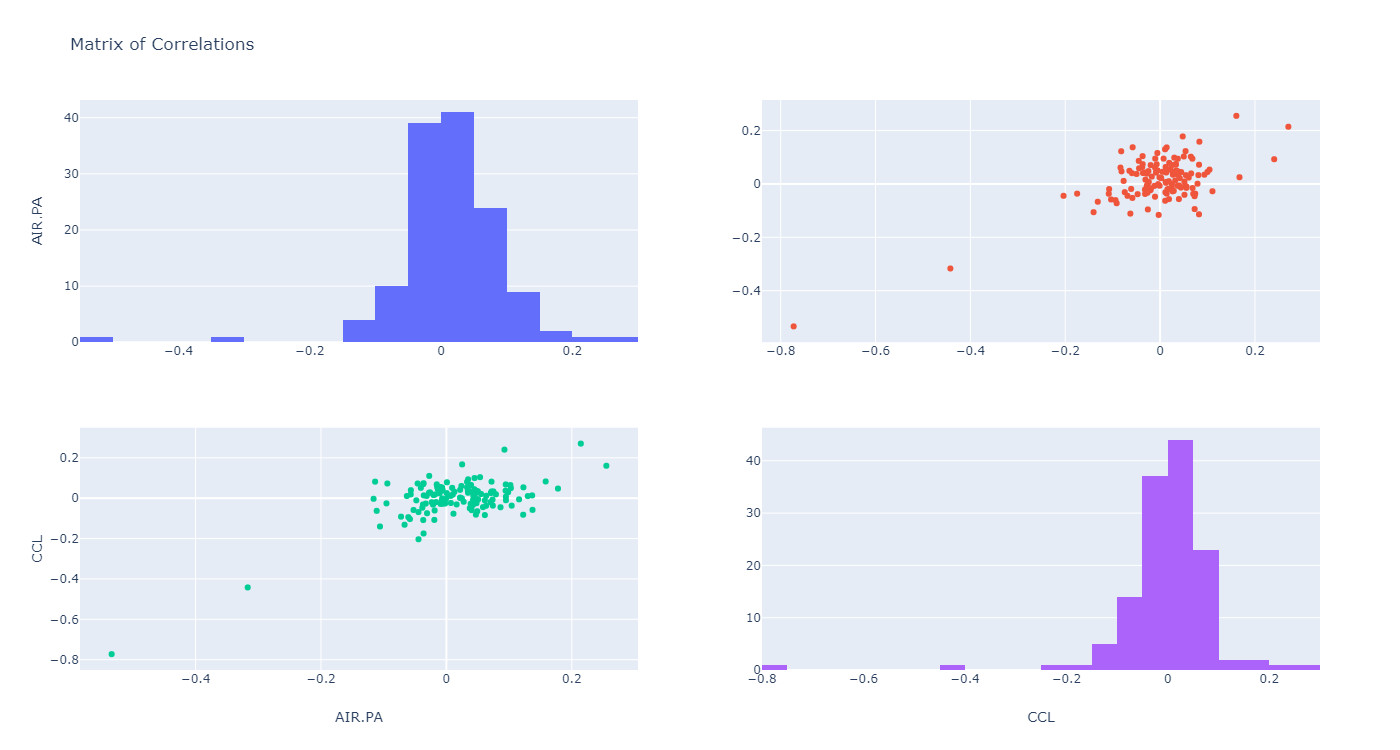
|  |  |
| --- | --- |
| **Indici più alti**  Mean: Amazon  Std Dev: Canadian Solar  Skewness: Carnival Corporation  Kurtosis: Carnival Corporation | **Indici più bassi**  Mean: Carnival Corporation  Std Dev: Walmart  Skewness: Canadian Solar  Kurtosis: Amazon |

Concludiamo così l’analisi dal punto di vista della statistica univariata.

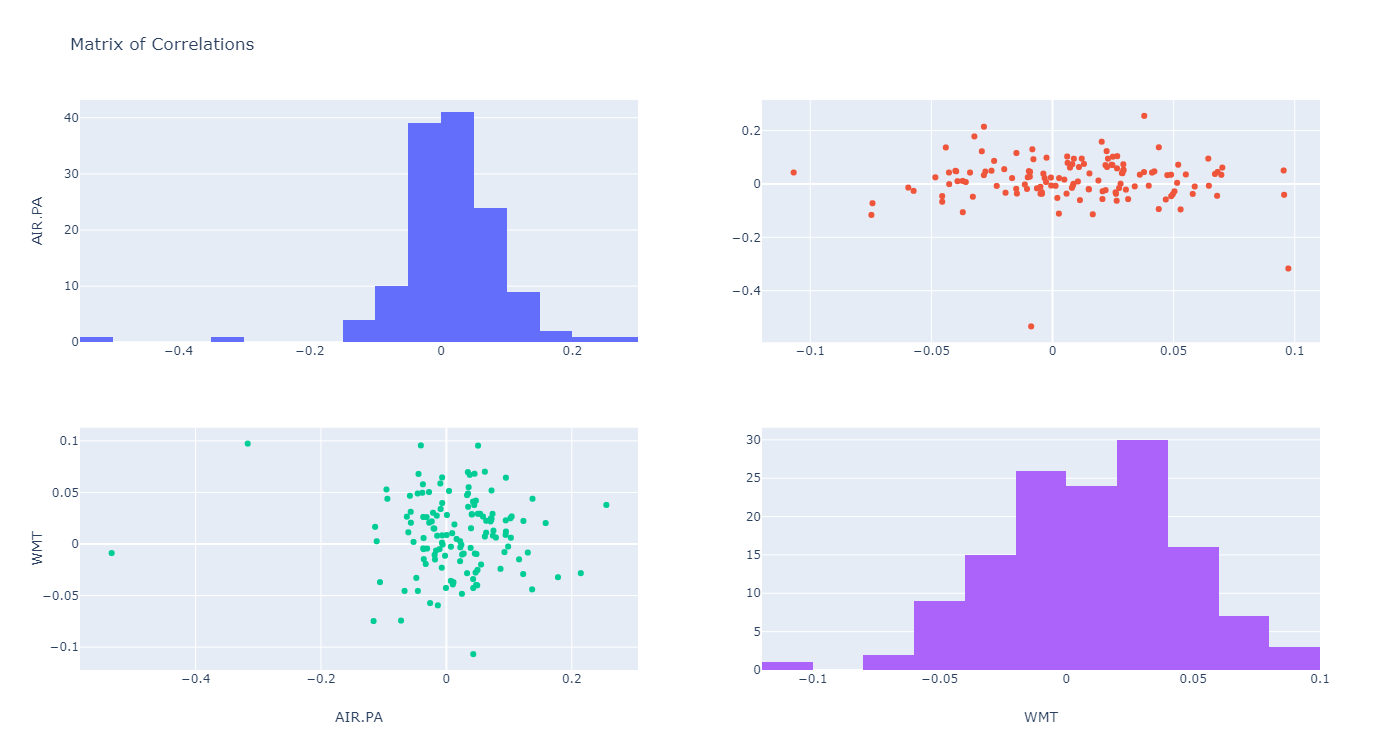
Ritornando invece all’analisi multivariata, quindi correlazioni e covarianze affrontate precedentemente, possiamo osservare ora un’altra modalità grafica per visualizzare la correlazione che intercorre tra le società.

Questa modalità prevede di usare gli scatter plot, ovvero i grafici a pallini, più precisamente maggiore è la correlazione tra due società e più i essi assumeranno una distribuzione a segmento diagonale; più le società non sono correlate e maggiore sarà la dispersione dei pallini all’interno del grafico.

Osserviamo ora gli scatter plot della coppia di società più correlata: Airbus e Carnival Corporation



Osserviamo ora gli scatter plot della coppia di società meno correlata: Airbus e Walmart



Possiamo notare che nella seconda immagine i pallini sono distribuiti in maniera molto più dispersiva rispetto alla prima.

**INDICE BETA**

Come ultimo argomento di “Descriptive Analysis” parleremo dell’indice **Beta,** questo indice viene calcolato grazie ai log ritorni della società che stiamo analizzando e grazie ai log ritorni dell’indice di mercato di riferimento, in questo caso un buon indice risulta esse lo Standard & Poor 500: realizzato da Standard & Poor's nel 1957, segue l'andamento di un paniere azionario formato dalle 500 aziende statunitensi a maggiore capitalizzazione.

Dal punto di vista statistico l’indice beta corrisponde al coefficiente angolare della retta che interpola linearmente i ritorni della società che stiamo analizzando in funzione dei ritorni dell’indice di riferimento.

Esso, se calcolato tra due titoli x e y (dove x risulta essere l’indice di mercato), risulta essere:

Covarianza(x,y) / Varianza(x)

L’informazione che ci fornisce è la misura della rischiosità sistematica dell'azione: esso misura la variazione attesa del rendimento del titolo per ogni variazione di un punto percentuale del rendimento di mercato. Il rendimento atteso di un titolo varia linearmente con il beta del titolo stesso.

Azioni con un beta superiore a 1 tendono ad amplificare i movimenti di mercato (l'attività è più rischiosa del mercato): in via generale si ritiene che le società con politiche imprenditoriali aggressive o con elevati livelli di indebitamento presentino i valori di beta più elevati.

Al contrario, azioni con beta compresi tra 0 e 1 tendono a muoversi nella stessa direzione del mercato (l'attività è meno rischiosa del mercato): si tratta generalmente di titoli emessi da società che operano nei settori tradizionali dell'economia.

Tale indice viene calcolato utilizzando un range passato di valori, questo range deve corrispondere alla nostra prospettiva di investimento. Per esempio: se abbiamo intenzione di tenere in portafoglio una determinata azione per un periodo breve (due anni) non andremo a calcolare il beta utilizzando i dati degli ultimi dieci anni ma lo andremo a calcolare sugli ultimi due.

Utilizzando l’applicazione web sviluppata osserviamo ora il beta delle nostre società calcolato nell’intero periodo che inizia il 1° gennaio 2020 e termina il 31 gennaio 2021:

Airbus: 1.517

Carnival Corporation: 1.745

Walmart: 0.337

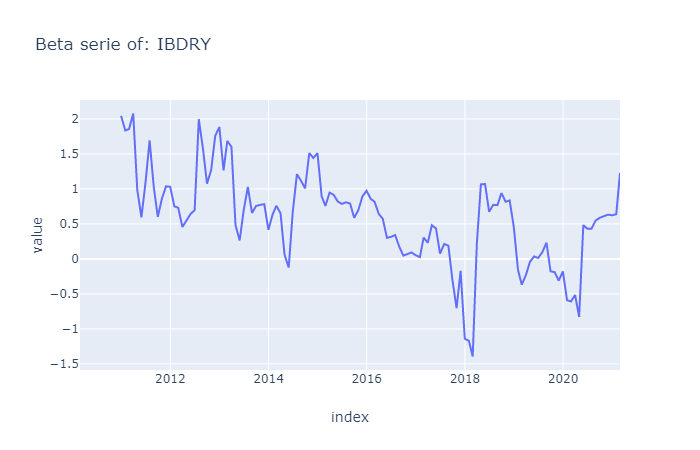
Amazon: 0.942

Canadian Solar: 1.876

Iberdrola: 0.726

Possiamo notare come ci sia una correlazione tra il Beta e la Volatilità delle società: Canadian Solar risultava essere la società più rischiosa, presenta infatti il Beta più elevato; al contrario Walmart risultava essere la società meno rischiosa e presenta infatti il beta più basso.

Questo indice possiamo osservarlo anche a scopo puramente grafico, selezionando un delta temporale su cui calcolarlo e plottarlo successivamente in un grafico a linea. Osserviamo ora come esempio la serie temporale del Beta di Canadian Solar utilizzando come delta 10 mesi:



**Predictive Analysis**

In questa sezione ho applicato una delle tecniche di machine learning apprese durante il corso, in particolare l’analisi previsionale è stata effettuata grazie all’algoritmo di Support Vector Machine.

Descrizione: Il Support Vector Machine (SVM) è un algoritmo di apprendimento automatico supervisionato che può essere utilizzato sia per scopi di classificazione che di regressione. È popolare in applicazioni quali l’elaborazione del linguaggio naturale, il riconoscimento vocale e delle immagini e la computer vision.

Nel nostro caso il suo scopo sarà quello di effettuare una regressione.

L’analisi previsionale è stata effettuata sugli ultimi 10 mesi del nostro periodo temporale di circa 10 anni. Potrebbe risultare insensato ma dal punto di vista didattico ci consente di osservare quanto la nostra previsione si è discostata dai dati reali.

Trattandosi di un algoritmo di apprendimento supervisionato ha bisogno di essere allenato con un dataset di dati presi in maniera randomica dall’intero dataset, nel nostro caso la dimensione del train set è circa il 70% dell’intera popolazione (80 mesi) mentre il restante 30% (30 mesi) verrà utilizzato come test set.

Nota bene: per intero dataset si intende tutta la serie storica dei log ritorni mensili tranne gli ultimi dieci mesi che andremo a predire successivamente.

In termini di dati quindi come test e train set sono stati presi i ritorni che vanno da gennaio 2010 fino ad aprile 2020, mentre i dati da predire incominciano da maggio 2020 fino a febbraio 2020.

L’algoritmo di SVM all’interno del linguaggio python assume questa connotazione:

* svm = SVR(kernel='rbf', **C**=Cs[a], **gamma**=gammas[b], **epsilon**=epsilons[c])

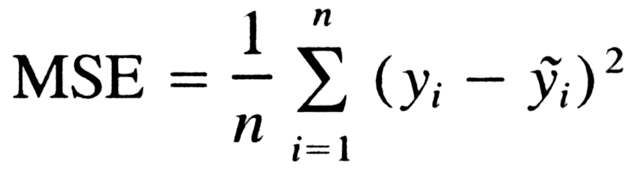
All’interno della parentesi possiamo trovare i parametri che servono al Support Vector Regressor per effettuare la sua regressione.

Per trovare i migliori parametri effettuiamo una ricerca detta “grid search” tale ricerca si basa sul concetto di effettuare il maggior numero di combinazioni possibili tra i vari parametri e selezionare alla fine il settaggio che ha fornito la misura di errore meno elevata.

In questo caso come misura di errore prendiamo in considerazione l’MSE ovvero “Mean square error” tradotto “errore quadratico medio”

In statistica indica la discrepanza quadratica media fra i valori dei dati osservati ed i valori dei dati stimati.

La sua formula è:

dove yi sono i dati osservati e y segnato i sono i dati stimati.

Trovata quindi la miglior combinazione di parametri si effettua la regressione per predire i futuri dieci mesi (che nel nostro caso già conosciamo) e successivamente unendo le due serie dei log ritorni, quella dei dati osservati + quella dei dati previsti possiamo plottarla in un grafico. Dalla web application sviluppata i migliori parametri risultano essere:

Airbus: [0.006294882198578257, **1000.0, 1.0, 0.1**]

Carnival Corporation: [0.008515408898791295, **10.0, 10.0, 0.0001**]

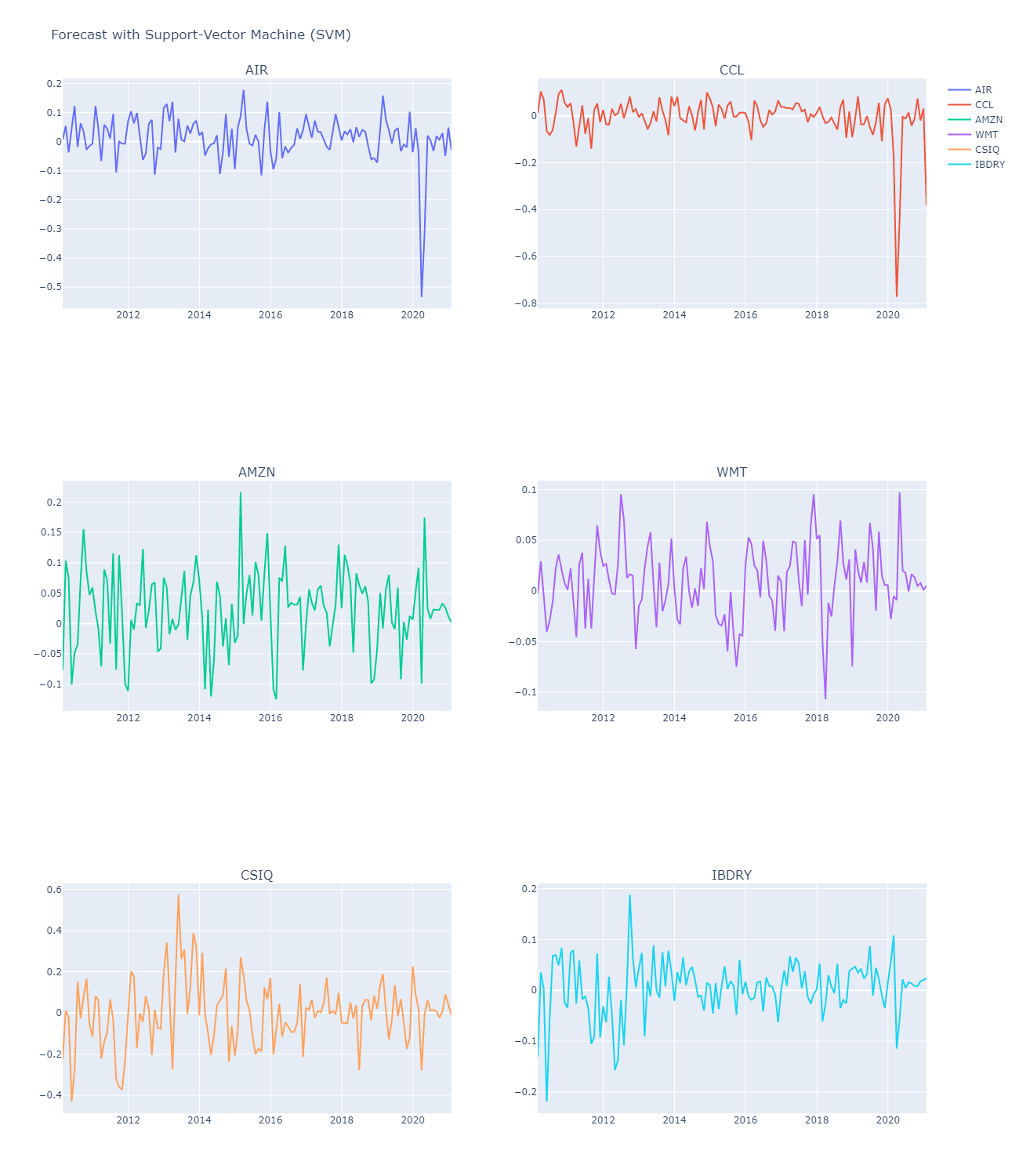
Amazon: [0.004434164915216299, **1.0, 1000.0, 0.1**]

Walmart: [0.0012456849439376584, **1000.0, 0.001, 0.01**]

Canadian Solar: [0.02502453947847072, **100.0, 0.001, 0.0001**]

Iberdrola: [0.0031045858090079544, **10.0, 0.01, 0.01**]

In rosso possiamo osservare invece l’MSE



La ricerca non è finita qua, in quanto i valori che abbiamo previsto ci serviranno ora (assieme ai dati passati) per l’ultimo step della nostra relazione ovvero la composizione del portafoglio ideale:

**Portfolio Management**

Per la composizione del portafoglio ideale utilizzeremo il modello media-varianza inventato da Markowitz.

Come abbiamo visto la media dei log ritorni della nostra società corrisponde al ritorno atteso che ci aspettiamo dal nostro investimento, tuttavia, calcolare la media utilizzando solo i dati passati potrebbe risultare poco efficiente, grazie alle tecniche di forecasting possediamo ora anche i valori futuri che abbiamo previsto e possiamo utilizzarli assieme ai valori passati per calcolare un nuovo ritorno atteso da utilizzare nel modello media-varianza.

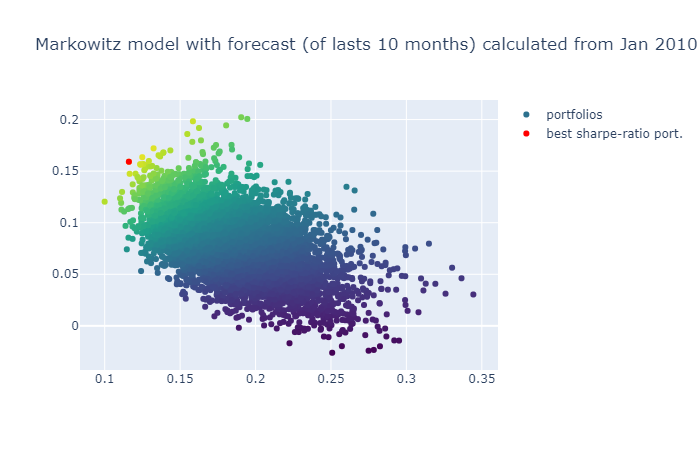
Lo stesso discorso lo applichiamo ovviamente anche per il calcolo della varianza, e quindi della volatilità/rischio della società.

L’idea di Markowitz è quella di trovare i pesi (ovvero in percentuale quanti soldi investire) da assegnare ad ogni società che compone il nostro portafoglio con la regola che il ritorno atteso del portafoglio deve essere massimo mentre la varianza deve essere minima.

Per trovare la miglior combinazione di pesi si utilizza un indice chiamato Sharpe Ratio, questo indice è calcolato nel seguente modo:

S.R. = ( Ritorno atteso portafoglio – Tasso free-risk ) / Deviazione standard portafoglio

Il modello che ho sviluppato calcola 6000 portafogli differenti, e per ogni portafoglio alloca in maniera randomica pesi differenti su ogni società (n.b. la somma dei pesi equivale sempre a 1 = 100%) e alla fine seleziona la combinazione che ha restituito lo Sharpe Ratio maggiore:

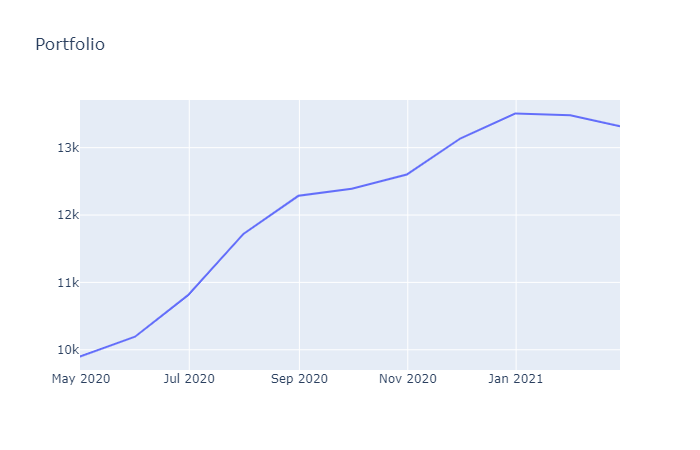


|  |  |
| --- | --- |
| Miglior portafoglio:  Ritorno atteso: 15.909%  Volatilità attesa 11.605% | AIR: 10.614%  CCL: 3.2837%  AMZN: 33.8613%  WMT: 48.4099%  CSIQ: 1.008%  IBDRY: 2.823% |

Come tasso free-risk ho considerato lo 0.1% ovvero il tasso di rendimento dei bond governativi statunitensi a breve scadenza.

Ipotizzando quindi un investimento di 10.000 euro nel portafoglio sopra calcolato, iniziato a maggio 2020 e terminato a febbraio 2021, il ritorno atteso sarebbe stato di circa 16%

Nella realtà però il ritorno effettivo sarebbe stato del 34,51% col seguente andamento:



Per un’analisi più accurata ho considerato una commissione d’ingresso pari all’1% dell’investimento e anche la tassazione italiana pari al 26% sul capital gain (ovvero il guadagno netto del nostro investimento):

Ritorno lordo: 13316.89€

Ritorno netto: 3416.89€

Ritorno netto (tassato al 26%): 2528.5€

**Conclusione**

In conclusione, possiamo quindi osservare come il modello media-varianza unito al forecast con il Support Vector Machine abbia sottostimato la performance reale, questo fenomeno possiamo spiegarlo in quanto l’arco temporale degli ultimi 10 mesi presi in considerazione corrisponde esattamente con la ripresa che i mercati hanno subito dopo il crash causato dal Covid-19.

Il nostro modello, tuttavia, non è a conoscenza di fenomeni che accadono nella realtà quali: una pandemia, la successiva ripresa dei mercati, l’influenza che il lockdown e lo stato di emergenza sanitaria hanno avuto sui i business delle aziende selezionate.

Il SVR si limita ad effettuare una regressione utilizzando i dati passati e di conseguenza ci possiamo aspettare una previsione futura tendenzialmente simile a come la nostra società ha performato in passato. Per osservare meglio quanto appena detto ho creato dei grafici che illustrassero la differenza tra i valori predetti dal nostro modello e i valori reali in termini di ritorni semplici:

