

# BUSINESS INTELLIGENCE PER I SERVIZI FINANZIARI

## PROGETTO DI LABORATORIO IN PYTHON

Alessio Martini

## INDICE

1. **Sommario dei dati utilizzati**
  - 1.1 Titoli analizzati e motivazione
  - 1.2 Funzioni download & Fusione/allineamento
  - 1.4 Presentazione dati
2. **Statistiche descrittive**
  - 2.1 Rendimento cumulato e CAGR per titolo
  - 2.2 Rendimenti semplici e logaritmici
  - 2.3 Statistiche annualizzate
  - 2.4 Varianze-covarianze e matrice di correlazione
3. **Analisi di previsione (ARIMA)**
  - 3.1 Split 80/30/10 e workflow
  - 3.2 ARIMA(1,0,1) vs auto-ARIMA
  - 3.3 Errori RMSE/MAE su test e forecast con ordine selezionato
  - 3.4 Commento: bassa prevedibilità
4. **Strategie di trading e backtesting - Analisi tecnica**
  - 4.1 Mean Reversion 5g (long/short)
  - 4.2 Mean Reversion 5g (long-only)
  - 4.3 SMA 50/200 (long-only)
  - 4.4 Random Forest (long-only)
  - 4.5 Confronto con Buy&Hold
5. **CAPM e Fattori Fama-French (3F)**
  - 5.1 Beta per ciascun titolo
  - 5.2 Fama-French 3F
  - 5.3 Confronto  $\beta$ (CAPM) vs  $\beta_{\_MKT}$ (FF)
  - 5.4 Esposizioni ai fattori
7. **Costruzione di portafoglio (media-varianza)**
  - 7.1 Setup: 108 mesi training; attese storiche vs CAPM
  - 7.2 Metodo analitico & Monte Carlo: frontiere efficienti
  - 7.3 Pesi dei portafogli ottimali & Equal Weight
  - 7.4 Beta di portafoglio
  - 7.5 Confronto OOS con Equal Weight e S&P500
  - 7.6 Metriche OOS
  - 7.7 Differenze tra attese storiche e CAPM
8. **Discussione dei risultati**
  - 8.1 Sintesi dei risultati su previsione, strategie e portafogli; implicazioni operative.

## **SOMMARIO DEI DATI UTILIZZATI**

### **Descrizione e motivazione della scelta dei titoli:**

Sono stati presi in esame otto titoli appartenenti all'indice S&P 500, rappresentativi di quattro settori differenti, oltre all'indice S&P 500 preso come benchmark. La scelta dei titoli è basata su diversificazione in settori, rilevanza economico-finanziaria e l'interesse di osservare come ciascun titolo abbia reagito agli eventi macroeconomici e geopolitici avvenuti negli ultimi dieci anni (2014–2024).

- Visa (V) e Mastercard (MA) - Pagamenti digitali:**

Il settore dei circuiti di pagamento elettronico è un oligopolio globale in mano a pochi grandi operatori; si tratta fondamentalmente di Visa e Mastercard che si spartiscono oltre il 75% delle transazioni interbancarie che avvengono al di fuori della Cina. Il settore ha barriere all'ingresso molto importanti dovute a economie di scala e effetti rete (un circuito ha più utilità quanto più ha esercenti e carte distribuiti), infrastrutture sicure/regolate e marchi noti e affidabili. Il settore è guidato da un forte trend strutturale di digitalizzazione e si è dimostrato abbastanza anticiclico, andando a riprendersi rapidamente dopo la breve interruzione nel corso del COVID dei volumi di pagamento.

- Johnson & Johnson (JNJ) e Pfizer (PFE) – Healthcare:**

Johnson & Johnson e Pfizer sono due grandi gruppi globali nel settore sanitario che operano in aree complementari:

- JNJ è un gruppo diversificato che opera nell'area farmaceutica, dei dispositivi medici e consumer; si è dimostrata molto solida e in grado di far crescere in modo costante i ricavi nel lungo periodo e con una volatilità inferiore a quella del mercato.
- Pfizer è un'azienda di ricerca e sviluppo in ambito farmaceutico più ciclica: nel corso della pandemia da COVID-19 ha avuto un forte impulso grazie al vaccino Comirnaty, ma ha poi visto il ritorno alla normalità con la fine della pandemia e pertanto il fatturato è sceso.

Il settore dell'Healthcare è un settore difensivo, in quanto di solito tende a mantenere la domanda anche in fasi di rallentamento economico, e consente di osservare comportamenti diversi in situazioni di crisi o in circostanze di emergenza sanitaria (si pensi a Pfizer nel COVID, e a J&J nei periodi di mercato ribassista).

- Lockheed Martin (LMT) e Caterpillar (CAT) - Industrials**

Lockheed Martin e Caterpillar sono due titoli industriali con modelli di business profondamente differenti, entrambi caratterizzati da una forte ciclicità: nel caso di Caterpillar legata principalmente all'andamento del ciclo economico, mentre per Lockheed Martin influenzata in misura rilevante dalla dinamica geopolitica.

- Lockheed Martin è il primo contractor del Dipartimento di Difesa USA, con ricavi molto stabili e garantiti da lunghi contratti di fornitura governativi. Il titolo tende a performare in caso di escalation delle tensioni geopolitiche (es. guerre o rialzi della spesa militare dei Paesi NATO).

- Caterpillar è leader mondiale delle macchine per construction e mining, molto sensibile al ciclo economico e alla domanda di infrastrutture e commodity. Performante in caso di forte crescita economica e di stimolo agli investimenti pubblici/privati.

- **Meta (META) e Amazon (AMZN) - Tech**

Meta e Amazon sono due tra i leader tecnologici mondiali con crescita a due cifre di lungo periodo, ma modelli di business differenti:

- Meta è uno dei maggiori colossi pubblicitari mondiali, proprietaria di Facebook, Instagram e WhatsApp. I suoi ricavi sono altamente ciclici ma crescono strutturalmente grazie alla digitalizzazione e alla crescita della spesa pubblicitaria online.
- Amazon è leader mondiale dell'e-commerce e tra i maggiori player cloud (AWS); a sua volta combina una componente ciclica (consumi online) e una componente con minore volatilità e più redditizia (cloud) che ha trainato gran parte della crescita negli ultimi anni.

Entrambe sono dotate di un elevato potenziale di crescita ma espongono il portafoglio a maggiore volatilità e ci consentono di analizzare come i titoli growth reagiscono in contesti di rallentamento del ciclo macroeconomico.

## **Funzioni utilizzate per download dei dati e fusione di serie:**

### **Download:**

- yfinance.download
- pandas\_datareader.data.DataReader (sorgente "famafrance")

### **Fusione/allineamento:**

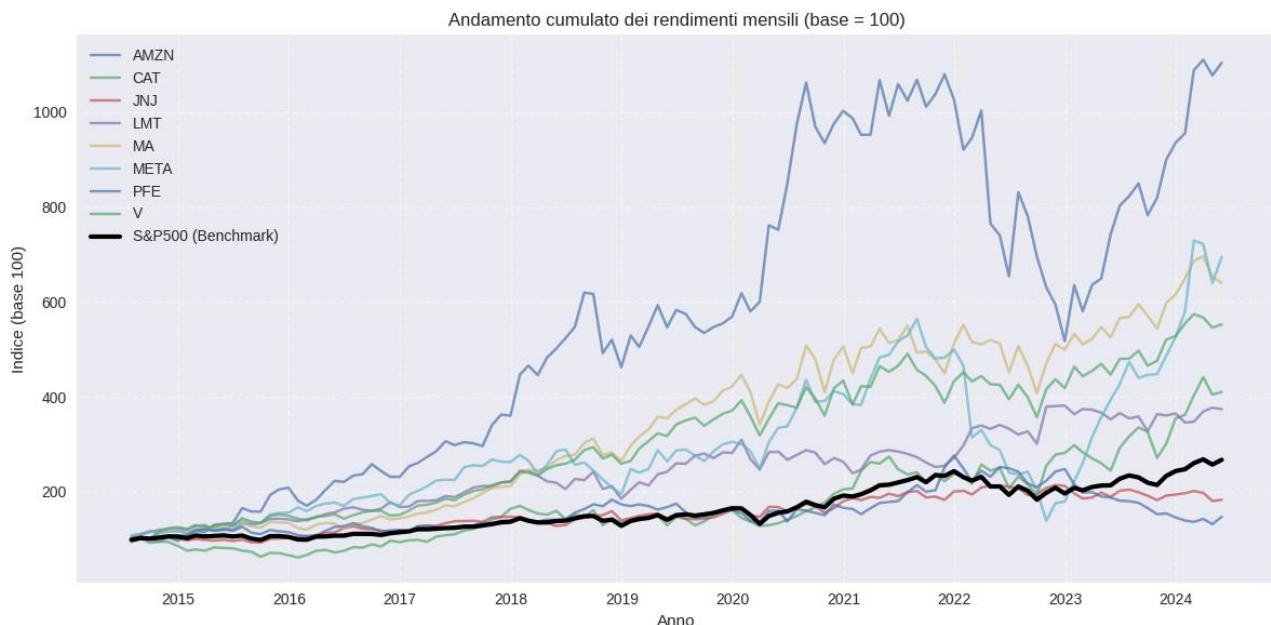
- pandas.concat (con axis=1, talvolta join="inner")
- pandas.DataFrame.merge
- pandas.DataFrame.reindex (con method="ffill" per allineare le date)
- pandas.Index.to\_timestamp (conversione PeriodIndex → Timestamp per l'allineamento)

## **Presentazione dei dati:**

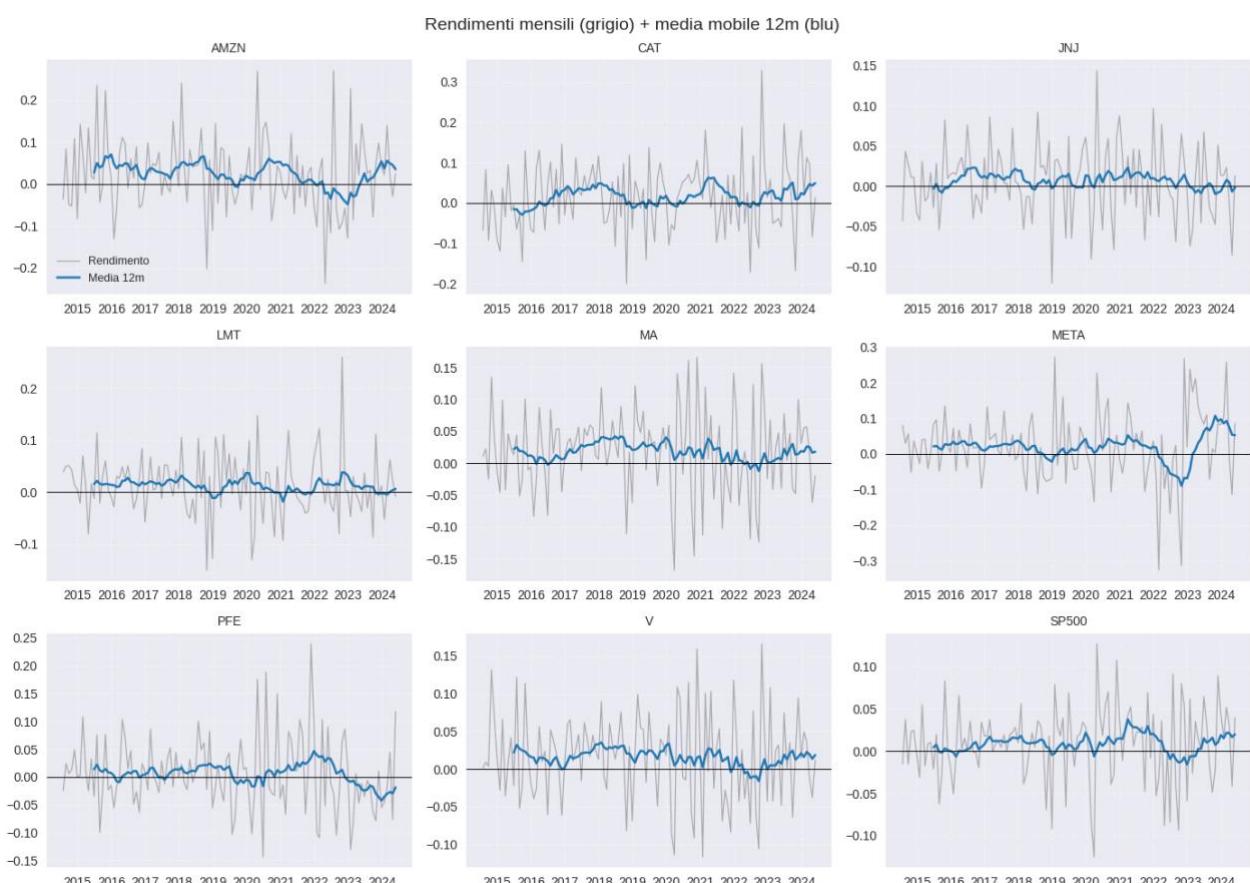
I dati giornalieri sono stati scaricati da Yahoo! Finance tramite la libreria *yfinance* e successivamente convertiti in frequenza mensile. La tabella seguente mostra le prime osservazioni disponibili per ciascun titolo e per l'indice S&P500 utilizzato come benchmark.

	AMZN	CAT	JNJ	LMT	MA	META	PFE	V	SP500
Date									
2014-06-02	15.442000	77.425217	74.707779	120.972313	71.922516	62.735828	17.723400	49.329597	1924.969971
2014-06-03	15.359500	77.969963	74.919807	120.810135	71.269081	62.526966	17.651819	48.821335	1924.239990
2014-06-04	15.339000	77.835632	75.080696	121.274536	71.241051	62.994411	17.681644	48.846745	1927.880005
2014-06-05	16.178499	79.813026	75.475540	122.608765	71.726463	62.845226	17.753229	49.029266	1940.459961
2014-06-06	16.483500	80.723389	75.446289	123.227928	72.314575	62.158989	17.550407	49.209473	1949.439941

Il grafico seguente mostra l'andamento cumulato dei rendimenti mensili dei titoli selezionati e dell'indice S&P 500, normalizzati a 100. La normalizzazione fa partire tutti i titoli dal medesimo livello di 100, rendendo il confronto tra l'andamento dei titoli visibilmente più semplice e immediato.



Qui di seguito l'andamento dei rendimenti mensili dei titoli selezionati e dell'indice S&P 500 (in grigio), affiancato alla relativa media mobile a 12 mesi (in blu). In questo modo è possibile osservare l'incertezza e la variabilità mensile dei rendimenti ma anche i trend di medio periodo attraverso la media mobile, che rende più leggibile l'andamento sottostante nonostante le variazioni mensili.



## STATISTICHE DESCRIPTIVE

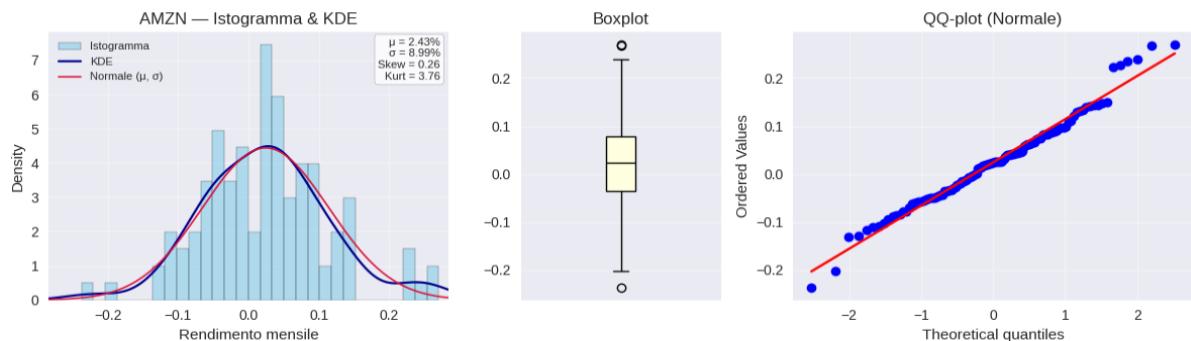
L'esame delle performance a 12 mesi evidenzia come i rendimenti medi più alti siano stati conseguiti dalle azioni tecnologiche Amazon (29,14%) e Meta (25,43%). Non a caso il settore tecnologico ha guidato la crescita dell'economia americana negli ultimi anni e oggi costituisce una grande fetta dell'indice S&P 500. Seguono Mastercard e Visa che pur essendo appartenenti al settore dei pagamenti sono rimaste al traino della tecnologia grazie anche alla crescente digitalizzazione dei consumi. I rendimenti più contenuti sono, invece, quelli dei titoli del settore sanitario con Pfizer (6,35%) e Johnson & Johnson (7,27%), che hanno registrato una crescita più contenuta nel periodo preso in esame.

	mean_ann	var_ann	std_ann	skewness	kurtosis
<b>AMZN</b>	0.2914	0.0970	0.3114	0.2624	3.7637
<b>CAT</b>	0.1866	0.0891	0.2984	0.2115	3.6697
<b>JNJ</b>	0.0727	0.0240	0.1550	-0.0235	3.3239
<b>LMT</b>	0.1544	0.0426	0.2065	0.3964	5.1380
<b>MA</b>	0.2148	0.0529	0.2301	-0.2405	3.0940
<b>META</b>	0.2543	0.1135	0.3368	-0.2558	5.0650
<b>PFE</b>	0.0635	0.0512	0.2262	0.6090	4.1071
<b>V</b>	0.1936	0.0410	0.2026	0.0457	2.8123

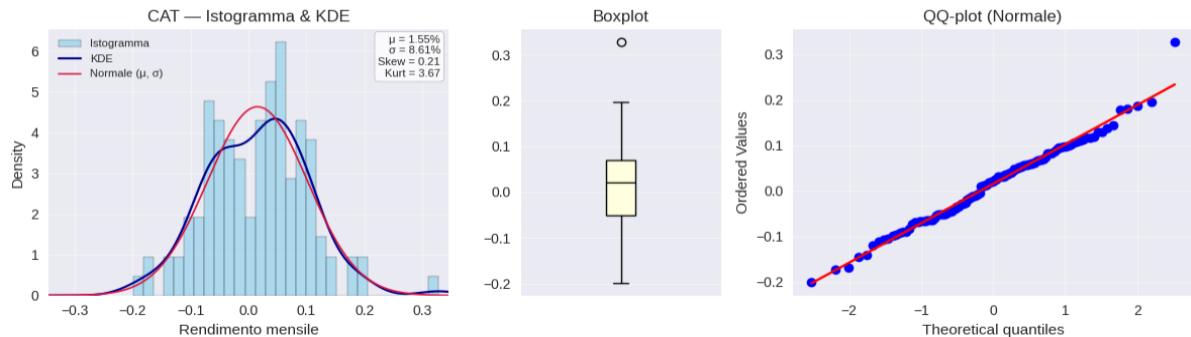
Per quanto riguarda la volatilità, Meta risulta il titolo più volatile (33,68%), seguita da Amazon (31,14%), Caterpillar (29,84%) e Mastercard (23,01%). All'estremo opposto si trovano Johnson & Johnson, che presenta la volatilità più contenuta (15,50%), e Visa (20,26%).

### Di seguito alcuni grafici sui rendimenti:

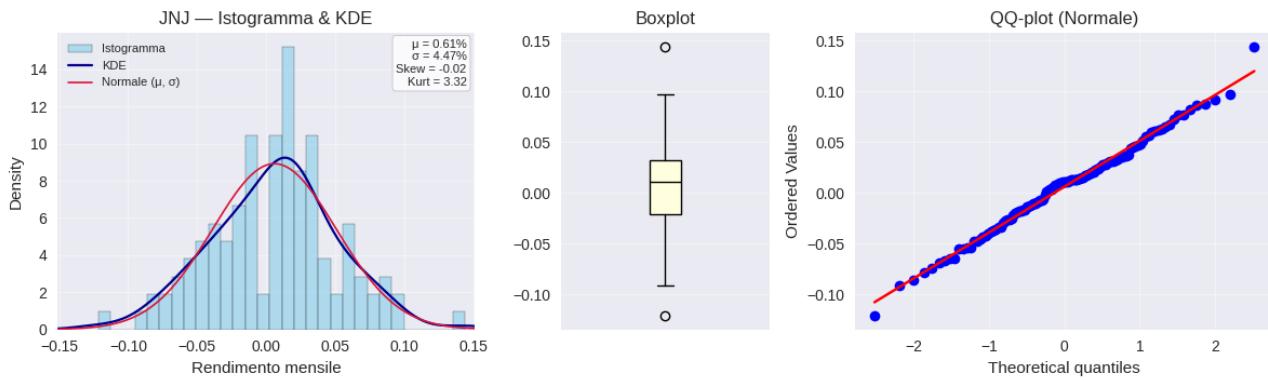
Diagnostica rendimenti mensili — AMZN



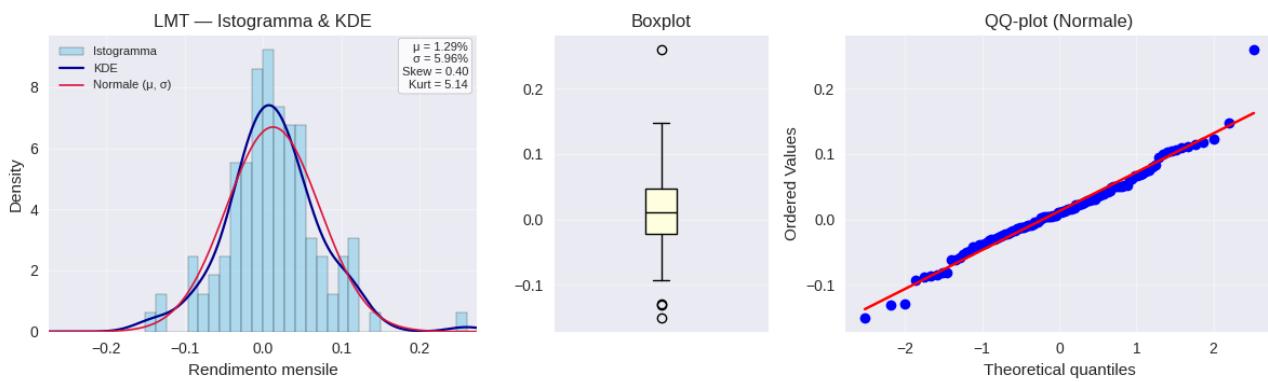
Diagnostica rendimenti mensili — CAT



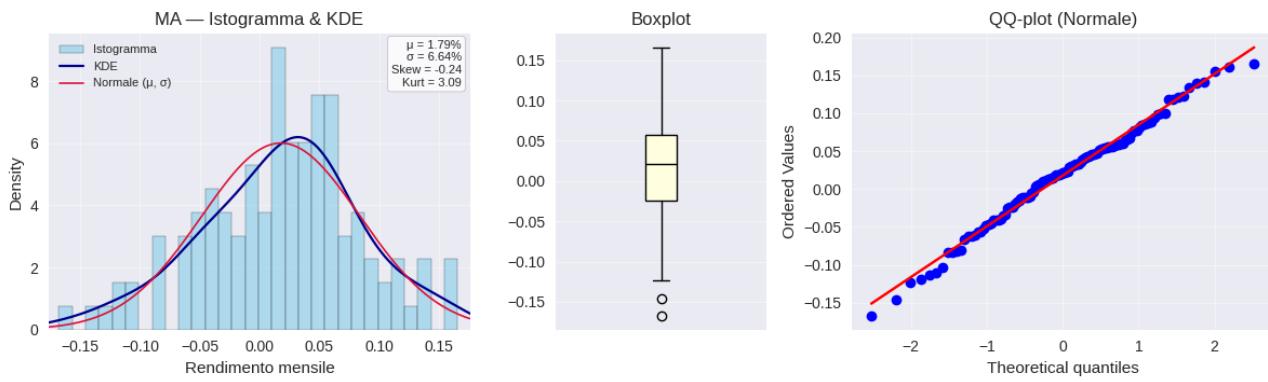
### Diagnostica rendimenti mensili — JNJ



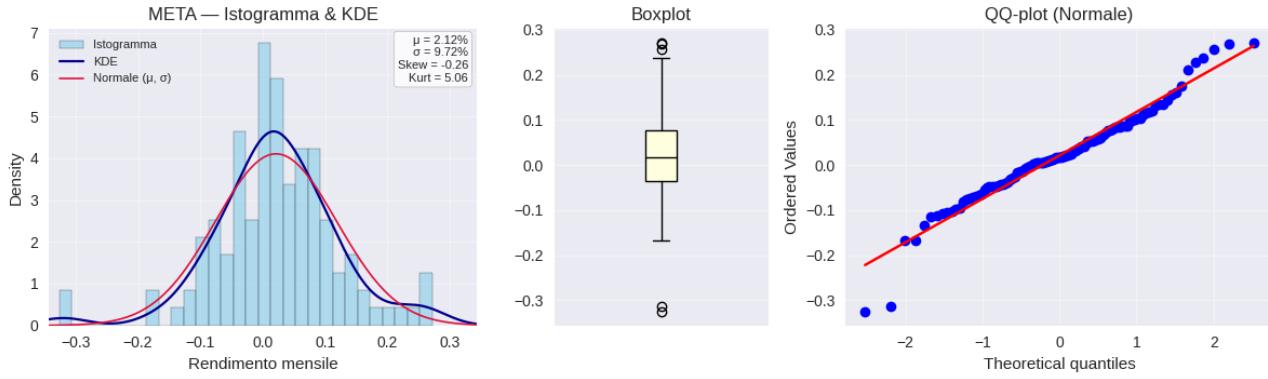
### Diagnostica rendimenti mensili — LMT



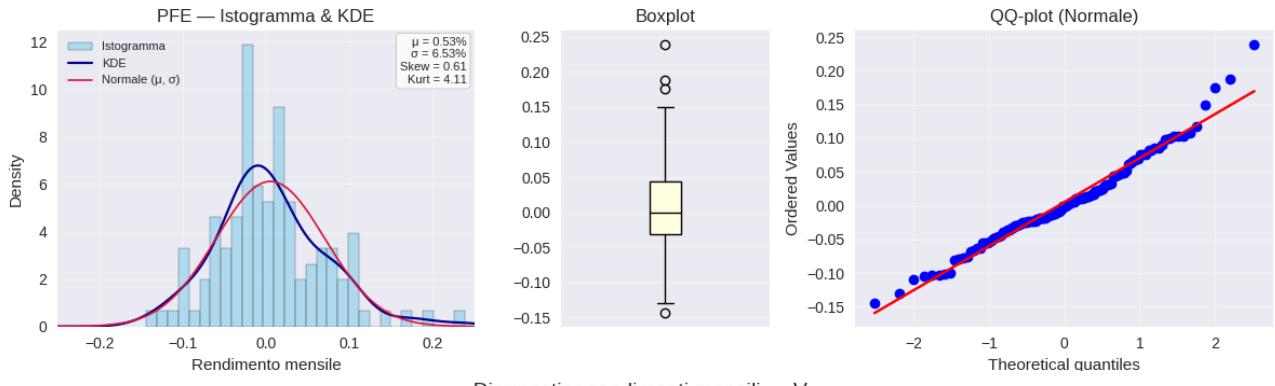
### Diagnostica rendimenti mensili — MA



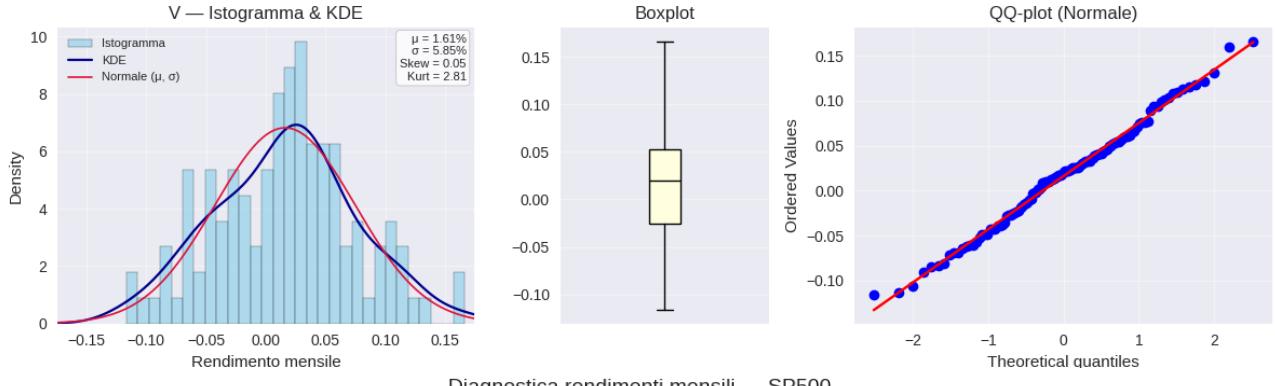
### Diagnostica rendimenti mensili — META



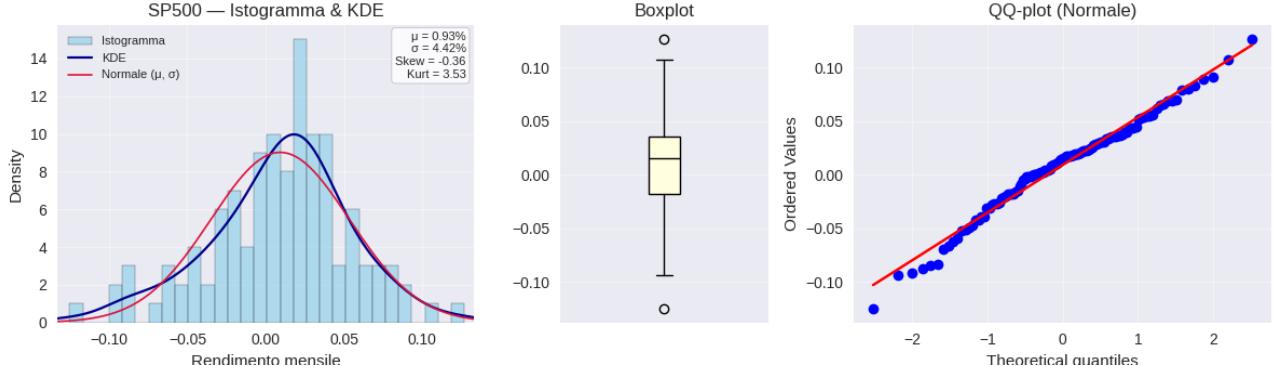
### Diagnostica rendimenti mensili — PFE



### Diagnostica rendimenti mensili — V



### Diagnostica rendimenti mensili — SP500



## ANALISI GRAFICI:

### Dispersione (volatilità)

I più dispersi sono i titoli tech che mostrano oscillazioni mensili ampie, anche Caterpillar è elevata; seguono i titoli dei pagamenti digitali e per ultimi gli healthcare che risultano i meno dispersivi (alcuni picchi isolati per Pfizer). Questo è in linea con le stime annualizzate di  $\sigma$  calcolate in precedenza.

### Normalità dei rendimenti

Dai QQ-plot e dall'excess kurtosis  $> 0$  in molte serie si vedono code più spesse della normale. Il Jarque-Bera indica come più vicini alla normale: Visa e Johnson & Johnson (p-value alti); mentre come più lontani: Lockheed Martin, Meta e Pfizer (p-value molto bassi => si rifiuta la normalità).

	<b>AMZN</b>	<b>CAT</b>	<b>JNJ</b>	<b>LMT</b>	<b>MA</b>	<b>META</b>	<b>PFE</b>	<b>V</b>
<b>JB_stat</b>	3.6376	2.6018	0.3472	22.8562	1.1259	19.7198	12.2410	0.3025
<b>JB_p</b>	0.1622	0.2723	0.8406	0.0000	0.5695	0.0001	0.0022	0.8596

In sintesi: l'assunzione di normalità non regge bene per diversi titoli, specie dove compaiono code/outlier.

## Outlier

I punti al di fuori dei baffi dei boxplot sono mesi anomali (outlier). Corrispondono a rendimenti sensibilmente distanti dalla media. Dal confronto dei grafici di diagnostica e dall'analisi dell'andamento dei rendimenti mensili emergono alcuni eventi di mercato che hanno scatenato dei movimenti eccezionali comuni o più marcati per alcune azioni:

- **Marzo 2020 - Shock COVID:** rendimenti molto negativi su quasi tutte le azioni e sull'indice S&P 500, legati al sell-off generalizzato dei mercati nella prima fase della pandemia.
- **Novembre 2020 - Prime notizie sui vaccini:** forte rimbalzo dei mercati azionari, con rialzi generalizzati e trainati, tra gli altri, da Pfizer (PFE) grazie ai successi sui vaccini.
- **Febbraio-marzo 2022 - Invasione Russia-Ucraina:** forte spike positivo per Lockheed Martin (LMT) e, in generale, per il settore difesa, mentre gli altri titoli sono stati pressionati.
- **2022 - Crollo di Meta (META):** molti mesi di pesanti rendimenti negativi sull'onda del calo dei ricavi pubblicitari, del deterioramento generato dalle policy idonee a proteggere la privacy (IDFA) e dei consistenti investimenti in Reality Labs.

Questi sono alcuni dei principali driver che hanno caratterizzato le performance di estremo segno delle serie. In tabella sono riepilogati i due mesi migliori e peggiori per ciascun titolo, con l'indicazione di data, rendimento e evento di mercato.

Ticker	Data (Top 1)	Evento (Top 1)	Rendimento (Top 1)	Data (Top 2)	Evento (Top 2)	Rendimento (Top 2)	Data (Bottom 1)	Evento (Bottom 1)	Rendimento (Bottom 1)	Data (Bottom 2)	Evento (Bottom 2)	Rendimento (Bottom 2)
AMZN	2022-07	Post-trimestrale; AWS solido	+27.06%	2020-04	Rimbalzo post-shock	+26.89%	2022-04	Q1 sotto attese; guida e-commerce debole	-23.75%	2018-10	Risk-off su tassi/growth	-20.22%
CAT	2022-10	Rally da minimi; rotazione ciclici (CPI reversal)	+32.81%	2023-06	Infra USA; backlog solido	+19.59%	2018-10	Risk-off su tassi/growth	-19.93%	2022-06	Fed 75 bps + QT; recessione	-17.18%
JNJ	2020-04	Rimbalzo post-shock	+14.42%	2021-12	Flussi difensivi di fine anno	+9.71%	2018-12	Christmas sell-off (Fed/crescita)	-12.15%	2020-02	Inizio sell-off COVID	-9.09%
LMT	2022-10	Q3 > attese; buyback	+25.99%	2020-04	Rimbalzo post-shock	+14.78%	2018-10	Risk-off su tassi/growth	-15.06%	2020-02	Inizio sell-off COVID	-13.06%
MA	2020-11	Vaccini / re-opening	+16.58%	2020-08	Riaperture; cross-border su	+16.10%	2020-03	Shock COVID	-16.78%	2020-10	Nuove restrizioni UE; travel debole	-14.55%
META	2019-01	Fed più dovish	+27.16%	2022-11	CPI sotto attese → rally risk	+26.77%	2022-02	Post-trim: guida debole; IDFA	-32.63%	2022-10	Post Q3: Reality Labs; guida debole	-31.34%
PFE	2021-11	Paxlovid: dati positivi	+23.92%	2020-07	Progressi vaccino	+18.83%	2020-06	Rotazione risk-on (difensivi deboli)	-14.38%	2023-01	Normalizzazione post-COVID; guida prudente	-13.03%
SP500	2020-04	Rimbalzo post-shock	+12.68%	2020-11	Vaccini / re-opening	+10.75%	2020-03	Shock COVID	-12.51%	2022-09	Fed molto hawkish (dots alti)	-9.34%
V	2022-10	Q4 FY22 > attese; cross-border solido	+16.61%	2020-11	Vaccini / re-opening	+15.94%	2021-01	Rotazione ciclici; travel debole	-11.85%	2020-03	Shock COVID	-11.36%

## Correlazioni

Nel nostro campione emerge una correlazione positiva intra-settore, con intensità diversa a seconda dei business. Nei pagamenti Visa-Mastercard sono nettamente i più correlati ( $r \approx 0,91$ ): la relazione è stabile nel tempo (rolling quasi sempre 0,8–0,95), segno di driver comuni su spesa con carta e

volumi cross-border. Nell'healthcare Pfizer–Johnson & Johnson mostrano una correlazione moderata ( $r \approx 0,54$ ) ma molto variabile: scende nelle fasi di shock quando prevalgono notizie specifiche (trial, guidance, cause legali). Nel tech/consumer Meta-Amazon hanno correlazione media ( $r \approx 0,47$ ), alta nei grandi rally del comparto (2020–21) e più bassa nel 2022 per gli shock idiosincratici di Meta (privacy iOS, ricavi ads).

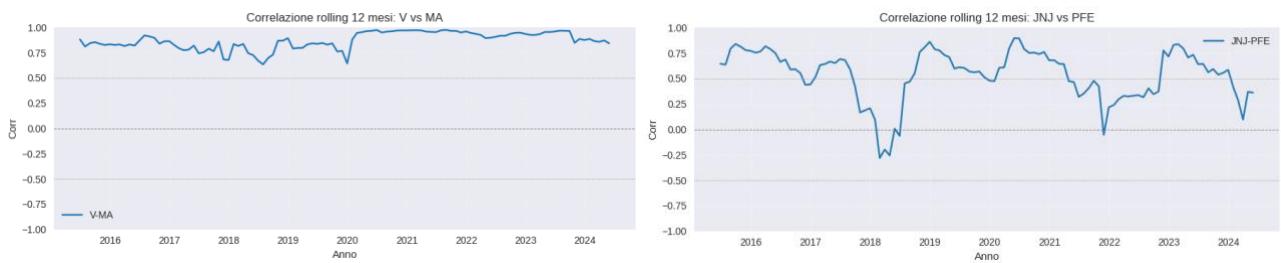
Guardando alla correlazione con l'indice S&P 500, i titoli più ciclici (V, MA, AMZN, META, CAT) stanno in area 0,6-0,7, mentre i difensivi (JNJ, PFE, in parte LMT) risultano meno correlati e più instabili: la sensibilità al mercato cala proprio durante turbolenze, quando entrano in gioco fattori settoriali/societari.

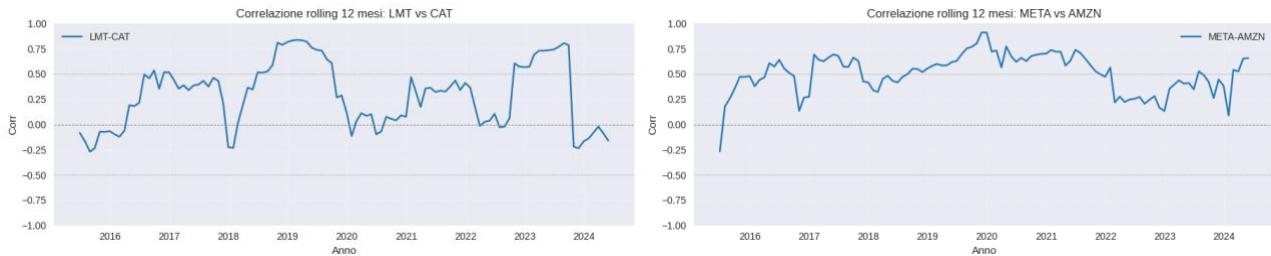


Dalla matrice di correlazione risulta che:

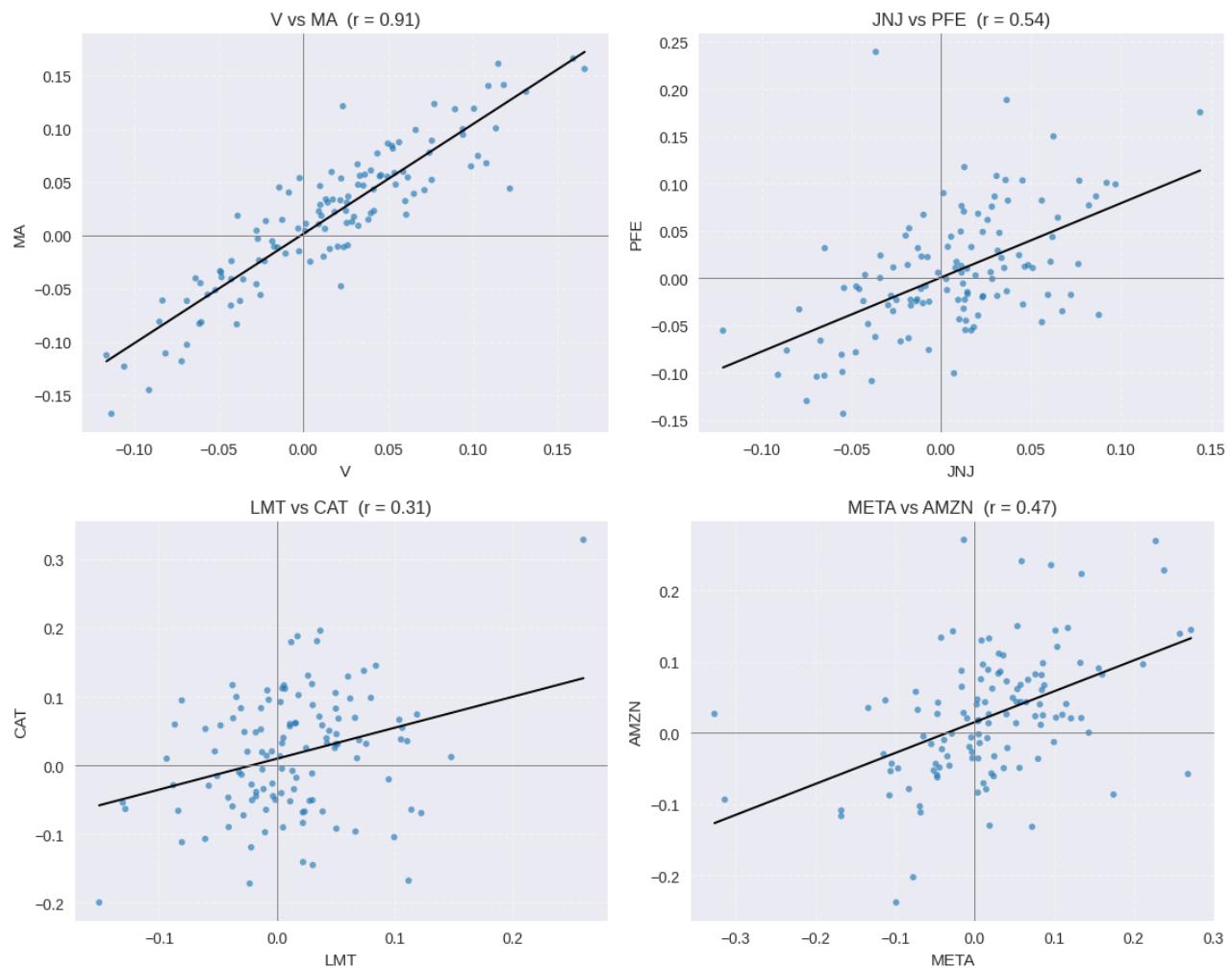
- **Più correlati:** V-MA (~0,91); molto alte anche le correlazioni di ciascuno di essi con l'S&P 500 (~0,7).
- **Meno correlati:** coppie “eterogenee” come LMT-META (~0,03), e più in generale META con alcuni difensivi (META-JNJ ~0,21; META-CAT ~0,17).

L'andamento nel tempo conferma che le correlazioni non sono costanti: tendono ad aumentare nei periodi di trend di mercato (risk-on) (es. rally vaccini di novembre 2020 o “Fed pivot” di fine 2023) e a ridursi o divergere negli shock in cui prevalgono driver specifici (es. guerra 2022: LMT sale per la domanda difesa mentre altri ciclici soffrono). Questo pattern è chiarissimo nei rolling: V-MA resta sempre alta; JNJ-PFE oscilla ampiamente; LMT-CAT passa da >0,7 in fasi di ciclo forte a valori vicino/sotto 0 quando i driver si separano.





Gli scatter plot confermano visivamente la relazione lineare: V-MA mostra una nuvola di punti stretta lungo la diagonale (forte linearità), mentre JNJ-PFE, LMT-CAT e META-AMZN presentano maggiore dispersione, coerente con correlazioni moderate/debolì. In presenza di rendimenti estremi, la nuvola tende a “stirarsi” e la correlazione si muove: si amplifica quando lo shock è sistemico e tocca tutti, si riduce quando l’evento è idiosincratico.



*Nota metodologica:* la matrice di varianze/covarianze e la matrice di correlazione sono state calcolate sui rendimenti mensili; la lettura dettagliata è quella mostrata nelle heatmap, mentre la dinamica temporale è catturata dai grafici di correlazione rolling a 12 mesi.

## ANALISI DI PREVISIONE (ARIMA)

Abbiamo stimato dei modelli per prevedere i rendimenti logaritmici mensili dei titoli/indice. È stata divisa la serie storica in: 80 mesi di training, 30 mesi di test e infine 10 mesi di previsioni out-of-sample. Sono stati utilizzati due modelli: un ARIMA(1,0,1) identico per tutti e un auto-ARIMA che seleziona l'ordine più efficiente. Dopo aver fatto il fit sul train abbiamo previsto il test; poi è stato rifittato il modello su train+test per ottenere un 10-step-ahead e confrontarlo con gli ultimi valori osservati. Le prestazioni sono misurate con RMSE e MAE sia sul test sia sui 10 mesi di forecast.

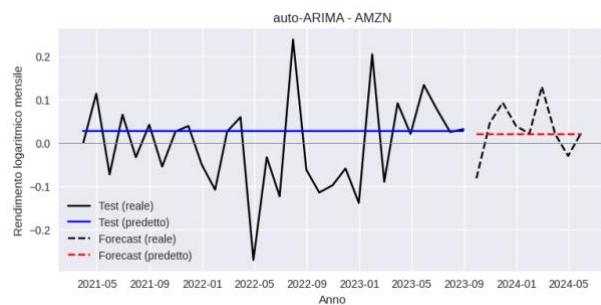
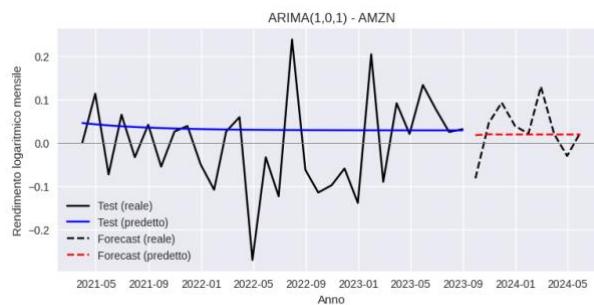
ARIMA

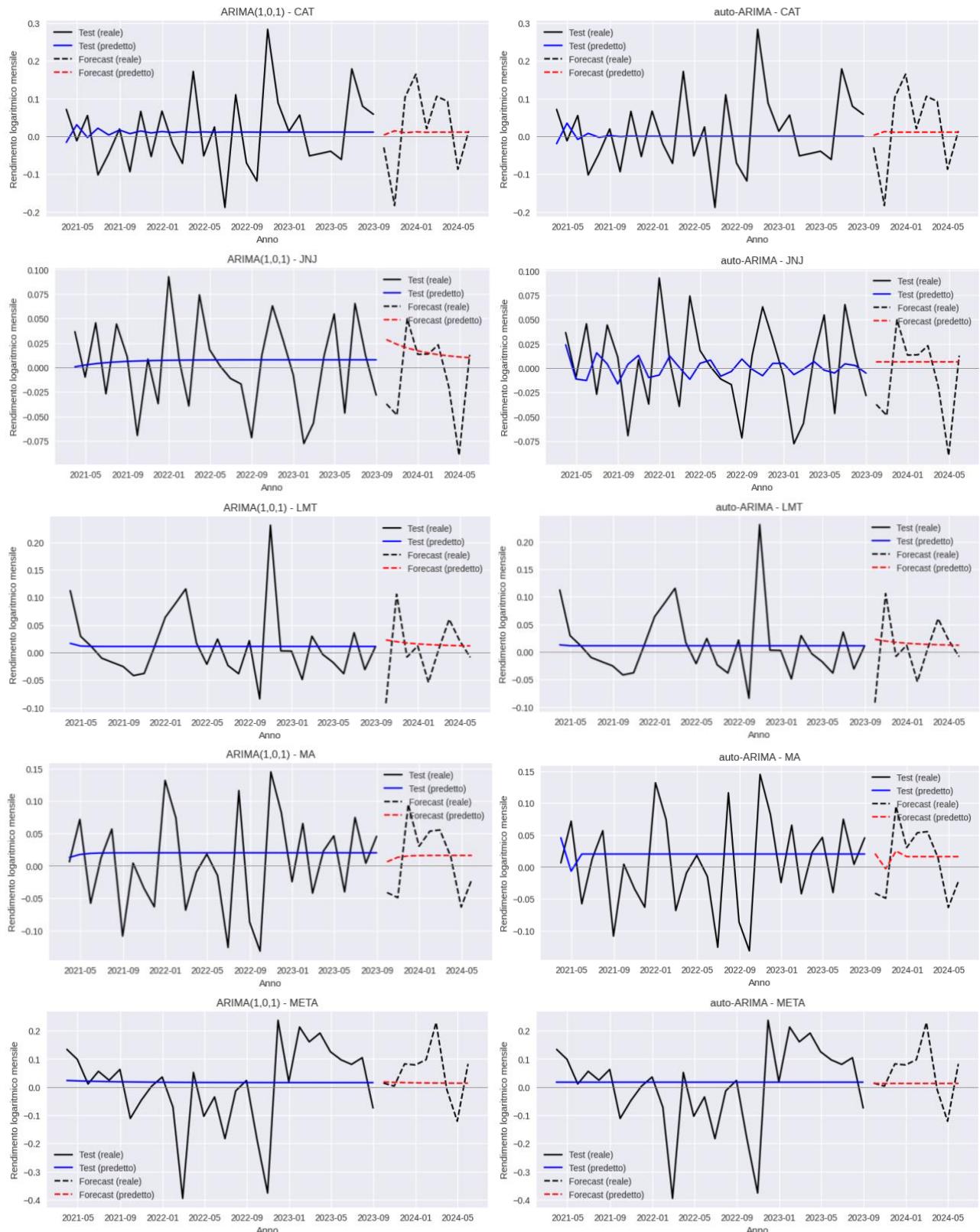
	Ticker	RMSE_test	MAE_test	RMSE_forecast	MAE_forecast
0	AMZN	0.1108	0.0856	0.0589	0.0430
1	CAT	0.0991	0.0808	0.1044	0.0851
2	JNJ	0.0447	0.0359	0.0491	0.0353
3	LMT	0.0605	0.0419	0.0566	0.0428
4	MA	0.0737	0.0604	0.0509	0.0443
5	META	0.1474	0.1074	0.0972	0.0746
6	PFE	0.0771	0.0613	0.0684	0.0605
7	V	0.0660	0.0559	0.0428	0.0330
8	SP500	0.0525	0.0448	0.0441	0.0396

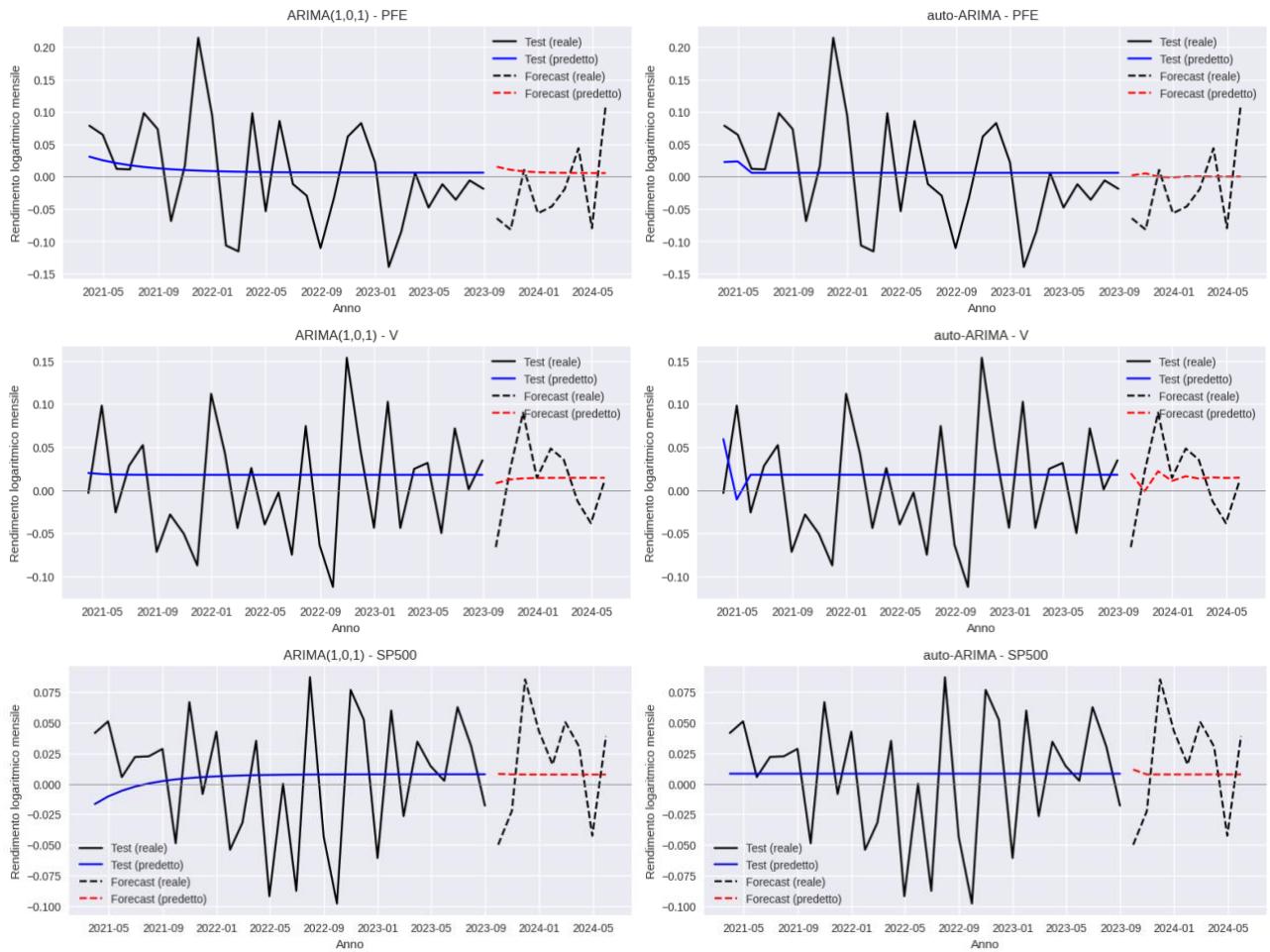
Auto-ARIMA

	Ticker	Order	RMSE_test	MAE_test	RMSE_forecast	MAE_forecast
0	AMZN	(0, 0, 0)	0.1097	0.0847	0.0593	0.0433
1	CAT	(2, 0, 0)	0.0997	0.0813	0.1040	0.0847
2	JNJ	(3, 0, 2)	0.0459	0.0360	0.0437	0.0336
3	LMT	(0, 0, 1)	0.0607	0.0420	0.0566	0.0428
4	MA	(0, 0, 2)	0.0748	0.0623	0.0492	0.0431
5	META	(0, 0, 0)	0.1475	0.1077	0.0978	0.0746
6	PFE	(0, 0, 2)	0.0778	0.0620	0.0652	0.0578
7	V	(0, 0, 2)	0.0682	0.0582	0.0441	0.0352
8	SP500	(0, 0, 0)	0.0509	0.0425	0.0446	0.0400

Dal punto di vista grafico, ogni titolo è stato rappresentato con le parti di test, forecast e parte reale degli ultimi 10 mesi con forecast multi-step. Le curve blu/rosse sono sovente quasi piatte: è il comportamento tipico di modelli che, su frequenza mensile, non riescono a trovare una forte struttura autoregressiva e tendono a riportare il multi-step forecast alla media di lungo periodo. I risultati tabellari confermano questa interpretazione. Su titoli “tranquilli” e difensivi (JnJ, V) e sull’indice S&P 500 gli errori sono relativamente contenuti (nel test RMSE  $\approx$  0,04–0,07, nel forecast  $\approx$  0,04–0,05). Al contrario su titoli più volatili o ciclici (Meta, Amzn, Cat) gli errori sono molto più elevati (nel test RMSE anche  $>$  0,10–0,15), perché picchi e rimbalzi causati da news e trimestrali non si possono prevedere con una dinamica puramente ARIMA. L’auto-ARIMA seleziona quasi sempre ordini molto parsimoniosi (in più casi (0,0,0) o (0,0,1/2)), segnale che la parte prevedibile è poca; solo su alcune serie più statiche (es. JnJ) sceglie ordini un po’ più ricchi (es. (3,0,2)) con un lieve beneficio in termini di errore. In generale ARIMA e auto-ARIMA si comportano molto similmente, qualche piccolo vantaggio locale per l’auto-ARIMA, ma niente di sostanziale. Nel complesso le evidenze sono coerenti con l’idea, ben nota in letteratura, che i rendimenti azionari mensili siano vicini a un processo scarsamente autocorrelato: la prevedibilità è scarsa, i multi-step forecast tendono alla media e non catturano gli shock (macro, guidance, risultati, fatti imprevisti, geopolitica). Su orizzonti temporali come questi modelli ARIMA “puri univariati” non sono in grado di prevedere i colpi di scena.







## STRATEGIE DI TRADING & BACKTESTING - Analisi tecnica

Abbiamo valutato le performance di quattro strategie su tutti i titoli e sull'S&P 500, tenendo conto di costi di transazione pari a 10 bps per cambio posizione. Per le strategie tecniche abbiamo utilizzato prezzi giornalieri, per la strategia di ML (Random Forest) rendimenti e caratteristiche mensili.

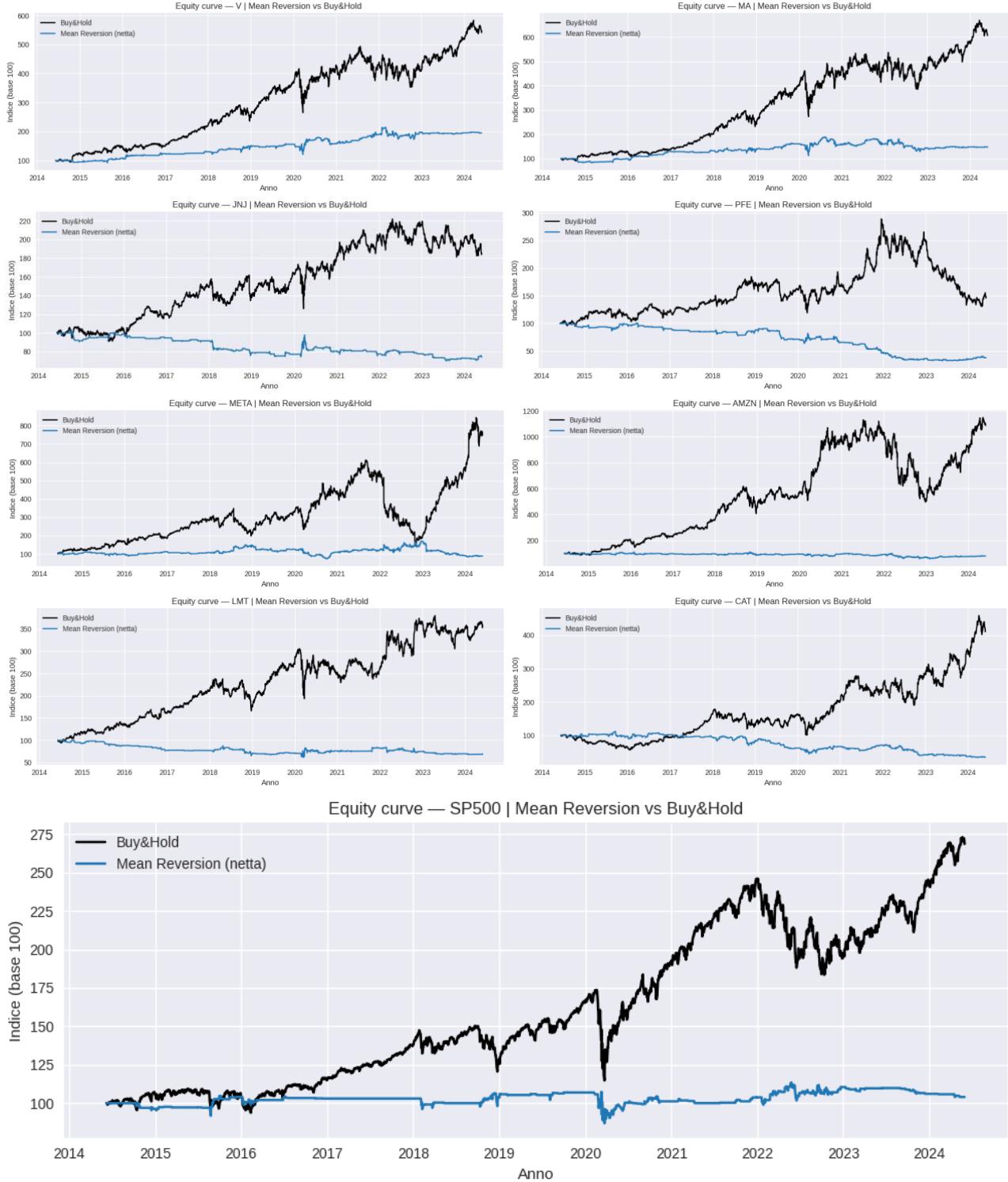
- **Mean Reversion 5 giorni (long/short):** si entra long quando il rendimento cumulato a 5 giorni scende sotto -3%, short quando supera +3%; la posizione effettiva entra con un giorno di ritardo.
- **Mean Reversion 5 giorni (long-only):** come sopra, ma long/flat.
- **SMA 50/200 (long-only):** long quando SMA50 > SMA200, flat altrimenti.
- **Random Forest (long-only):** classificatore mensile con caratteristiche derivate dai prezzi del titolo (RSI 6m, gap dalla SMA12, volatilità 3m, momentum 3m, lag dei rendimenti), split temporale 70 mesi train + 30 mesi test; si entra long quando il modello prevede rendimento > 0 (probabilità > 0,5), altrimenti si resta flat.

Il benchmark è Buy&Hold sul titolo/indice. Per ogni strategia riportiamo CAGR, volatilità annualizzata, Sharpe, Max Drawdown, equity curve e un grafico "volatilità vs Sharpe" con il posizionamento medio delle strategie.

### Risultati sintetici per strategia

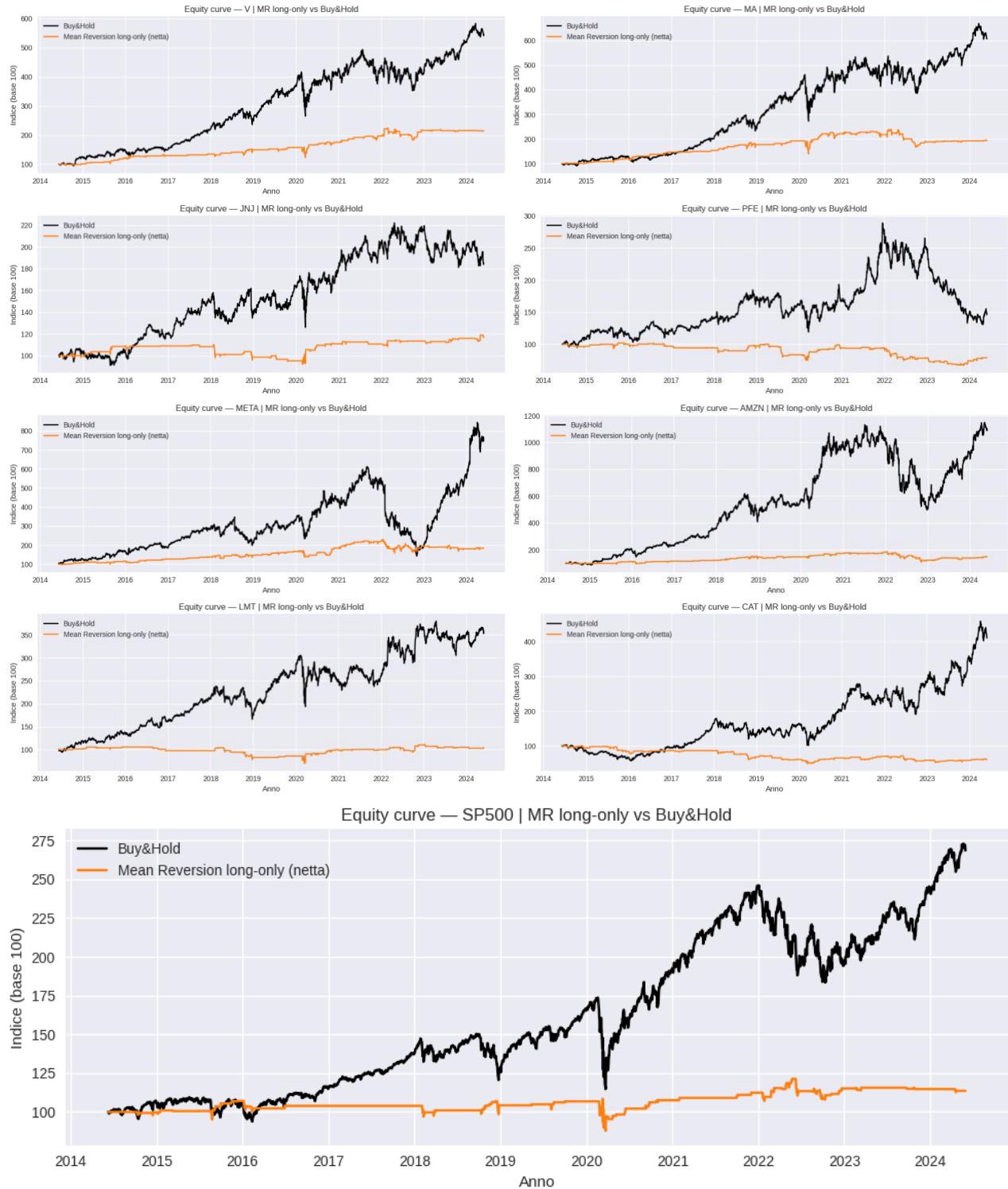
#### 1) Mean Reversion 5g (long/short):

Questa è senza dubbio la peggiore. Il trend rialzista di lungo periodo dei mercati azionari ed il costo dello short erodono le performance: il CAGR risulta negativo su gran parte dei titoli (ad es. CAT – 9,9%, PFE –9,3%, JNJ –2,9%) e modesto/leggermente positivo solo in qualche caso (ad es. V 7,0%, MA 3,9%, S&P 500 ~0,4%). Anche lo Sharpe medio è debole o negativo e le equity line appaiono livellate rispetto al Buy&Hold. In definitiva, il rimbalzo di breve periodo non appare né sistematico né abbastanza diffuso, e i costi finiscono per eroderlo completamente.



## 2) Mean Reversion 5g (long-only):

Rimuovere lo short aiuta: il profilo migliora e il CAGR diventa positivo su diversi titoli growth e sull'indice (AMZN 4,0%, META 6,3%, MA 6,8%, V 8,0%, S&P 500 1,3%). Tuttavia si pone ancora al di sotto del Buy&Hold quasi ovunque. Positivo è il rischio: volatilità più contenuta (ad es. V 0,15 contro 0,25 del BH) e MaxDD ridotto (V -22,8% contro -36,4%).



### 3) Trend-following SMA 50/200 (long-only):

Si tratta dell'approccio più solido e convincente tra quelli sperimentati. Il CAGR è positivo su quasi tutti i titoli (fanno eccezione JNJ -0,5% e PFE -1,4%) e molto vicino al Buy&Hold sui growth/tech:

AMZN 24,2% contro 27,1%, META 21,1% contro 22,4%. In cambio restituisce volatilità e drawdown più contenuti (ad es. AMZN MaxDD -39,6% contro -56,2%). Sul paniere S&P 500 la strategia realizza un 6,1% contro un 10,4% del BH, ma con rischio più benigno.



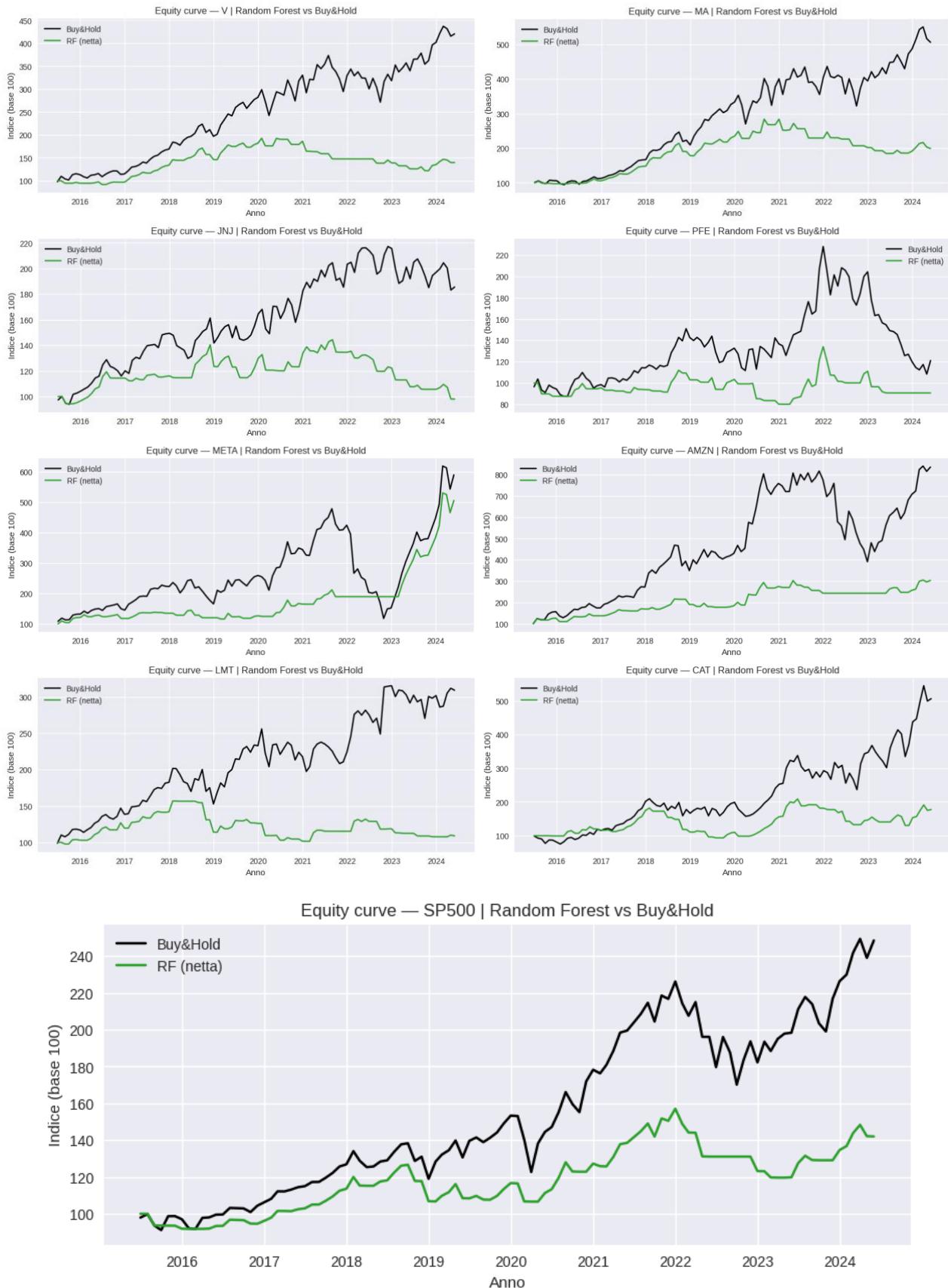
#### 4) Random Forest (long-only, mensile):

I risultati appaiono disomogenei, ma rivelano spunti interessanti su alcuni titoli:

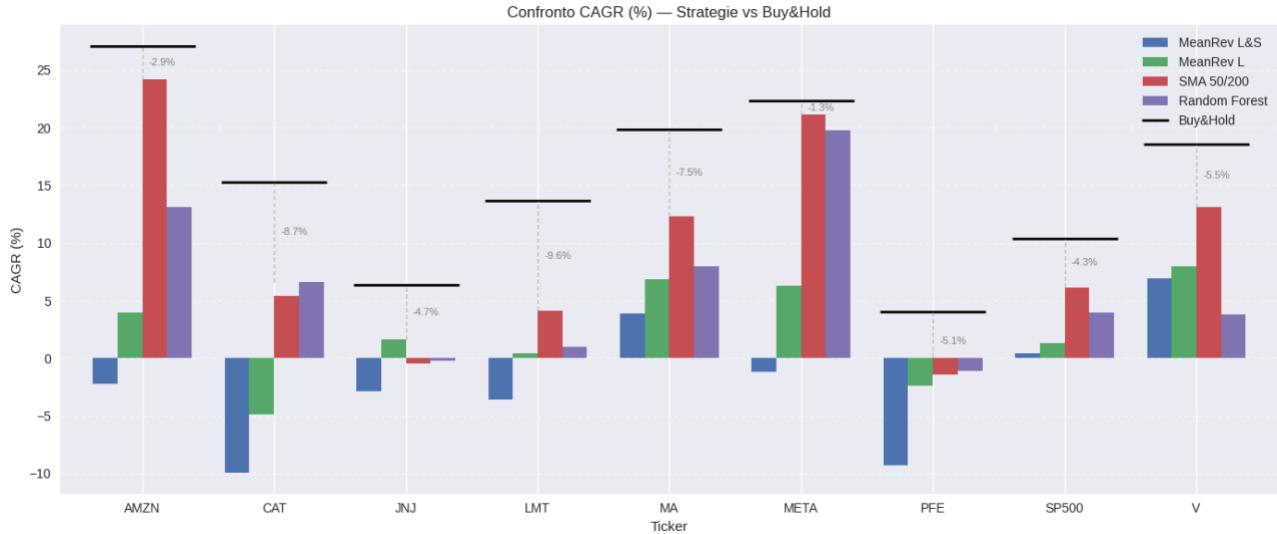
- META 19,7% e AMZN 13,1% (Sharpe circa pari a 0,96 e 0,72) con MaxDD in misura inferiore al BH.

- Su CAT (6,6%) e MA (8,0%) il modello apporta valore rispetto alle mean-reversion.
- Debole su JNJ (-0,2%) e PFE (-1,1%).

Mediamente il RF non prevale sul Buy&Hold, ma ne attenua volatilità e drawdown e, rispetto a strategie di trading fisse, è in grado di adeguarsi meglio ai cambi di regime su alcuni growth.



## Confronto trasversale con Buy&Hold:



- Per quanto riguarda i puri rendimenti (CAGR) il Buy&Hold rimane il migliore: AMZN ~27%, META ~22%, V ~18,6%, MA ~19,8%, S&P 500 ~10,4%.
  - La SMA 50/200 è la migliore seconda piazza “regolare”: spesso ancora “prima tra le seconde” dietro il BH (AMZN 24,2%, META 21,1%, V 13,1%, MA 12,3%, S&P 500 6,1%), ma con drawdown e volatilità minori.
  - Le mean-reversion sono in difficoltà: la long/short soffre costi di transazione e bias rialzista; la long-only è più dignitosa ma comunque molto sottoperformante rispetto a BH.
  - Random Forest si colloca a metà: non è vincente, ma migliora il profilo rischio/rendimento su alcuni titoli (META, AMZN) e fornisce spunti per sviluppi con segnali macro o di sentimento.



Il grafico Volatilità vs Sharpe conferma questa lettura: il punto della media BH si trova in alto a destra (Sharpe più alto, vol più alta), la SMA si colloca poco sotto ma più a sinistra (meno rischiosa), la mean-reversion long-only è ancora più a sinistra ma con uno Sharpe modesto, mentre la long/short è la meno efficiente. Il RF sta nel quadrante intermedio, vicino alla SMA ma con dispersione più ampia per ticker

## Conclusioni operative

- Se si vuole massimizzare il rendimento, il Buy&Hold rimane imbattuto.
- Se si vuole migliorare il profilo rischio/rendimento, la SMA 50/200 è la soluzione più semplice ed efficace: riduce MaxDD in modo significativo con un relativamente modesto costo di CAGR.
- Mean-reversion va usata prudentemente: meglio long-only come filtro di timing in drawdown, non come strategia principale.
- Random Forest è promettente su titoli a regimi marcati (META/AMZN).

## CAPM & FATTORI DI RISCHIO

L'analisi è condotta sui rendimenti logaritmici mensili dei titoli, allineati all'S&P 500 (proxy di mercato). Per il CAPM assumiamo un T-Bill 1Y = 4% annuo come risk-free e un rendimento atteso di mercato = 8% annuo.

Il rendimento atteso via CAPM è quindi:  $E[R_i] = \{r_f + \beta_i [E(R_m) - r_f]\} = \{4\% + 4\% \cdot \beta_i\}$ .

### Beta rispetto al mercato (CAPM):

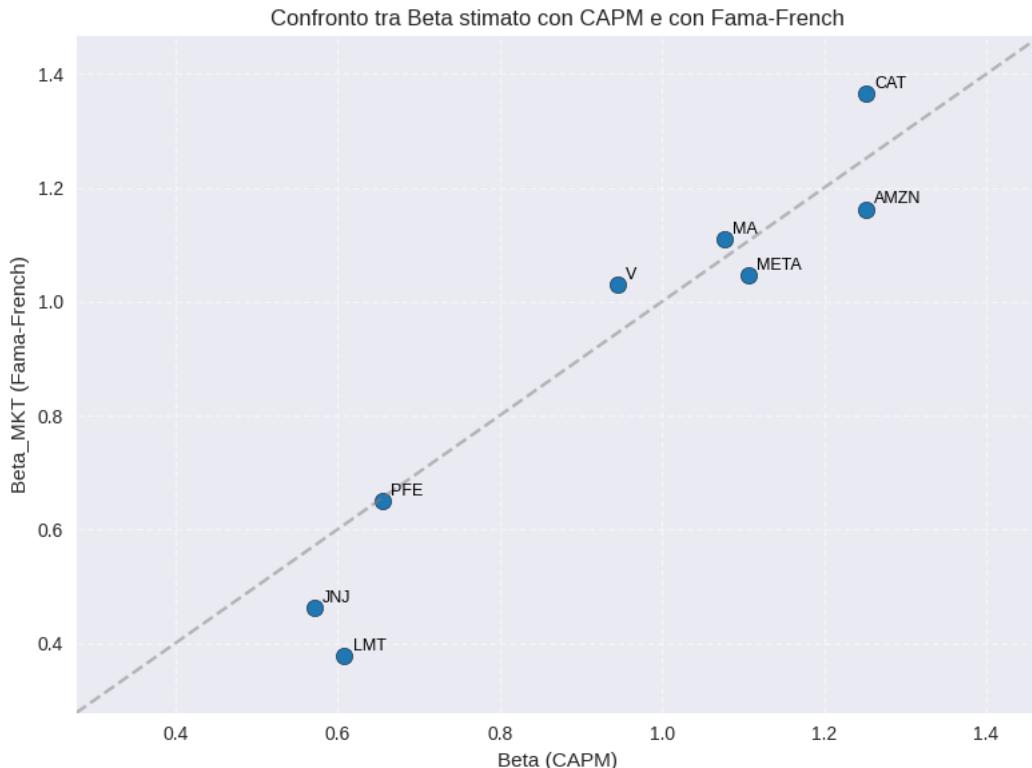
Dalla regressione  $R_i$  su  $R_m$  (con intercetta) otteniamo:

- **Beta più alti ( $\beta>1$ ):** AMZN ≈ 1,25, CAT ≈ 1,25, META ≈ 1,11, MA ≈ 1,08. Sono i titoli più ciclici/growth, quindi più sensibili al mercato.
- **Beta prossimi a 1:** V ≈ 0,95 (leggermente sotto-mercato).
- **Beta difensivi ( $\beta<1$ ):** PFE ≈ 0,66, LMT ≈ 0,61, JNJ ≈ 0,57.

Rendimenti attesi CAPM (annui): AMZN ≈ 9,0%, CAT ≈ 9,0%, META ≈ 8,4%, MA ≈ 8,3%, V ≈ 7,8%, PFE ≈ 6,6%, LMT ≈ 6,4%, JNJ ≈ 6,3%. Coerentemente, i titoli a  $\beta>1$  richiedono un premio più alto; i difensivi hanno attese più contenute.

	Tabella integrata CAPM + Fama-French:								
	Ticker	Alpha_x	Beta	E(R_i)_CAPM	Alpha_y	Beta_MKT	Beta_SMB	Beta_HML	R2
0	AMZN	0.0098	1.2520	0.0901	-0.0042	1.1627	0.0328	-0.8510	0.6479
1	CAT	0.0015	1.2515	0.0901	-0.0004	1.3659	0.1439	0.8529	0.6323
2	META	0.0071	1.1071	0.0843	0.0045	1.0457	-0.8064	-0.7792	0.2584
3	MA	0.0067	1.0778	0.0831	-0.0108	1.1102	-0.7068	0.3405	0.5248
4	V	0.0066	0.9460	0.0778	-0.0078	1.0294	-0.5703	0.2080	0.5610
5	PFE	-0.0022	0.6550	0.0662	-0.0111	0.6503	-0.6493	0.0598	0.1996
6	LMT	0.0061	0.6085	0.0643	-0.0052	0.3776	-0.5566	0.5750	0.2987
7	JNJ	0.0003	0.5718	0.0629	-0.0073	0.4627	-0.2627	0.2262	0.2666

## Confronto Beta CAPM vs Beta di mercato nel modello Fama-French:

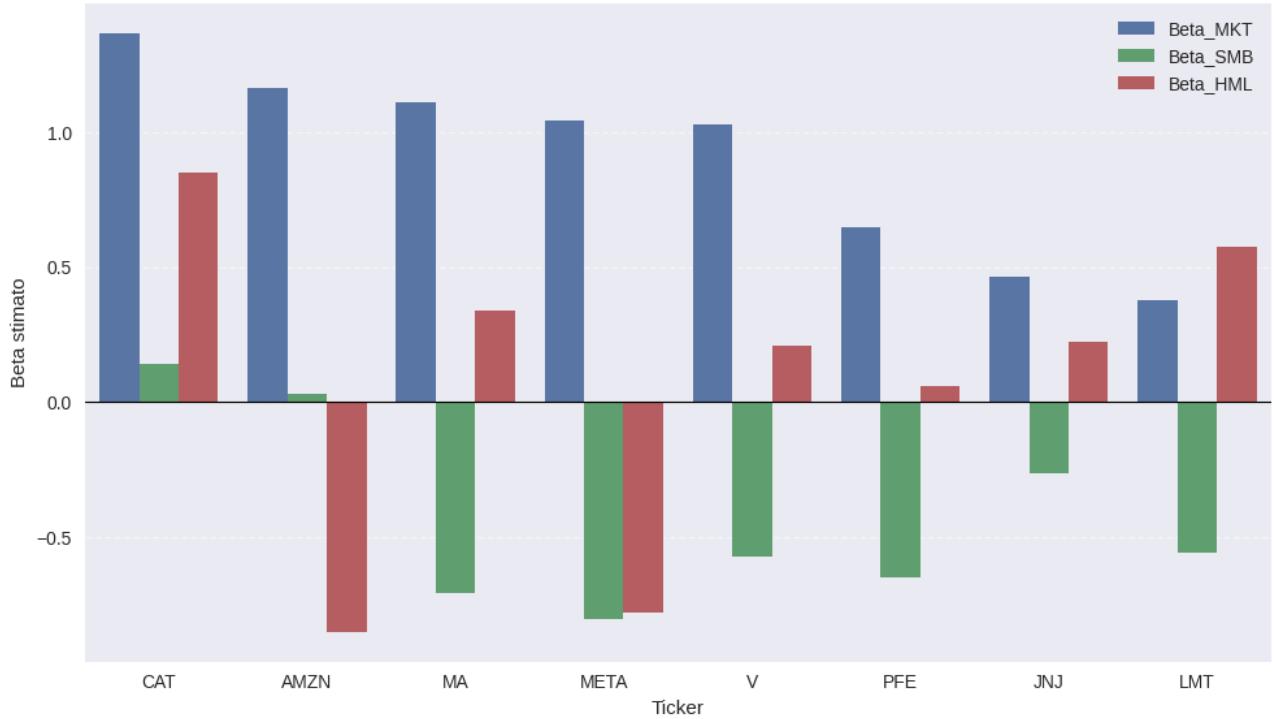


Nel grafico “Beta (CAPM) vs Beta\_MKT (Fama-French)” i punti si distribuiscono lungo la diagonale (coerenza fra stime), con alcune deviazioni interessanti:

- **CAT:**  $\beta_{\text{CAPM}} \sim 1,25$  vs  $\beta_{\text{MKT}} \sim 1,37 \rightarrow$  FF attribuisce ancora più sensibilità al mercato (controllando size/value).
- **AMZN:** 1,25 vs 1,16  $\rightarrow$  FF leggermente più basso (parte della varianza è spiegata dai fattori).
- **V:** 0,95 vs 1,03  $\rightarrow$  FF segnala una sensibilità di mercato un filo maggiore.
- **LMT:** 0,61 vs 0,38  $\rightarrow$  nel FF la market-beta si riduce molto; per defense la componente “value/size” spiega una quota importante della dinamica rispetto all’indice.
- In generale, **MA/META** sono molto vicini tra i due approcci; **PFE/JNJ** restano sotto-mercato in entrambi.

## Esposizioni ai fattori Fama-French (3F):

Esposizione ai fattori Fama-French 3



Stimiamo  $R_i - RF$  su (Mkt – RF, SMB, HML) con intercetta. I principali risultati (vedi barre nel grafico “Esposizione ai fattori Fama-French 3”):

- **Fattore di mercato (Mkt-RF):** in linea con i  $\beta$  CAPM: AMZN~1,16, CAT~1,37, META~1,05, MA~1,11, V~1,03; difensivi più bassi (PFE~0,65, JNJ~0,46, LMT~0,38).
- **SMB (size):** quasi tutti i titoli mostrano  $\beta_{SMB}$  negativi (MA ~-0,71; META ~-0,81; V ~-0,57; PFE ~-0,65; LMT ~-0,56; JNJ ~-0,26), coerente con il fatto che sono large/mega-cap. CAT è una delle poche eccezioni con  $\beta_{SMB}$  positivo (~+0,14); AMZN è circa neutro.
- **HML (value):** forte tilt growth ( $\beta_{HML} < 0$ ) per AMZN (~-0,85) e META (~-0,78); tilt value ( $\beta_{HML} > 0$ ) per CAT (+0,85) e LMT (+0,58); MA (+0,34) e V (+0,21) mostrano un leggero value; JNJ (+0,23) e PFE (+0,06) sono debolmente value/difensivi.
- **Qualità del fit ( $R^2$ ):** buona per AMZN (~0,65) e CAT (~0,63); media per V (~0,56) e MA (~0,52); più bassa per META (~0,26), PFE (~0,20) e JNJ (~0,27); in questi casi una quota maggiore di rendimento resta idiosincratica.

## Lettura complessiva

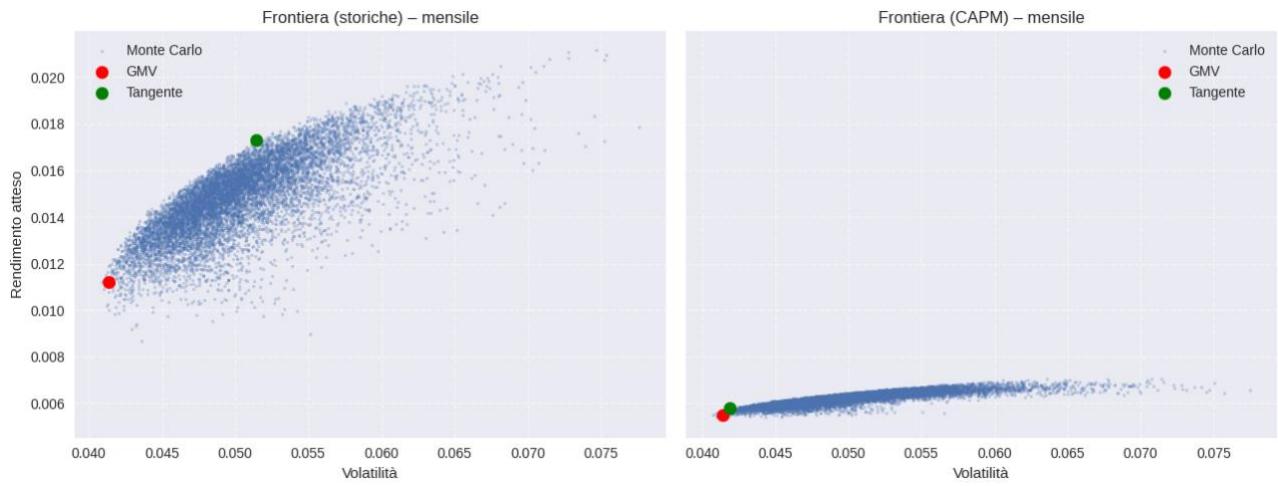
- La tassonomia growth vs value e large vs small che emerge dai fattori è molto pulita: AMZN/META sono growth-mega cap ( $HML < 0$ ,  $SMB \leq 0$ ); CAT/LMT value-large ( $HML > 0$ ,  $SMB < 0$ ), con CAT anche un po’ “meno mega” ( $SMB > 0$ ); MA/V risultano market-driven, leggermente value, molto large.
- Le differenze tra  $\beta$  CAPM e  $\beta_{MKT}$  derivano dal fatto che il modello a 3 fattori isola size/value: quando una parte dell’esposizione al mercato “nasconde” tilt strutturali (es. value forte su CAT, defense su LMT),  $\beta_{MKT}$  si sposta rispetto al  $\beta$  semplice.

## COSTRUZIONE DI PORTAFOGLIO (MEDIA–VARIANZA)

Lavoriamo su rendimenti mensili, con training sui primi 108 mesi e valutazione out-of-sample sul periodo successivo. Costruiamo due insiemi di rendimenti attesi:

1. **storici** (media campionaria con matrice di covarianza stimata),
2. **CAPM** (con  $r_f=4\%$  ed  $E[R_M] = 8\%$  annui, poi convertiti a mensile).

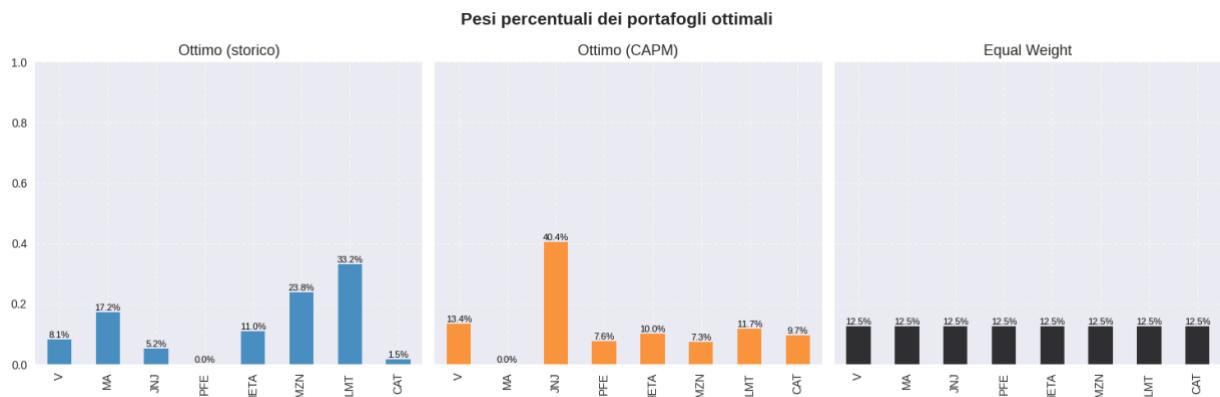
Ottimizziamo con vincoli long-only e somma pesi = 1, usando sia la soluzione analitica (GMV) e portafoglio tangente) sia una simulazione Monte Carlo per mappare la frontiera (cfr. figure “Frontiera (storiche)” e “Frontiera (CAPM)”).



**Cosa si vede sulle frontiere:**  
 La frontiera basata su attese storiche è più “aperta” e consente un punto tangente a rendimento atteso più alto; quella CAPM è più schiacciata per via del fatto che le attese sono simili tra i titoli e quindi l'aumento del premio a parità di rischio è più contenuto.

### Pesi dei portafogli ottimali (tangenti):

- Con le attese storiche l'ottimo si concentra su pochi titoli (in particolare LMT e AMZN, seguiti da MA), cioè i titoli che nei 108 mesi di training hanno dimostrato di avere il più alto premio/efficienza.
- Con le attese CAPM l'ottimo diventa più difensivo: il peso si sposta verso JNJ e, in generale, su profili di rischio inferiori. I relativi grafici a barre (tre pannelli) rendono immediato il confronto con il portafoglio Equal Weight(12,5% ciascuno).



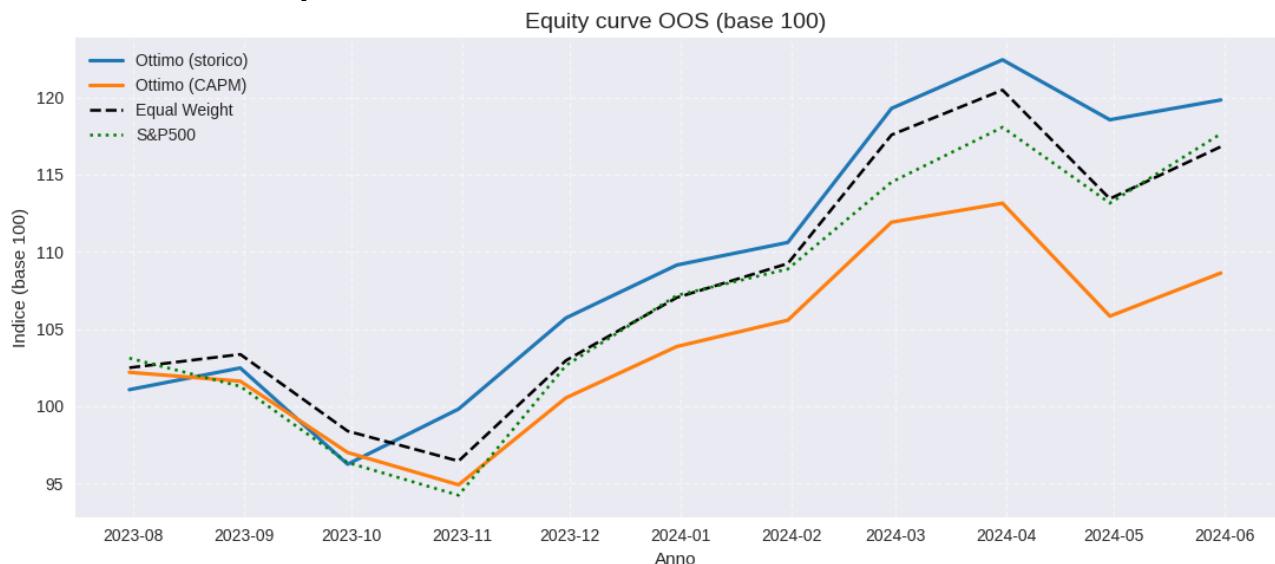
## Beta del portafoglio (vs S&P 500):

Calcoliamo il  $\beta$  in due modi: somma pesata dei  $\beta$  dei singoli titoli (training) e regressione dei rendimenti di portafoglio contro l'indice nel periodo OOS.

- **Somma pesata:** Storico ~0,93, CAPM ~0,80, Equal ~0,93.
- **Regressione OOS:** Storico ~0,69 (il più basso), CAPM ~0,88, Equal ~0,93.

L'ottimo storico, pur “contabile” vicino a 1, ex-post si comporta da beta più basso: seleziona combinazioni che hanno smorzato i movimenti dell'indice nel nostro OOS. L'Equal rimane vicino al mercato; l'ottimo CAPM rimane intermedio.

## Confronto out-of-sample



Nell'equity curve OOS (base 100) l'ottimo storico prevale, l'Equal Weight segue a breve distanza, mentre l'ottimo CAPM resta indietro. In sintesi:

- **Ottimo (storico):** rendimento annuo composto OOS più alto e Sharpe migliore ( $\approx 1,55$ ), con drawdown più contenuto rispetto al benchmark.
- **Equal Weight:** profilo molto competitivo (Sharpe circa 1,22), conferma la solidità del benchmark “semplice”.
- **Ottimo (CAPM):** più difensivo per costruzione, ma senza rendimento nel campione di test (Sharpe circa 0,72).
- **S&P 500:** market proxy, con Sharpe poco sotto 1,26 e drawdown più elevato di quello del miglior in-sample.

## Conclusioni

- Le attese storiche hanno premiato una costruzione più “aggressiva” ma efficiente: miglior rapporto rischio/rendimento nel nostro OOS e  $\beta$  effettivo inferiore.
- Le attese CAPM hanno prodotto pesi più cauti e una frontiera meno ripida; il risultato è più robusto ma meno redditizio nel nostro campione.
- Il portafoglio effettivo Equal Weight è un ottimo baseline: semplice, diversificato, vicino al mercato e a rendimento solido.

## DISCUSSIONE DEI RISULTATI

In sintesi, i rendimenti mensili mostrano bassa prevedibilità: gli ARIMA (anche auto-ARIMA) tendono alla media e faticano a cogliere shock e trimestrali. Le distribuzioni presentano code spesse e outlier (non perfettamente normali), mentre le correlazioni sono coerenti con il settore: molto elevate tra Visa–Mastercard, più instabili tra difensivi e tech. Nel confronto tra strategie, il Buy&Hold resta il più redditizio nel campione; la SMA 50/200 è l'alternativa più robusta sul profilo rischio/rendimento (volatilità e drawdown più bassi a fronte di un modesto sacrificio di CAGR); mean-reversion soffre, soprattutto in versione long/short; Random Forest offre spunti su alcuni growth ma non batte sistematicamente BH. Sul CAPM/Fama-French emergono chiaramente tilt growth (AMZN, META) e value (CAT, LMT), con beta di mercato più alti per i ciclici e più bassi per i difensivi. Nella costruzione di portafoglio, l'ottimo con attese storiche ha fornito il miglior OOS nel nostro periodo, l'Equal Weight si conferma un benchmark solido e vicino al mercato, mentre l'ottimo CAPM è più prudente ma meno performante. Operativamente: per massimizzare rendimento restano centrali esposizione al mercato e orizzonte lungo; per contenere rischio, regole semplici di trend-following (SMA) migliorano la traiettoria senza complicazioni eccessive.