

# Tesina Serie Storiche

Alessio Splendiani

September 2024

# 1 Stellantis Serie Storica

## 1.1 Introduzione

Il presente studio si propone di analizzare la serie storica delle quotazioni di **Stellantis**, conglomerato automobilistico nato dalla fusione tra *PSA Group* e *Fiat Chrysler Automobiles* (FCA) nel gennaio 2021. Sono stati presi come riferimento i dati giornalieri delle quotazioni che vanno dal 18 Gennaio 2021 al 28 Giugno 2024, e abbiamo tenuto conto in particolare, dei **valori di chiusura aggiustati**.

Il **valore di chiusura aggiustato** è il prezzo di chiusura di un'azione corretto per tenere conto di dividendi, frazionamenti azionari e altri eventi societari. Riflette il valore reale dell'azione, includendo l'impatto di queste operazioni straordinarie. Per modellare e prevedere l'evoluzione futura dei prezzi, abbiamo seguito l'approccio proposto da **Box e Jenkins**.

## 1.2 Analisi preliminare del dataset

Stiamo esaminando il seguente dataset *Stellantis*<sup>1</sup> preso da Yahoo Finance.

	Date	Open	High	Low	Close	Adj.Close	Volume
2	2021-01-19	13.18221	13.52386	13.05504	13.16702	10.167957	34981621
3	2021-01-20	13.44984	13.45743	13.17841	13.23725	10.222189	21595725
4	2021-01-21	13.34165	13.40618	12.95064	12.95064	10.000859	19585832
5	2021-01-22	12.85384	12.87472	12.44195	12.56533	9.703311	23123548
6	2021-01-25	12.66972	12.77032	11.88202	11.98072	9.251858	21308618
7	2021-01-26	11.95794	12.28821	11.92567	12.21608	9.433610	15668469
8	2021-01-27	12.33756	12.35274	11.71119	11.79471	9.108213	21156645
9	2021-01-28	11.67133	12.06993	11.35055	12.06993	9.320748	20647766
10	2021-01-29	11.83836	12.10030	11.68272	11.89151	9.182969	13953977
11	2021-02-01	12.10030	12.11928	11.80609	11.96364	9.238666	10479299
12	2021-02-02	12.09650	12.40020	12.07183	12.32807	9.520090	16271807
13	2021-02-03	12.62227	12.70199	12.27302	12.35085	9.537681	17417693
14	2021-02-04	12.40209	12.51028	12.27302	12.44195	9.608036	13898836
15	2021-02-05	12.45144	12.77412	12.34705	12.68301	9.794188	17565417
16	2021-02-08	12.81777	12.99620	12.73236	12.90699	9.967147	12954107

Figure 1: Stellantis Data

<sup>1</sup>Puoi accedere al dataset *Stellantis* tramite il seguente link: <https://yhoo.it/3uocUQy>.

Dopo aver presentato il dataset e la sua struttura, il passo successivo è visualizzare l'andamento temporale delle quotazioni di Stellantis attraverso il prezzo di chiusura aggiustato. In particolare ho trasformato la serie in forma logaritmica così da migliorare la comprensione del comportamento della serie.

```
1 #trasformazione logaritmica
2 LCloseStellantis<-log(Stellantis$Adj.Close)
3
4 #Grafico log AdjClose
5 plot(Stellantis$Date,LCloseStellantis,col=4,type = "l",main="Prezzo
6      di chiusura aggiustato 2021-2024",
       xlab="Date", ylab = "Log Adj Close")
```



Figure 2: Serie Log Adj Close Stellantis

Si può osservare un trend generalmente crescente, con alcune fluttuazioni, ed un picco significativo all'inizio del 2024, seguito da un calo più marcato. La presenza del trend indica che siamo in presenza di una serie non stazionaria.

Per procedere con la nostra analisi è necessario ricondurre la serie ad una situazione di stazionarietà debole, in cui si vanno a rispettare le seguenti condizioni:

- **Media costante:**

$$E[X_t] = \mu, \quad \forall t$$

dove  $X_t$  è la serie temporale e  $\mu$  è una costante.

- **Varianza costante:**

$$\text{Var}(X_t) = \sigma^2, \quad \forall t$$

dove  $\sigma^2$  è una costante.

- **Covarianza dipendente solo dal ritardo:**

$$\text{Cov}(X_t, X_{t+k}) = \gamma(k), \quad \forall t$$

dove  $\gamma(k)$  è la funzione di autocovarianza che dipende solo dal ritardo  $k$ , e non dal tempo  $t$ .

Analizziamo ora i grafici di ACF e PACF.

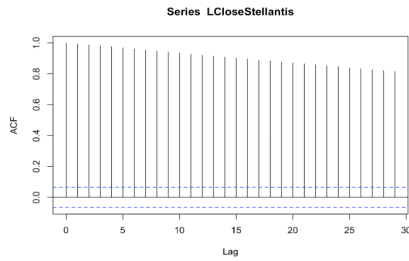


Figure 3: ACF

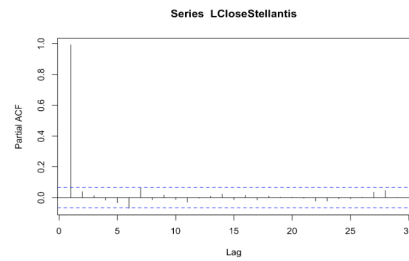


Figure 4: PACF

Il grafico dell'ACF mostra un andamento che decresce lentamente, ed il grafico della PACF ci porta ad ipotizzare che il processo abbia ordine 1 con coefficiente positivo.

### 1.3 Augmented Dickey-Fuller test

Per identificare la non stazionarietà della serie si effettua un **Augmented Dickey-Fuller test**.

```

1 #Test regression "none"
2 ADF_test1 <- ur.df(LCloseStellantis, type = c("none"), lags=10,
   selectlags = c("AIC"))
3 summary(ADF_test1)

```

Value of test-statistic is: 1.0399

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

Figure 5: Test regression "none"

Il test nella versione "none" serve a verificare la presenza di un **Random walk**. Confrontando il valore della statistica test (1.0399) con i valori critici, possiamo osservare che il valore della statistica test è superiore a tutti i valori critici (1%, 5%, e 10%). Di conseguenza **non possiamo rifiutare l'ipotesi nulla** e quindi siamo in presenza di un processo **random walk** e di una serie non stazionaria.

```

1 #Test regression "Drift"
2 ADF_test2 <- ur.df(LCloseStellantis, type = c("drift"), lags=10,
   selectlags = c("AIC"))
3 summary(ADF_test2)

```

Value of test-statistic is: -1.6415 2.0494

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.43	-2.86	-2.57
phi1	6.43	4.59	3.78

Figure 6: Test regression "drift"

Il test nella versione "drift" serve a verificare la presenza di un **Drift**. I risultati mostrano che per un livello di significatività fissato all'1% il test ci porta ad accettare l'ipotesi nulla sia per tau2 che per phi1. Quindi possiamo osservare la **presenza di un processo di Random walk** e **l'assenza di una componente di drift**.

```

1 #Test regression "Trend"
2 ADF_test3 <- ur.df(LCloseStellantis, type = c("trend"), lags=10,
   selectlags = c("AIC"))
3 summary(ADF_test3)

```

Value of test-statistic is: -2.0985 1.991 2.2836

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.96	-3.41	-3.12
phi2	6.09	4.68	4.03
phi3	8.27	6.25	5.34

Figure 7: Test regression "trend"

Il **test nella versione "trend"** serve a verificare la presenza di un **Trend**. Il test per un livello di significatività fissato all'1% ci porta ad accettare tutte le ipotesi nulle, quindi possiamo dire di essere in **presenza di un processo con Random Walk con esclusione sia della componente drift che del trend deterministico**.

## 1.4 Serie storica con Integrazione di ordine 1

Poiché è stato riscontrato che il processo segue un Random Walk, si procede con la differenziazione della serie con l'integrazione di ordine 1 per cercare di ottenere una serie stazionaria.

```

1 # Differenziazioni
2 Diff_Stellantis = diff(LCloseStellantis, lag=1)
3
4 Diff_Stellantis_ADF1 <- ur.df(Diff_Stellantis, type = c("none"),
   lags=10, selectlags = c("AIC"))
5 summary(Diff_Stellantis_ADF1)

```

Value of test-statistic is: -12.3751

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.58	-1.95	-1.62

Figure 8: Test regression diff. "none"

I risultati del **test "none"** dopo aver differenziato la serie di ordine 1, ci mostrano come ad un livello di significatività dell'1% il valore della statistica

test (-12.3751) è di molto inferiore al valore critico (-2.58), questo ci porta a rifiutare l'ipotesi nulla. **La serie dopo la differenziazione non presenta un Random walk.**

```
1 Diff_Stellantis_ADF2 <- ur.df(Diff_Stellantis, type = c("drift"),
2   lags=10, selectlags = c("AIC"))
3 summary(Diff_Stellantis_ADF2)
4 Diff_Stellantis_ADF3 <- ur.df(Diff_Stellantis, type = c("trend"),
5   lags=10, selectlags = c("AIC"))
6 summary(Diff_Stellantis_ADF3)
```

Value of test-statistic is: -12.4294 77.2505

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.43	-2.86	-2.57
phi1	6.43	4.59	3.78

Value of test-statistic is: -12.4276 51.4856 77.2232

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.96	-3.41	-3.12
phi2	6.09	4.68	4.03
phi3	8.27	6.25	5.34

Figure 9: Test regression diff. "drift"

Figure 10: Test regression diff. "trend"

I risultati sia del **test "none"** differenziato e sia del **test "trend"** differenziato, ci portano a rifiutare l'ipotesi nulla. Questo indica chiaramente che, dopo la differenziazione di ordine 1, la serie è diventata stazionaria. Possiamo ora osservare la nostra serie differenziata.

```
1 # Creare il grafico
2 plot(Stellantis$Date[-1], Diff_Stellantis, type="l", col="blue",
3   main="Prezzo di chiusura aggiustato 2021-2024",
4   xlab="Date", ylab="Adj Close")
5 par(mfrow = c(1, 2))
6 acf(Diff_Stellantis)
7 pacf(Diff_Stellantis)
```

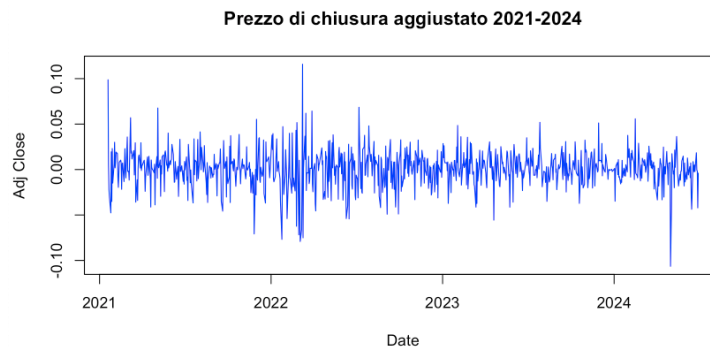


Figure 11: Serie Stellantis con Integrazione ordine 1

Possiamo osservare come la serie sia stata ricondotta ad una condizione di stazionarietà. Analizziamo ora i grafici di **ACF** e **PACF**.

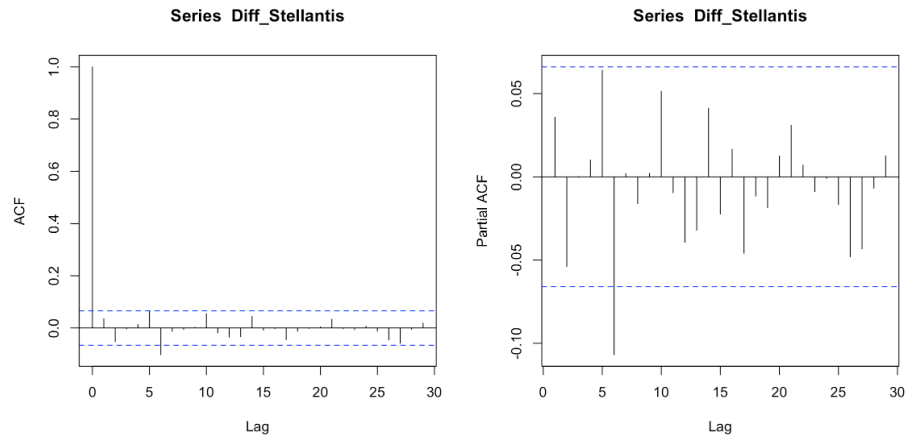


Figure 12: ACF e PACF

Nel **grafico dell'ACF** della serie differenziata, non emergono valori significativi oltre le bande di confidenza (rappresentate dalle linee tratteggiate blu) per la maggior parte dei lag. Tuttavia, sembra esserci un possibile valore significativo al lag 6. Questo indica che, per la maggior parte dei lag, non ci sono correlazioni significative tra la serie e i suoi lag, confermando che la serie differenziata ha raggiunto un livello di stazionarietà.

Nel **grafico della PACF**, si osserva una situazione simile: l'unico valore significativo è localizzato al lag 6. Ciò suggerisce che esiste una correlazione parziale residua tra i valori della serie a questo lag specifico, ma al di fuori di questo, non ci sono altre correlazioni significative.



## 2 Stima dei Modelli ARIMA e Confronto

L'obiettivo di questa analisi è identificare il modello ARIMA che meglio rappresenta l'evoluzione della serie storica differenziata di ordine 1. Sono stati stimati tre modelli **ARIMA** con **diversi ordini (p,d,q)**, ovvero **ARIMA(0,1,1)**, **ARIMA(1,1,0)** e **ARIMA(1,1,1)**. In questa sezione viene presentato un confronto tra i modelli basato sui seguenti criteri: valore dei parametri, valore dell'AIC, analisi dei residui, e test diagnostici di normalità e indipendenza.

```
1 #stimo il modello arima(0,1,1)
2 arima_model1 <- Arima(LCloseStellantis, order = c(0,1,1), include.
   constant = TRUE)
3 summary(arima_model1)
4 coeftest(arima_model1)
5
6 #stimo il modello arima(1,1,0)
7 arima_model2 <- Arima(LCloseStellantis, order = c(1,1,0), include.
   constant = TRUE)
8 summary(arima_model2)
9 coeftest(arima_model2)
10
11 #stimo il modello arima(1,1,1)
12 arima_model3 <- Arima(LCloseStellantis, order = c(1,1,1), include.
   constant = TRUE)
13 summary(arima_model3)
14 coeftest(arima_model3)
```

Le stime dei parametri per ciascun modello sono riportate di seguito, insieme ai relativi errori standard e valori di significatività statistica.

Modello	Parametro	Stima	Errore standard	p-value
ARIMA(0,1,1)	MA(1)	0.0411	0.0361	0.2546
	Drift	0.0008	0.0007	0.2688
ARIMA(1,1,0)	AR(1)	0.0366	0.0341	0.2827
	Drift	0.0008	0.0007	0.2659
ARIMA(1,1,1)	AR(1)	-0.7347	0.1171	<b>3.49e-10***</b>
	MA(1)	0.7873	0.1049	<b>5.99e-14***</b>
	Drift	0.0008	0.0007	0.2623

Table 1: Risultati del test z per i modelli ARIMA.

Dai risultati ottenuti, si può notare che il modello ARIMA(1,1,1) ha i parametri più significativi, indicando che probabilmente cattura meglio la struttura dei dati rispetto ai modelli ARIMA(0,1,1) e ARIMA(1,1,0).

Andiamo ora a fare il confronto anche in base all'**AIC**. L'**AIC** è una misura di qualità del modello, un valore basso indica un modello migliore.

Modello	AIC
ARIMA(1,1,1)	-4355.107
ARIMA(0,1,1)	-4353.181
ARIMA(1,1,0)	-4353.040

Table 2: Confronto dei modelli ARIMA sulla base dell'AIC.

Il modello **ARIMA(1,1,1)** è il migliore sia in termini di significatività dei parametri che di minimizzazione dell'AIC, ed è quindi il modello da preferire per la previsione della serie storica di Stellantis.

## 2.1 Analisi dei residui

Osserviamo come si distribuiscono i residui per i rispettivi modelli considerati.

```

1 #MODELLO ARIMA(0,1,1)
2 residui_arima_model1 <- arima_model1$residuals
3 plot(Stellantis$Date, residui_arima_model1, type="l", xlab="Date",
4       ylab="Residui", main="Grafico dei Residui ARIMA (0,1,1)")
5 abline(h=0, col= "red")

```

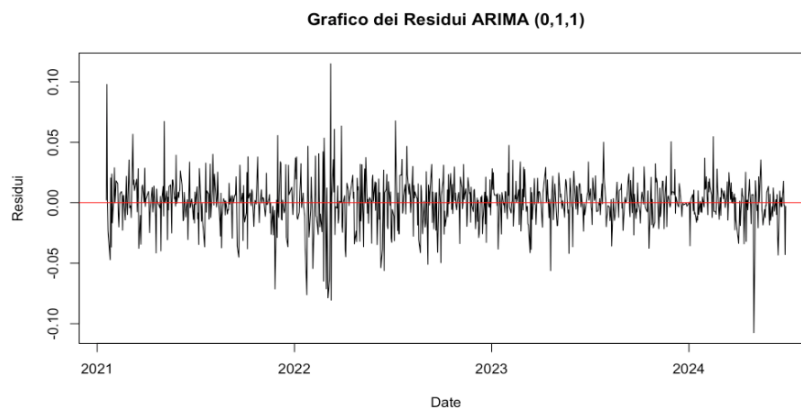


Figure 13: Residui ARIMA (0,1,1)

```

1 #Modello ARIMA (1,1,0)
2
3 residui_arima_model2 <- arima_model2$residuals
4 plot(Stellantis$Date, residui_arima_model2, type="l", xlab="Date",
5       ylab="Residui", main="Grafico dei Residui ARIMA (1,1,0)")
6 abline(h=0, col= "red")

```

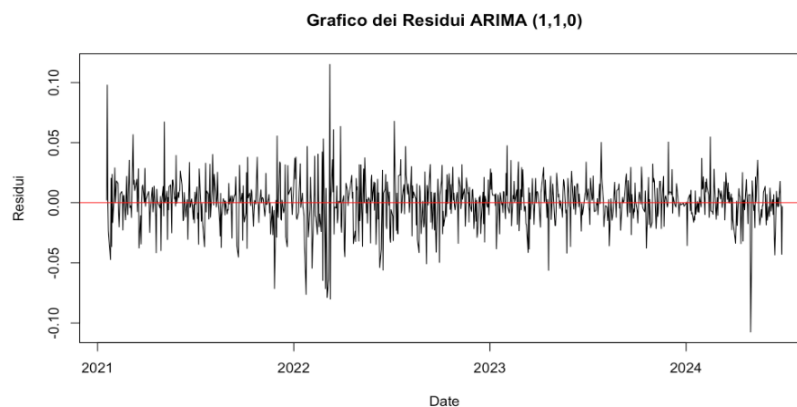


Figure 14: Residui ARIMA (1,1,0)

```

1 #MODELLO ARIMA(1,1,1)
2 par(mfrow = c(1, 1))
3
4 residui_arima_model3 <- arima_model3$residuals
5 plot(Stellantis$Date, residui_arima_model3, type="l", xlab="Date",
6      ylab="Residui", main="Grafico dei Residui ARIMA (1,1,1)")
7 abline(h=0, col= "red")

```

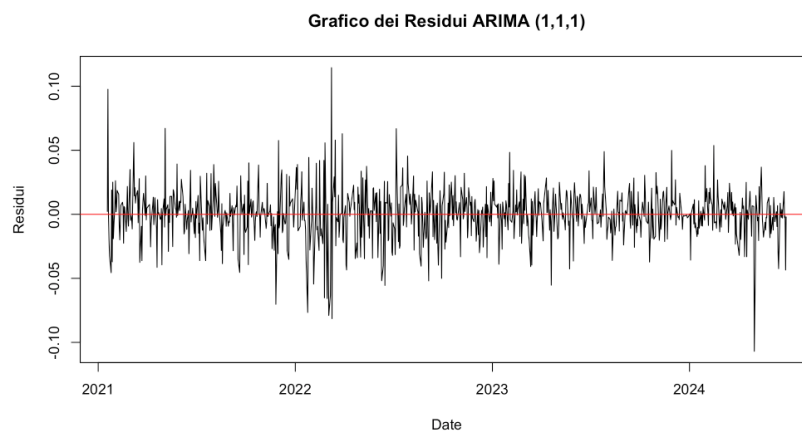


Figure 15: Residui ARIMA (1,1,1)

I grafici dei residui sembrerebbero essere stazionari, anche se è possibile osservare la presenza di alcuni valori anomali.

Andiamo a confrontare ora i grafici di ACF E PACF dei residui.

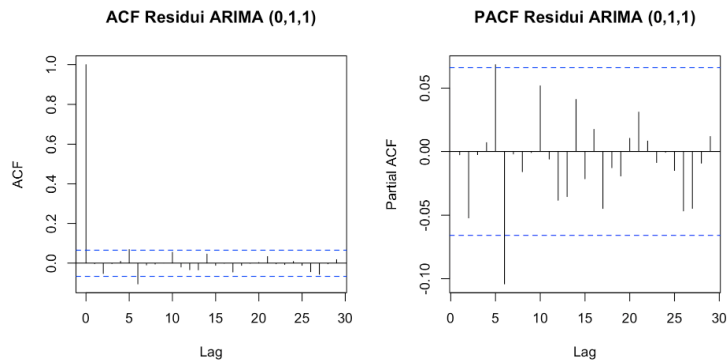


Figure 16: ARIMA(0,1,1)

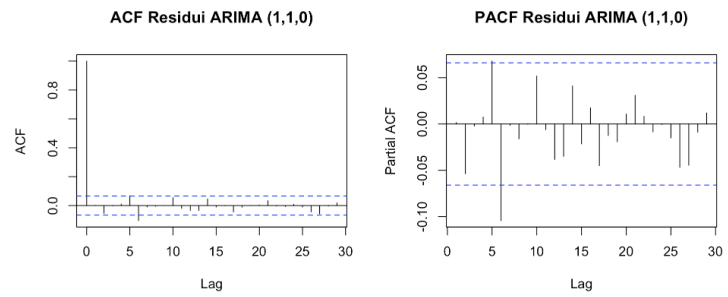


Figure 17: ARIMA(1,1,0)

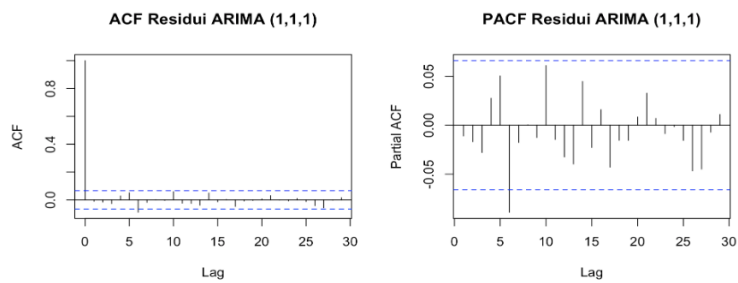


Figure 18: ARIMA(1,1,1)

Attraverso degli istogrammi sono andato a rappresentare anche la distribuzione dei residui.

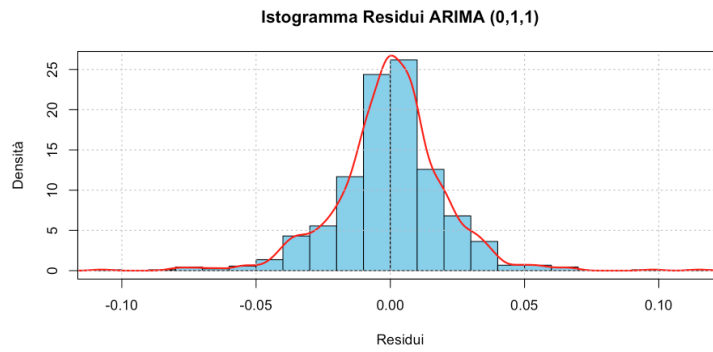


Figure 19: ARIMA(0,1,1)

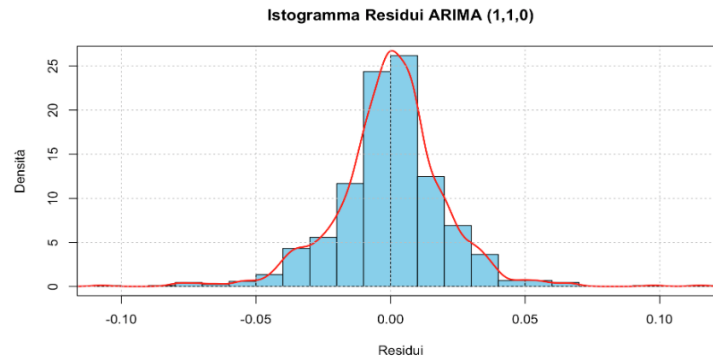


Figure 20: ARIMA(1,1,0)

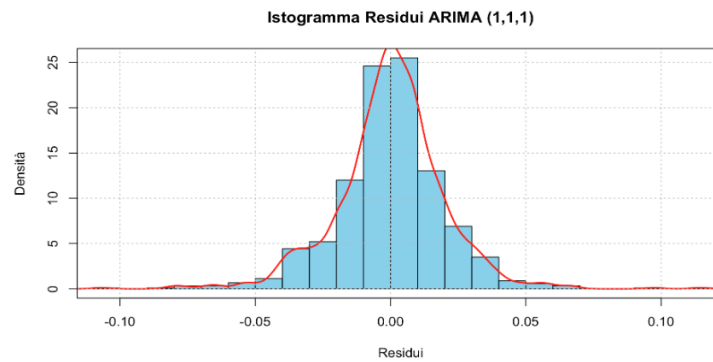


Figure 21: ARIMA(1,1,1)

Dai grafici notiamo che i residui non sono distribuiti normamente.

Table 3: Confronto tra i test di normalità e indipendenza dei residui per i modelli ARIMA

Test	ARIMA(0,1,1)	ARIMA(1,1,0)	ARIMA(1,1,1)
<b>Jarque-Bera</b> (Normalità)	$X^2 = 455.79$ $p < 2.2 \times 10^{-16}$	$X^2 = 456.33$ $p < 2.2 \times 10^{-16}$	$X^2 = 442.93$ $p < 2.2 \times 10^{-16}$
<b>Box-Pierce</b> (Lag 8)	$X^2 = 16.39$ $p = 0.0218$	$X^2 = 16.45$ $p = 0.0213$	$X^2 = 11.47$ $p = 0.1193$
<b>Ljung-Box</b> (Lag 8)	$X^2 = 16.52$ $p = 0.0208$	$X^2 = 16.58$ $p = 0.0203$	$X^2 = 11.57$ $p = 0.1156$
<b>Box-Pierce</b> (Res. quad. Lag 8)	$X^2 = 98.69$ $p < 2.2 \times 10^{-16}$	$X^2 = 98.50$ $p < 2.2 \times 10^{-16}$	$X^2 = 93.62$ $p < 2.2 \times 10^{-16}$
<b>Ljung-Box</b> (Res. quad. Lag 8)	$X^2 = 99.34$ $p < 2.2 \times 10^{-16}$	$X^2 = 99.14$ $p < 2.2 \times 10^{-16}$	$X^2 = 94.23$ $p < 2.2 \times 10^{-16}$

Nel **Test di Jarque-Bera (Normalità)** per tutti e tre i modelli ARIMA, il p-value è estremamente basso. Questo suggerisce il rifiuto dell'ipotesi di normalità dei residui. I residui non seguono una distribuzione normale.

Nei **Test di Box-Pierce e Ljung-Box** ricordando che sono utilizzati per verificare l'indipendenza dei residui, possiamo vedere che per **ARIMA(0,1,1)** e **ARIMA(1,1,0)**, i p-value sono inferiori a 0.05 (circa 0.02), indicando la presenza di autocorrelazione significativa nei residui a lag 8. Per **ARIMA(1,1,1)**, i p-value sono superiori a 0.05 (circa 0.11-0.12), suggerendo che non c'è evidenza significativa di autocorrelazione nei residui a lag 8. Questo indica che il modello **ARIMA(1,1,1)** sembra performare meglio in termini di indipendenza dei residui rispetto agli altri due modelli.

Infine nei **Test di Box-Pierce e Ljung-Box** sui residui al quadrato possiamo osservare che per tutti i modelli, i p-value sono estremamente bassi, indicando una forte evidenza di effetti ARCH (eteroschedasticità condizionale autoregressiva) nei residui.

Tenendo conto di tutte le analisi svolte, dai risultati forniti anche dall'AIC possiamo concludere che il modello che rappresenta meglio i nostri dati è il modello **ARIMA (1,1,1)**.

### 3 Previsioni del modello ARIMA

Mostriamo le previsioni ottenute dal modello ARIMA(1,1,1) per la serie storica per i valori di chiusura aggiustati delle azioni Stellantis. Il modello è stato stimato sui primi 794 che corrispondono al nostro **training set**, escludendo gli ultimi 88 dati (**test set**).

```
1 #Previsione stimo togliendo gli ultimi 88 dati (10% campione
  2021-2024)
2 model.prevl<- Arima(window(LCloseStellantis,end=794),order=c(1,1,1)
  , include.constant = TRUE)
3 summary(model.prevl)
4 coeftest(model.prevl)
```

```
Series: window(LCloseStellantis, end = 794)
ARIMA(1,1,1) with drift
```

Coefficients:

	ar1	ma1	drift
	-0.7294	0.7764	0.0011
s.e.	0.1322	0.1208	0.0007

```
sigma^2 = 0.0004211: log likelihood = 1958.16
AIC=-3908.33 AICc=-3908.27 BIC=-3889.62
```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	1.240921e-05	0.02046829	0.01472264	-0.0038647	0.5779175	0.9933631	-0.01551035

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )
ar1	-0.7293910	0.1322070	-5.5170	3.448e-08 ***
ma1	0.7763955	0.1207769	6.4283	1.290e-10 ***
drift	0.0011265	0.0007479	1.5062	0.132

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dai risultati possiamo osservare che:

- L'errore quadratico medio (**RMSE**) sul set di training è di 0.02, indicando una discreta precisione delle previsioni nel nostro campione.
- Il **MAPE** (Mean Absolute Percentage Error) di 0.58% suggerisce che, in media, le previsioni del modello si discostano dello 0.58% dai valori effettivi nel set di training.
- **L'ACF1** (autocorrelazione al primo lag dei residui) è molto vicina a zero (-0.01), indicando che il modello ha catturato bene la struttura di autocorrelazione della serie.

A seguire si può osservare la previsione sul test set.

```

1 # Previsione a 88 passi
2 forecast1 <- forecast(model.prev1,h=88)
3 par(mfrow = c(1, 1))
4 plot(forecast(model.prev1, h = 88),main="Previsioni con Modello
   ARIMA (1,1,1)")
5 print(forecast(model.prev1,h=88))
6 lines(LCloseStellantis)
7 # Aggiungi una linea verticale rossa per separare le osservazioni
   dalle previsioni
8 abline(v = length(LCloseStellantis) - 88, col = "red", lty = "
   dashed")

```

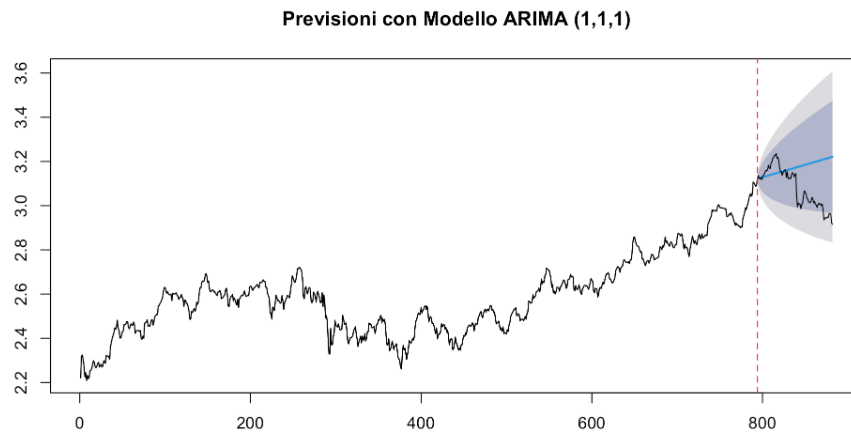


Figure 22: Previsione modello ARIMA(1,1,1)

Il modello non sembra prevedere in modo ottimale i dati probabilmente a causa dell'eteroschedasticità dello stesso. Andiamo a valutare l'accuracy del nostro **test set**.

```

1 # out-of-sample multi-step forecasts
2 accuracy(forecast(model.prev1,h=88),
3          (window(LCloseStellantis,start=795)))

```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	1.240921e-05	0.02046829	0.01472264	-0.0038647	0.5779175	0.9933631	-0.01551035	NA
Test set	-9.080272e-02	0.14593802	0.11296543	-3.0554990	3.7477075	7.6219799	0.96121499	7.943279

Figure 23: Accuracy Test

Analizzando questi risultati, possiamo trarre le seguenti conclusioni:

- **Errore Medio (ME):** Il valore negativo (-0.09) indica che il modello tende a sottostimare leggermente i valori reali nel set di test.



- **Errore Quadratico Medio (RMSE)**: Il valore di 0.14 per il set di test è significativamente più alto rispetto al set di training (0.020), indicando una diminuzione dell'accuratezza nelle previsioni out-of-sample.
- **Errore Percentuale Medio (MPE)**: Il valore negativo (-3.05%) conferma la tendenza del modello a sottostimare i valori reali.
- **Errore Percentuale Assoluto Medio (MAPE)**: Il MAPE del 3.75% sul set di test è significativamente più alto rispetto allo 0.58% del set di training, indicando una perdita di precisione nelle previsioni out-of-sample.
- **ACF1**: L'alto valore di autocorrelazione al primo lag (0.96) nei residui del set di test indica che il modello non ha catturato completamente la struttura di dipendenza temporale nei dati out-of-sample.

L'analisi delle previsioni del modello ARIMA(1,1,1) per il prezzo di chiusura delle azioni Stellantis ha rivelato che il modello ha mostrato una performance significativamente peggiore sul test set rispetto al training. Inoltre, le previsioni tendono a sottostimare i valori reali. Per migliorare il nostro modello possiamo esplorare l'utilizzo di modelli più sofisticati come **ARCH** e **GARCH**.

## 4 Modelli ARCH e GARCH

I modelli **ARCH** e **GARCH** sono utilizzati per modellare e prevedere la volatilità dei dati finanziari, in particolare nelle serie temporali economiche, come i rendimenti delle azioni. Questi modelli sono importanti perché tengono conto dell'eteroschedasticità condizionale, ovvero delle variazioni non costanti nella volatilità nel tempo, un fenomeno comune nei mercati finanziari.

- **ARCH**: Utilizza una struttura autoregressiva per modellare la volatilità come una funzione quadratica dei residui passati. In altre parole, la varianza del processo in un dato periodo dipende dai quadrati dei residui dei periodi precedenti.
- **GARCH**: Estende il modello ARCH aggiungendo una componente autoregressiva nella varianza condizionata. La varianza condizionata in un periodo dipende non solo dai quadrati dei residui passati (come nel modello ARCH) ma anche dalle varianze passate, migliorando la capacità del modello di catturare la persistenza della volatilità.

```

1 # Modello ARCH
2 r2 <- Diff_Stellantis^2
3 archmodel <- arima(x = r2, order = c(1, 0, 0), method = "CSS")
4
5 summary(archmodel)
6 coeftest(archmodel)

```

```

Call:
arima(x = r2, order = c(1, 0, 0), method = "CSS")

Coefficients:
    ar1  intercept
    0.1404    4e-04
s.e.  0.0315    0e+00

sigma^2 estimated as 8.332e-07: part log likelihood = 4916.04

Training set error measures:
      ME      RMSE      MAE  MPE MAPE  MASE  ACF1
Training set -1.168816e-09 0.0009122707 0.0004607025 -Inf  Inf 0.8035193 0.01179303

z test of coefficients:

              Estimate Std. Error z value Pr(>|z|)
ar1      1.4037e-01 3.1541e-02  4.4504  8.57e-06 ***
intercept 4.0440e-04 4.9146e-05  8.2285 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figure 24: Summary ARCH

Figure 25: Coeftest ARCH

I valori di errore del training set, come **RMSE** (0.000912) e **MAE** (0.000461), sono piuttosto piccoli, indicando che il modello si adatta bene ai dati. L'**ACF1** vicino a zero (0.0118) suggerisce che i residui del modello non mostrano autocorrelazione significativa, confermando una buona specificazione del modello. Il test z mostra che il coefficiente ar1 è altamente significativo con  $p < 0.001$ , evidenziando l'eteroschedasticità e la relazione con i rendimenti passati.

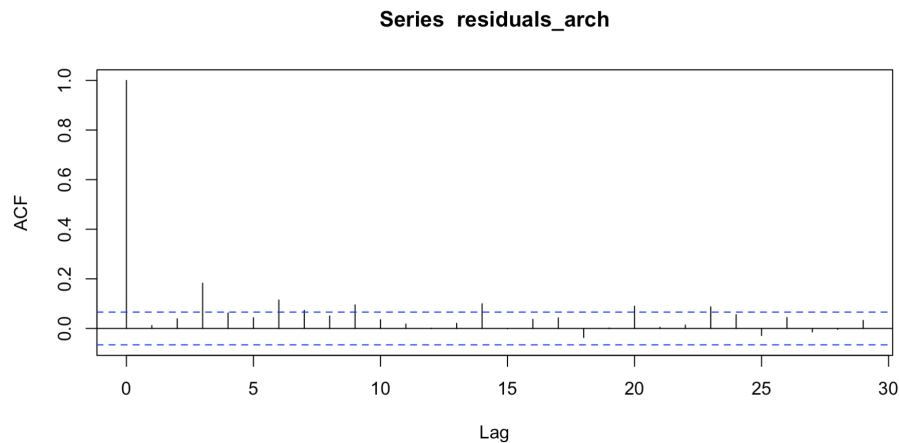


Figure 26: ACF ARCH

Questi risultati indicano che il modello non è riuscito a catturare tutta la struttura di dipendenza nei dati. Questo è un segnale che il modello ARCH, da solo, potrebbe non essere sufficiente per descrivere correttamente la volatilità condizionata della serie temporale.

In particolare, il fatto che ci sia ancora una certa autocorrelazione residua può giustificare il passaggio a un modello **GARCH**, che include anche un termine autoregressivo per la varianza condizionata, migliorando la capacità del modello di catturare la persistenza della volatilità.

```

Call:
garch(x = Diff_Stellantis, order = c(1, 1, 1))

```

Model:

GARCH(1,1)

Residuals:

Min	1Q	Median	3Q	Max
-6.15329	-0.47072	0.06332	0.57252	3.77888

Coefficient(s):

	Estimate	Std. Error	t value	Pr(> t )
a0	1.054e-05	4.494e-06	2.345	0.019 *
a1	2.880e-02	7.173e-03	4.015	5.95e-05 ***
b1	9.435e-01	1.772e-02	53.247	< 2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Diagnostic Tests:

Jarque Bera Test

data: Residuals

X-squared = 298.66, df = 2, p-value < 2.2e-16

Box-Ljung test

data: Squared.Residuals

X-squared = 0.60996, df = 1, p-value = 0.4348

> coeftest(GARCH\_1)

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )
a0	1.0537e-05	4.4938e-06	2.3448	0.01904 *
a1	2.8798e-02	7.1733e-03	4.0146	5.954e-05 ***
b1	9.4351e-01	1.7720e-02	53.2467	< 2.2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

I risultati del modello **GARCH(1,1)** mostrano quanto segue:

- **Coefficiente a0:** È significativo con un valore di  $p = 0.019$ , indicando che c'è una componente di varianza fissa nel modello.
- **Coefficiente a1:** (autoregressivo della volatilità): Il valore è 0.0288, altamente significativo, il che conferma che la volatilità passata ha un impatto sulla volatilità corrente.
- **Coefficiente b1:** Molto alto, pari a 0.9435, e fortemente significativo, suggerendo che c'è una forte persistenza della volatilità nel tempo.

- **Jarque-Bera:** Indica che i residui non seguono una distribuzione normale.
- **Box-Ljung** sui residui al quadrato: Con un  $p = 0.4348$ , non si rileva autocorrelazione significativa nei residui al quadrato.

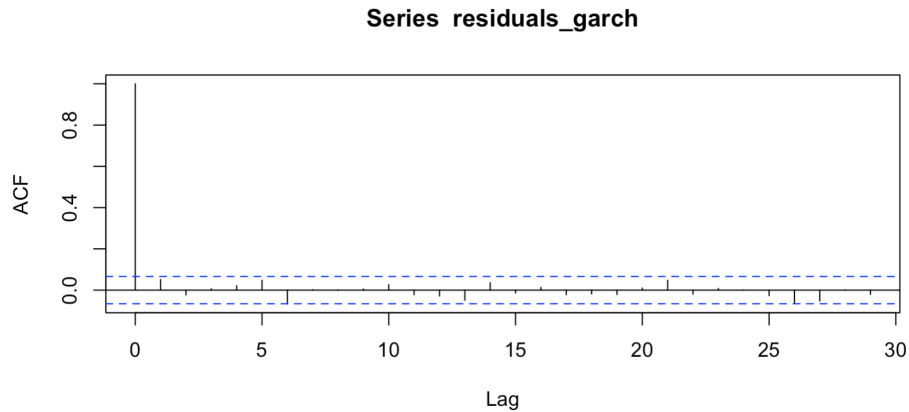


Figure 27: ACF GARCH

Il grafico **ACF** dei residui del modello **GARCH** indica che non c'è autocorrelazione significativa tra i residui, suggerendo che il modello GARCH ha catturato efficacemente la volatilità del modello.

```
> #Test di Ljung-Box e Box-Pierce per verificare l'autocorrelazione dei residui al quadrato
> Box.test(residuals_garch^2, lag = 4, type = c("Ljung-Box"), fitdf = 2)
```

Box-Ljung test

```
data: residuals_garch^2
X-squared = 2.4905, df = 2, p-value = 0.2879
```

```
> Box.test(residuals_garch^2, lag = 4, type = c("Box-Pierce"), fitdf = 2)
```

Box-Pierce test

```
data: residuals_garch^2
X-squared = 2.4775, df = 2, p-value = 0.2897
```

Figure 28: Box-Ljung test - Box-Pierce test

Applicando i test di **Box-Pierce** e **Ljung-Box** sui residui al quadrato del modello GARCH, possiamo osservare che risultano non essere significativi e quindi ci permettono di concludere che il modello stimato non contenga eteroschedasticità.

## 5 Conclusioni

In questo studio, abbiamo analizzato la serie storica dei prezzi di chiusura giornalieri di **Stellantis** dal 2021 al 2024, applicando diverse tecniche di modellizzazione. L'approccio seguito è stato quello classico di **Box-Jenkins**, che ci ha permesso di identificare il miglior modello **ARIMA** per descrivere l'andamento dei dati. Tra i modelli ARIMA considerati (**ARIMA(1,1,1)**, **ARIMA(0,1,1)** e **ARIMA(1,1,0)**), **il modello ARIMA(1,1,1) è risultato essere il più appropriato**, grazie ai suoi migliori risultati in termini di adattamento e significatività dei parametri.

Tuttavia, l'analisi non può fermarsi alla semplice previsione, che a causa dell'eteroschedasticità presente nel modello non si è dimostrata efficace. Questo ci ha portato ad approfondire l'analisi introducendo i modelli **ARCH** e **GARCH**, per modellare l'eteroschedasticità della serie.

L'applicazione del modello **ARCH** ha evidenziato alcune problematiche: l'ACF dei residui del modello mostrava ancora una leggera autocorrelazione, suggerendo che la dipendenza nella volatilità non fosse completamente catturata. Di conseguenza, abbiamo esteso l'analisi utilizzando il modello **GARCH(1,1)**, che si è dimostrato più flessibile e adatto a rappresentare la persistenza nella varianza.

**L'ACF** dei residui del modello **GARCH** ha mostrato una riduzione significativa dell'autocorrelazione rispetto al modello **ARCH**, confermando che il **GARCH è più efficace** nel catturare la dinamica della volatilità dei prezzi di Stellantis.