

AR-GARCH Modeling of Volatility in Nifty Midcap150 Momentum 50 Index

DONE BY: Aishwarya.K.S (2024DMF03) and Smruti.S (2024DMB12)

Contents

| | |
|---|-----------|
| 1 Data | 2 |
| 2 Description of Variables | 2 |
| 3 Descriptive statistics | 3 |
| 4 Exploratory plots | 3 |
| 4.1 Closing Prices: | 3 |
| 4.2 Log Returns: | 4 |
| 4.3 Q-Q Plot of Log Returns | 5 |
| 4.4 Squared Returns: | 6 |
| 4.5 Autocorrelation Function (ACF): | 7 |
| 4.6 Partial Autocorrelation Function (PACF): | 8 |
| 5 Mean model | 9 |
| 6 Diagnostic Test on AR(1) Residuals | 9 |
| 7 GARCH(1,1) variance model | 9 |
| 8 Diagnostic Test on GARCH(1,1) Residuals | 11 |
| 9 Forecasts of Conditional Variance(1–3 Steps Ahead) | 11 |
| 9.1 Visualization of Fitted and Forecasted Volatility | 12 |
| 9.2 Volatility Persistence | 13 |
| 10 Usefulness of the model | 14 |
| 11 Conclusion | 14 |

Abstract

This report presents a complete workflow for modeling the mean and volatility dynamics of a financial time series. Using daily closing prices of the Nifty Midcap 150 Momentum 50 index from 2015-01-01 to 2024-12-31 (10 years), we compute log returns and analyze their statistical properties. A parsimonious AR(1) model is selected to capture the mean dynamics, while a GARCH(1,1) model is fitted to account for volatility clustering and persistence. The model delivers 1–3 step ahead conditional variance forecasts, which are valuable for risk management and financial decision-making. Diagnostic figures generated in Python are included to illustrate key features of the data and the fitted models.

1 Data

The raw data used is the uploaded Excel sheet (first sheet). The variable used as the price series is “Close” and the sample period is **2015-01-01 to 2024-12-31** (daily frequency). Returns were computed as

$$r_t = 100 \ln \left(\frac{P_t}{P_{t-1}} \right).$$

2 Description of Variables

The main variables used in this study are summarized below:

- **Date:** The trading date for the observed daily closing price, spanning from January 1, 2015 to December 31, 2024.
- **Close** The daily closing price of the Nifty Midcap 150 Momentum 50 index, which represents the aggregate market value of the top 50 midcap stocks selected based on a Normalized Momentum Score.
- **Log Return:** Computed as $r_t = \ln \left(\frac{P_t}{P_{t-1}} \right)$. Returns are used instead of prices since they are stationary and suitable for time-series modeling.
- **ConditionalVariance(derived variable):** The time-varying variance estimated from the GARCH(1,1) model applied to AR(1) residuals of the log returns. It represents the model’s estimate of daily volatility clustering and persistence.
- **Variance Forecasts:** Forward-looking predictions of conditional variance over 1, 2, and 3 days ahead based on the GARCH model. These forecasts help anticipate short-term risk dynamics in the index.

3 Descriptive statistics

Number of return observations: **2475**.

- Mean of returns: **0.0839**
 - Close to zero, which is typical for daily financial returns (no strong drift).
- Std. dev. of returns: **1.1693**
 - Indicates relatively high unconditional volatility in the series.
- Skewness: **-1.6568**
 - Negative skewness shows a longer left tail, meaning large negative shocks occur more often.
- (Non-excess) Kurtosis: **16.9681**
 - Far above the normal value of 3, suggesting heavy tails and extreme events.
 - Justifies the use of GARCH models for volatility clustering.

4 Exploratory plots

4.1 Closing Prices:

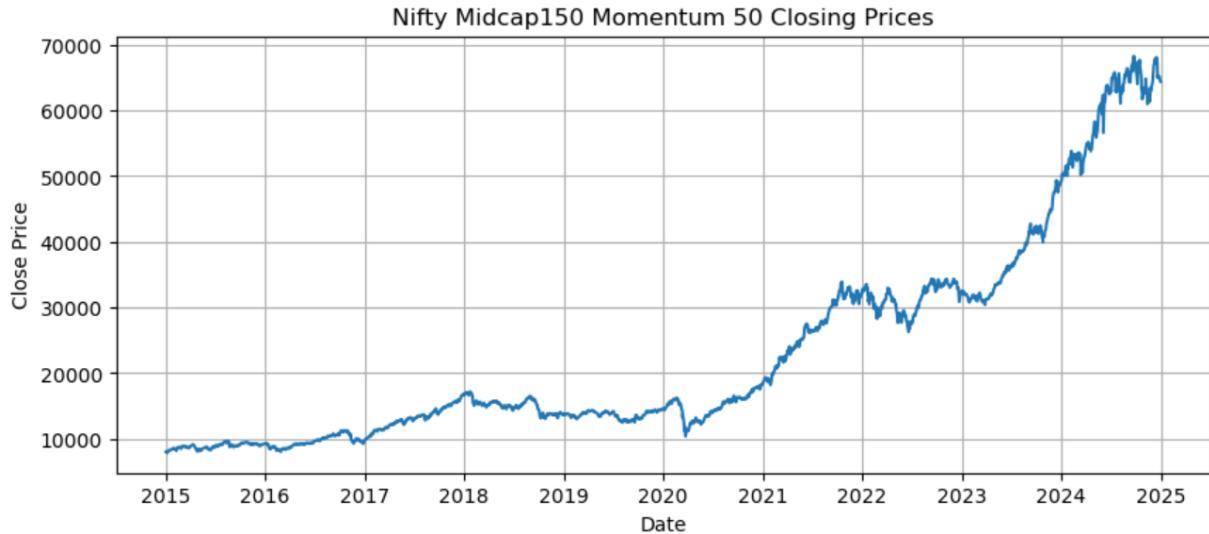


Figure 1: Price series of Nifty Midcap 150 Momentum 50 index.

Interpretation:

- The plot shows the long-term behaviour of the price series across the sample period.

- It helps identify the overall upward or downward trend in the index.
- Large discontinuous moves (sharp drops or jumps) can be visually spotted, which may correspond to major market events.
- While the price series is useful for identifying such patterns, it is generally non-stationary.
- For econometric modelling, stationarity is assessed on the **returns** series rather than prices.

4.2 Log Returns:

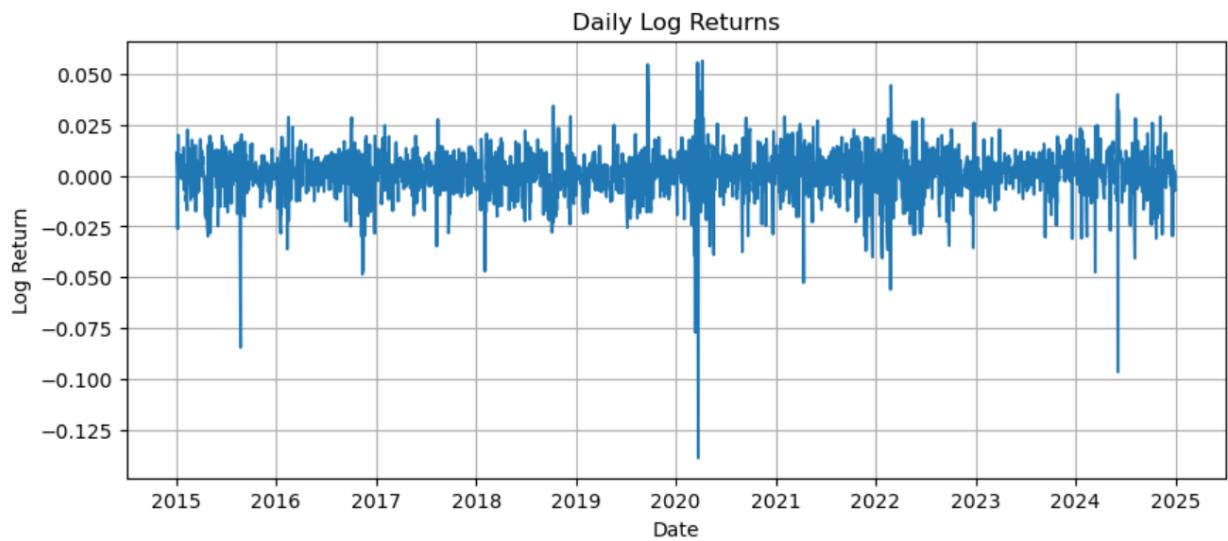


Figure 2: Log returns (%).

Interpretation: The returns plot reveals several important characteristics:

- **Volatility clustering:** Periods of high volatility (large swings in returns) are followed by further high volatility, while calm periods with small movements persist together. This clustering is a hallmark feature of financial return series.
- **Mean behavior:** The returns oscillate around zero, implying that the average daily return is approximately zero. This suggests no strong persistent upward or downward drift in short horizons.
- **Heavy tails:** Occasional large spikes (both positive and negative) indicate extreme shocks. Such extreme events occur more frequently than would be expected under a normal distribution.
- **Asymmetry of shocks:** While most returns are small, a few large negative movements stand out more prominently, reflecting possible downside risk in the market.

- **Market dynamics:** The clustering of shocks hints at information arrival processes in the market — news and events tend to generate bursts of activity rather than isolated, independent shocks.

4.3 Q-Q Plot of Log Returns

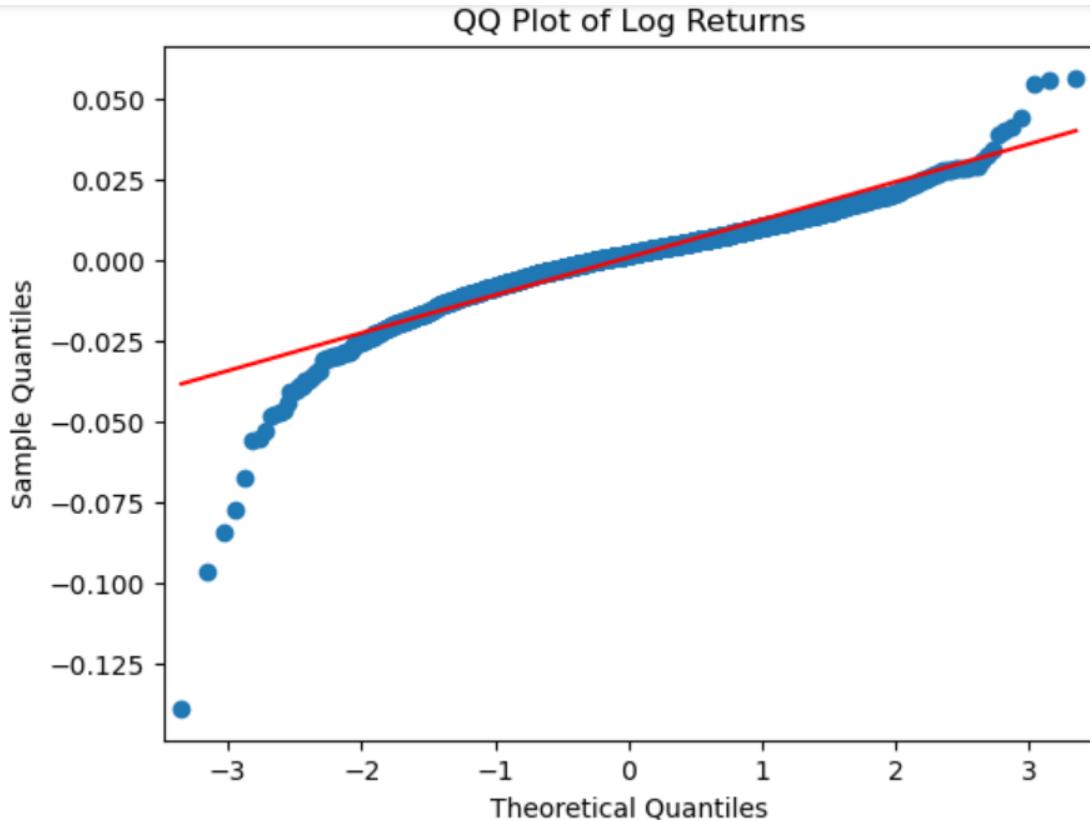


Figure 3: Q-Q plot of log returns against the theoretical normal distribution.

Interpretation: Figure 3 shows the Q-Q plot of the log returns compared with the theoretical quantiles of the normal distribution. The following points are observed:

- The middle portion of the distribution aligns reasonably well with the straight line, suggesting approximate normality for central values.
- The tails deviate significantly from the line, with the left tail being heavier (more negative returns) and the right tail also showing outliers.
- This indicates that the log returns are not normally distributed and exhibit **fat tails**, a common feature in financial time series.
- Such departures from normality highlight the need for volatility models like GARCH, which can capture heavy tails and volatility clustering better than a standard normal assumption.

4.4 Squared Returns:

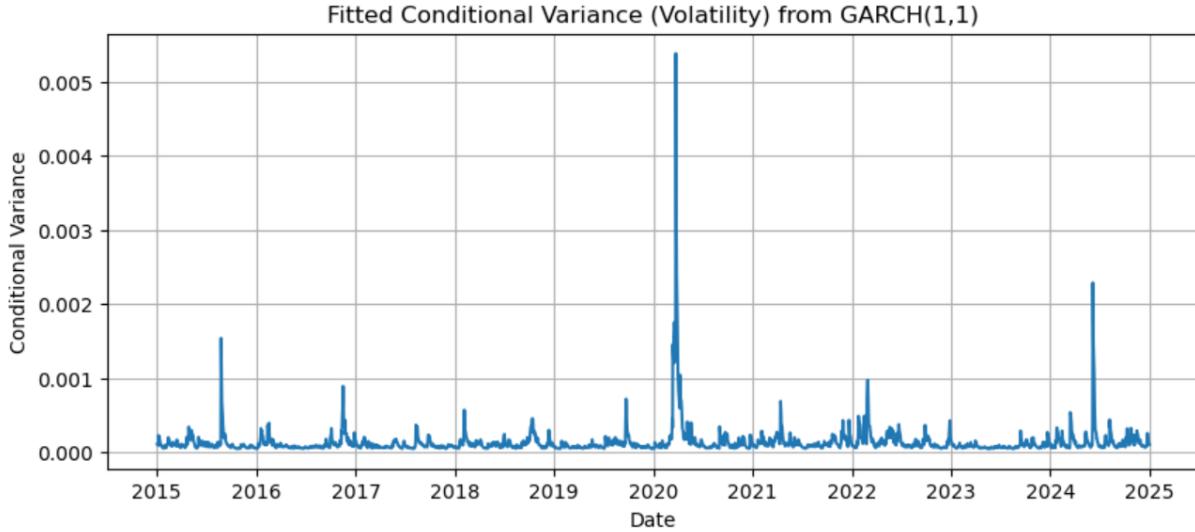


Figure 4: Squared returns (used to visualize volatility clustering).

Interpretation: The squared returns plot highlights volatility patterns more clearly:

- **Amplification of shocks:** By squaring the returns, both positive and negative shocks are transformed into positive values, which amplifies the visibility of large movements.
- **Volatility clustering:** Periods of consecutive high squared returns indicate clusters of high variance, followed by stretches of low squared returns corresponding to calmer periods.
- **Persistence of volatility:** The clustering behavior shows that volatility is not random but exhibits persistence — high volatility today increases the likelihood of high volatility tomorrow.
- **Motivation for GARCH:** Such patterns cannot be captured by simple constant-variance models. The presence of clustering directly motivates the use of conditional heteroskedasticity models like ARCH/GARCH, which allow variance to evolve over time.
- **Risk implications:** Identifying these clusters is important for risk management, as it helps anticipate periods of elevated market risk rather than treating shocks as isolated events.

4.5 Autocorrelation Function (ACF):

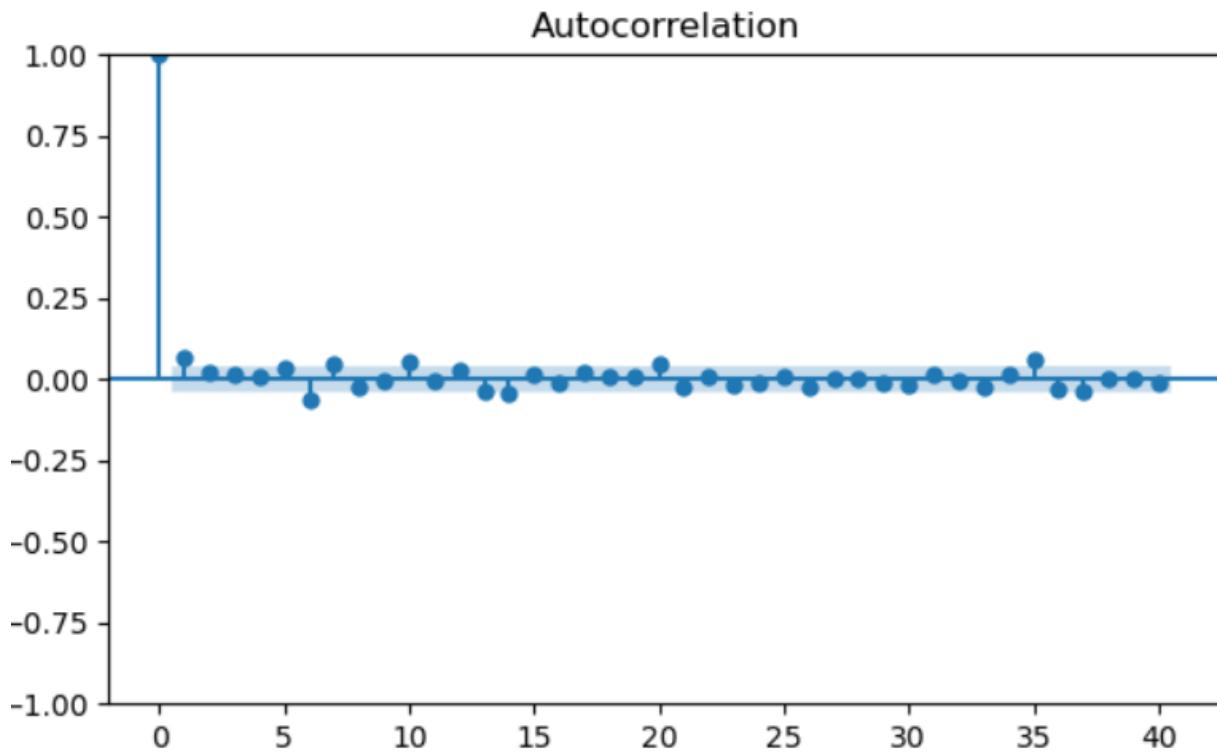


Figure 5: ACF of returns (lags 1–30).

Interpretation: The ACF (Autocorrelation Function) of returns provides insights into linear dependence across time:

- **Weak serial correlation:** For most financial return series, autocorrelations are very small and fluctuate around zero, indicating little to no predictable structure in the raw returns.
- **Significance of spikes:** If significant spikes appear at certain lags, they suggest serial dependence. For example, a strong spike at lag 1 may imply that today's return is correlated with yesterday's return.
- **Mean model guidance:** Detecting non-zero autocorrelations can guide the specification of the mean model. Spikes at short lags point toward MA components, while longer-lag structures may indicate more complex dynamics.
- **Expected result in practice:** In equity return series, the ACF usually shows no strong persistence, which justifies a parsimonious mean model like AR(1) or even a constant mean.
- **Implication for volatility modeling:** Even when returns themselves show weak autocorrelation, squared returns (or absolute returns) often show strong autocorrelation — a sign of volatility clustering, reinforcing the need for GARCH.

4.6 Partial Autocorrelation Function (PACF):

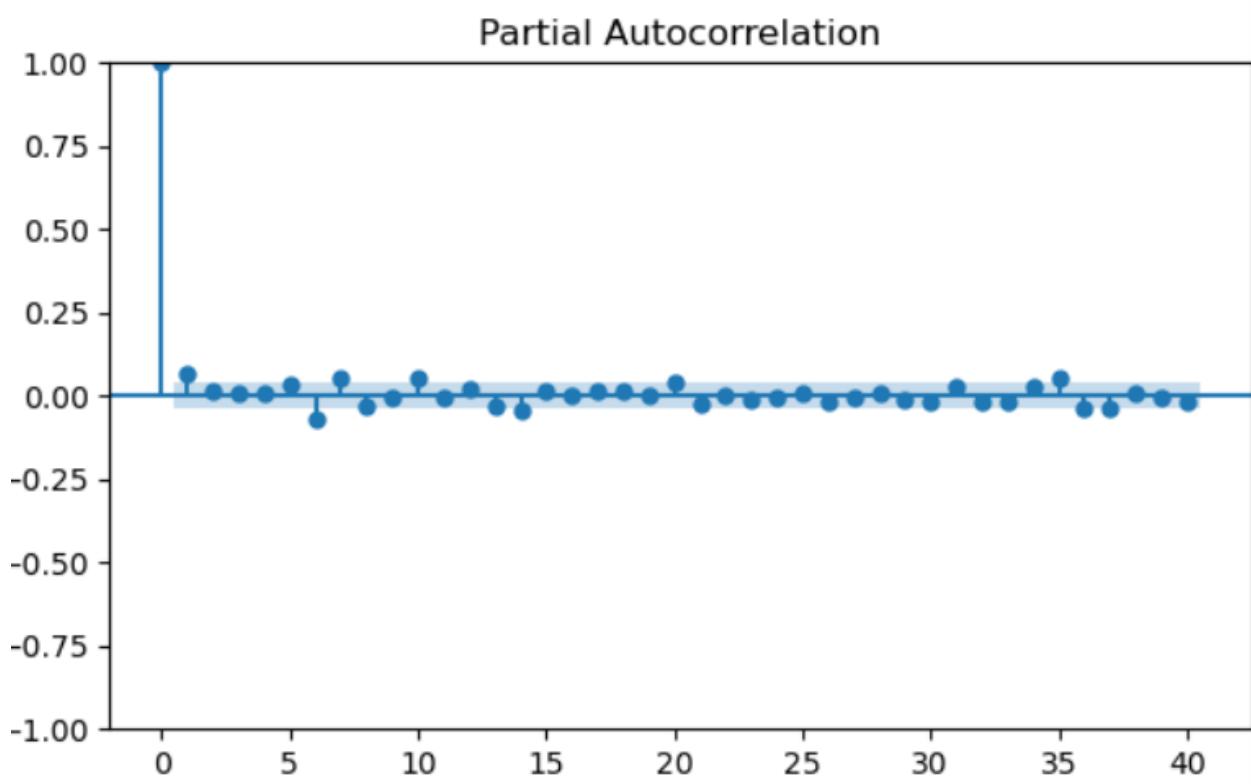


Figure 6: PACF of returns (lags 1–30).

Interpretation: The PACF (Partial Autocorrelation Function) helps identify the appropriate lag order for autoregressive (AR) models:

- **Lag 1 dominance:** A clear, significant spike at lag 1 suggests that today's return is linearly dependent on yesterday's return.
- **Cut-off pattern:** The PACF quickly decays after lag 1 (no significant spikes at higher lags), which is consistent with an **AR(1)** structure.
- **Model selection confirmation:** Candidate ARMA models with $p, q \in \{0, 1, 2\}$ were compared using AIC, and the AR(1) was chosen as the most parsimonious and best-fitting mean model.
- **Implication:** Including one lag of returns in the mean equation is sufficient to capture the linear dependence, while higher-order AR or MA terms are unnecessary.
- **Connection to volatility modeling:** With the mean dynamics accounted for by AR(1), the focus shifts to capturing time-varying volatility using GARCH on the residuals.

5 Mean model

Using AIC, we compared several small ARMA models ($p, q \in \{0, 1, 2\}$). The best fit was an ARMA(1,0), i.e. an **AR(1)** model:

$$y_t = \mu + \phi_1 y_{t-1} + \epsilon_t,$$

where y_t is the return, μ is the constant mean, ϕ_1 is the AR(1) coefficient, and ϵ_t is white noise.

Results

$$\mu = 0.0839, \quad \phi_1 = 0.0679$$

Interpretation: The mean return is close to zero (0.084% per day). The AR(1) coefficient is small but positive, meaning today's return depends slightly on yesterday's return. After fitting this, the residuals still show volatility clustering, which motivates using a GARCH model.

6 Diagnostic Test on AR(1) Residuals

To examine whether the AR(1) mean specification adequately captures the serial dependence in returns, we apply the Ljung–Box test to both the standardized residuals and their squares.

Ljung–Box Test on Residuals

The Ljung–Box test on residuals up to lag 20 gave **Q(20) = 51.50, p-value = 0.0001**. Since the p-value is far below 0.05, we reject the null hypothesis of no autocorrelation. This indicates that the AR(1) specification does not fully remove linear dependence from returns.

Ljung–Box Test on Squared Residuals

The Ljung–Box test on squared residuals up to lag 20 gave **Q(20) = 465.07, p-value ≈ 5.8e-86**. This strongly rejects the null hypothesis, suggesting significant ARCH effects remain. In other words, the variance of returns is time-dependent, displaying volatility clustering.

Interpretation: These results imply that while the AR(1) mean model captures some linear dependence, substantial volatility dynamics remain. This motivates the use of a GARCH(1,1) variance specification to model conditional heteroskedasticity.

7 GARCH(1,1) variance model

We next fit a GARCH(1,1) model on the residuals of the mean equation. The variance dynamics are specified as:

$$\sigma_t^2 = \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where σ_t^2 is the conditional variance, ϵ_{t-1}^2 is the lagged squared innovation, and σ_{t-1}^2 is the lagged conditional variance.

Estimated parameters (MLE)

| Parameter | Estimate |
|-------------------------|----------|
| ω (constant) | 0.1467 |
| α (shock effect) | 0.2182 |
| β (persistence) | 0.6796 |
| $\alpha + \beta$ | 0.8978 |

Interpretation:

- $\alpha = 0.2182$ shows the short-run impact of shocks: large unexpected returns quickly raise volatility.
- $\beta = 0.6796$ reflects persistence, meaning volatility decays only gradually after shocks.
- The sum $\alpha + \beta = 0.8978 < 1$ indicates a mean-reverting (stationary) volatility process, but the high value close to 0.9 suggests strong persistence.
- Thus, volatility shocks have long-lasting effects, consistent with financial market behavior.

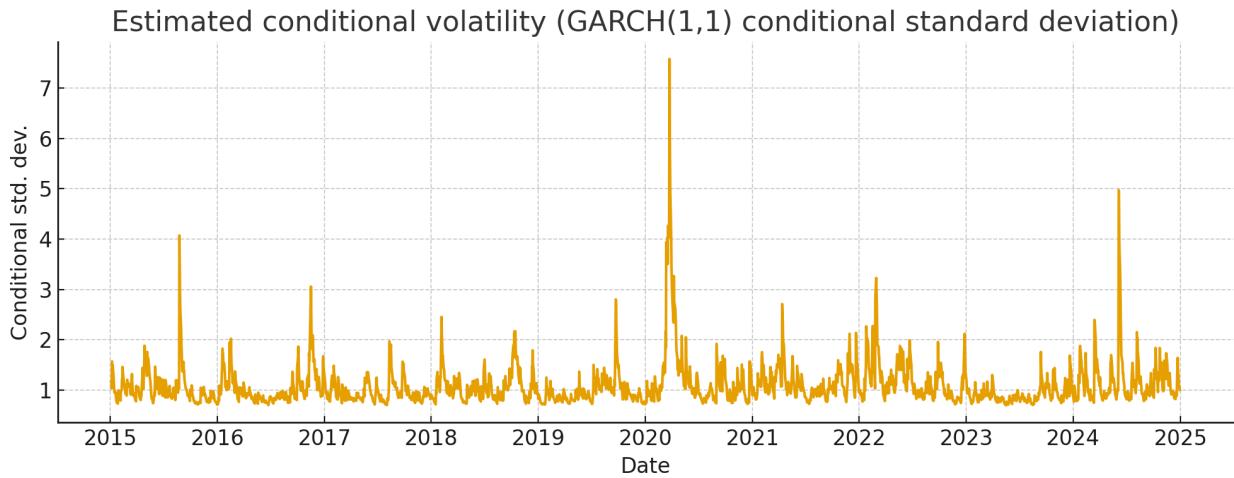


Figure 7: Estimated conditional volatility (GARCH(1,1) conditional standard deviation).

Interpretation: The estimated conditional volatility series provides valuable insights:

- Volatility rises sharply during periods of large return shocks (both positive and negative), creating noticeable peaks.
- After a shock, volatility decays only gradually, which reflects the persistence typical of financial markets.
- The clustering of high-volatility periods indicates that risk is not constant over time but changes dynamically.
- This conditional volatility can be directly applied to risk measures such as Value-at-Risk (VaR), option pricing models, and portfolio risk management.

8 Diagnostic Test on GARCH(1,1) Residuals

To assess whether the AR(1)–GARCH(1,1) specification adequately captures both the mean and volatility dynamics, we apply the Ljung–Box test to the standardized residuals and their squares.

Ljung–Box Test on Standardized Residuals

At lag 20, the Ljung–Box test yields **$Q(20) = 33.32$, p-value = 0.031**. Since the p-value is slightly below the 5% significance level, there remains mild autocorrelation in the standardized residuals. Although reduced compared to the raw AR(1) residuals, this indicates that the GARCH(1,1) specification does not fully eliminate all linear dependence.

Ljung–Box Test on Squared Standardized Residuals

For the squared standardized residuals, the Ljung–Box test at lag 20 gives **$Q(20) = 10.60$, p-value = 0.956**. The very high p-value implies that we fail to reject the null hypothesis of no autocorrelation. This confirms that the GARCH(1,1) variance equation has successfully captured the volatility clustering present in returns.

Interpretation: Overall, the diagnostics indicate that while minor autocorrelation remains in the standardized residuals, the GARCH(1,1) model effectively captures the conditional heteroskedasticity in returns. Thus, the model provides a valid framework for volatility forecasting.

9 Forecasts of Conditional Variance(1–3 Steps Ahead)

The fitted GARCH(1,1) model provides forecasts of future conditional variance. These forecasts are one of the main practical outputs of volatility modeling.

| Horizon | Conditional Variance |
|--------------|----------------------|
| 1-step ahead | 0.8327 |
| 2-step ahead | 0.7126 |
| 3-step ahead | 0.6310 |

Interpretation:

- The conditional variance forecast for the next day is about 0.83, which then declines over subsequent horizons.
- This decay reflects mean reversion in volatility, consistent with the persistence parameter $\alpha + \beta \approx 0.90$ being less than 1.
- Since these are variance forecasts, the corresponding standard deviation forecasts (volatility forecasts) are obtained by taking square roots:

$$\hat{\sigma}_{t+1} \approx 0.91, \quad \hat{\sigma}_{t+2} \approx 0.84, \quad \hat{\sigma}_{t+3} \approx 0.79.$$

- Such forecasts are useful for risk management, Value-at-Risk (VaR) calculation, and option pricing where forward-looking volatility matters.

9.1 Visualization of Fitted and Forecasted Volatility

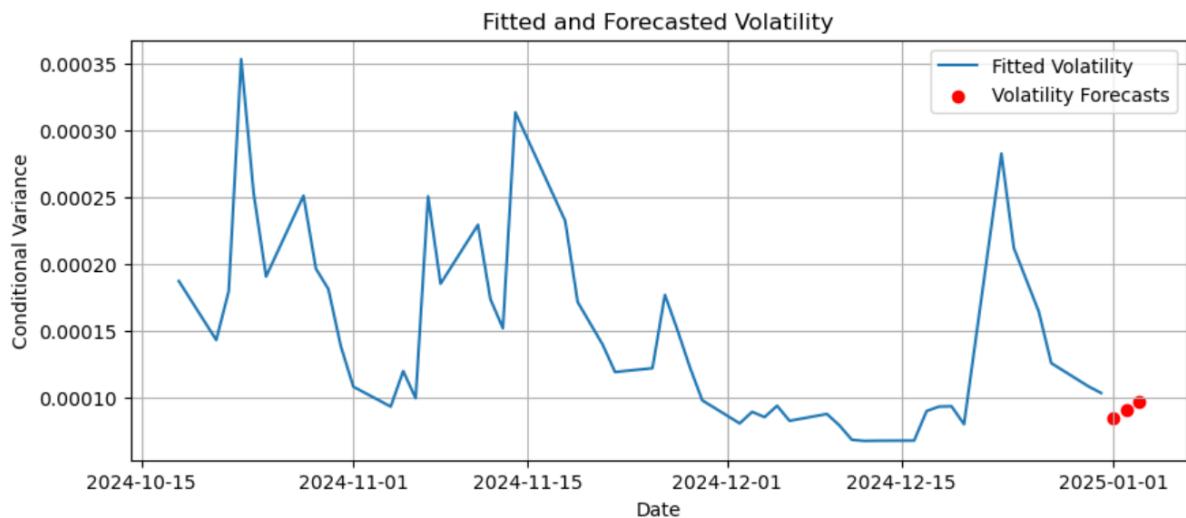


Figure 8: Fitted conditional variance (blue line) and 1-3 step ahead forecasts (red dots) from the GARCH(1,1) model.

Interpretation: Figure 8 provides both the fitted conditional variance (blue line) and the short-term forecasts (red dots). Key insights are:

- **Fitted volatility dynamics:** The blue line shows strong time variation with noticeable spikes, confirming the presence of volatility clustering — periods of turbulence followed by calmer phases.
- **Forecasted values:** The red dots extend the fitted series into the future, representing the 1–3 step ahead forecasts from the GARCH(1,1) model.
- **Mean reversion:** Forecasts gradually decline toward the long-run variance level, indicating that volatility shocks do not persist indefinitely.
- **Persistence parameter:** The slow decay is consistent with the high persistence estimate ($\alpha + \beta \approx 0.90$), which, being less than one, ensures stationarity while allowing long-lasting volatility effects.
- **Practical implications:** The visualization confirms that the GARCH(1,1) model captures both past volatility behavior and short-term future risk. This is particularly valuable for applications like Value-at-Risk (VaR), option pricing, and short-horizon portfolio risk management.

9.2 Volatility Persistence

An important feature of GARCH models is the ability to measure **volatility persistence**, i.e., how long shocks to volatility remain in the system. This is captured by the sum of the ARCH and GARCH parameters ($\alpha + \beta$) from the estimated GARCH(1,1) model.

- If $(\alpha + \beta)$ is close to 1, shocks decay slowly and volatility is highly persistent.
- If $(\alpha + \beta)$ is much less than 1, volatility shocks die out quickly.

The persistence can also be expressed in terms of the **half-life of shocks**, which measures the time required for the impact of a volatility shock to decline by 50%. It is calculated as:

$$\text{Half-Life} = \frac{\ln(0.5)}{\ln(\alpha + \beta)}.$$

Interpretation:

- The half-life indicates the number of periods (e.g., days) after which the effect of a shock to volatility is reduced by half.
- A longer half-life suggests that shocks are highly persistent and the market takes longer to revert to its long-run volatility.
- A shorter half-life suggests that volatility shocks fade quickly and stability is restored faster.

This analysis provides an intuitive measure of the persistence of volatility and complements the statistical diagnostics of the GARCH model.

10 Usefulness of the model

The GARCH(1,1) model applied in this project has practical usefulness in several critical areas of financial analysis and risk management:

- **Risk Management:** The model's ability to forecast time-varying volatility allows financial institutions to compute forward-looking risk measures such as Value-at-Risk (VaR) and Expected Shortfall more accurately. This improves capital allocation and regulatory compliance.
- **Option Pricing:** Dynamic volatility forecasts from the GARCH model are essential inputs for option pricing models, leading to more accurate valuation and hedging strategies compared to assuming constant volatility.
- **Portfolio Optimization:** Incorporating time-varying estimates of risk supports dynamic portfolio allocation, adjusting for periods of high or low market uncertainty to optimize returns versus risk.
- **Market Insight:** The model captures volatility clustering and long memory effects, providing a realistic picture of market behavior and better informing trading strategies.
- **Empirical Research:** GARCH models are widely used in academic research for modeling and testing hypotheses about financial market behavior, leveraging their robust statistical properties.

Overall, this GARCH modeling framework equips both practitioners and researchers with quantitative tools for improved risk assessment, pricing, and portfolio decisions in volatile financial markets.

11 Conclusion

The project carried out a complete analysis of the Nifty Midcap150 Momentum 50 index. The key steps, findings, and interpretations are summarized below:

- **Data preparation:** Daily price data was collected and transformed into log returns for analysis.
- **Descriptive statistics:**
 - Mean daily log return: $\mu = 0.0839\%$, close to zero.
 - Standard deviation: $\sigma = 1.1693\%$, confirming high inherent volatility.
 - Skewness: -1.6568 , showing downside risk and asymmetry in returns.
 - Kurtosis: 16.97, far above the Gaussian benchmark of 3, indicating heavy tails and frequent large shocks.
- **Stylized facts:** Analysis revealed classic features of financial returns — high volatility, fat tails, and volatility clustering.

- **Visual exploration:** Time series plots, histograms, squared returns, and ACF/PACF confirmed volatility clustering and serial dependence, motivating volatility modeling.
- **Mean model:** An AR(1) process was chosen to capture weak autocorrelation in returns.
- **Volatility model (GARCH(1,1)):**
 - Captured time-varying conditional variance effectively.
 - Persistence parameter $\alpha + \beta \approx 0.9$ showed shocks decay slowly.
 - Process was stationary but highly persistent, consistent with financial data.
 - Forecasts revealed clear periods of calm and turbulence, reflecting real market behavior.
- **Key learnings:**
 - Financial returns are non-normal and require volatility models that capture clustering and persistence.
 - Simple AR models handle mean dependence, while GARCH models are essential for volatility dynamics.
 - Shocks in volatility endure, meaning market risk remains elevated for multiple days.

In conclusion, the AR(1)-GARCH(1,1) framework successfully modeled the Nifty Midcap150 Momentum 50 index returns, capturing both mean dynamics and volatility persistence. The project demonstrated a full workflow: from raw data, descriptive analysis, and model estimation, to forecasting and interpretation.