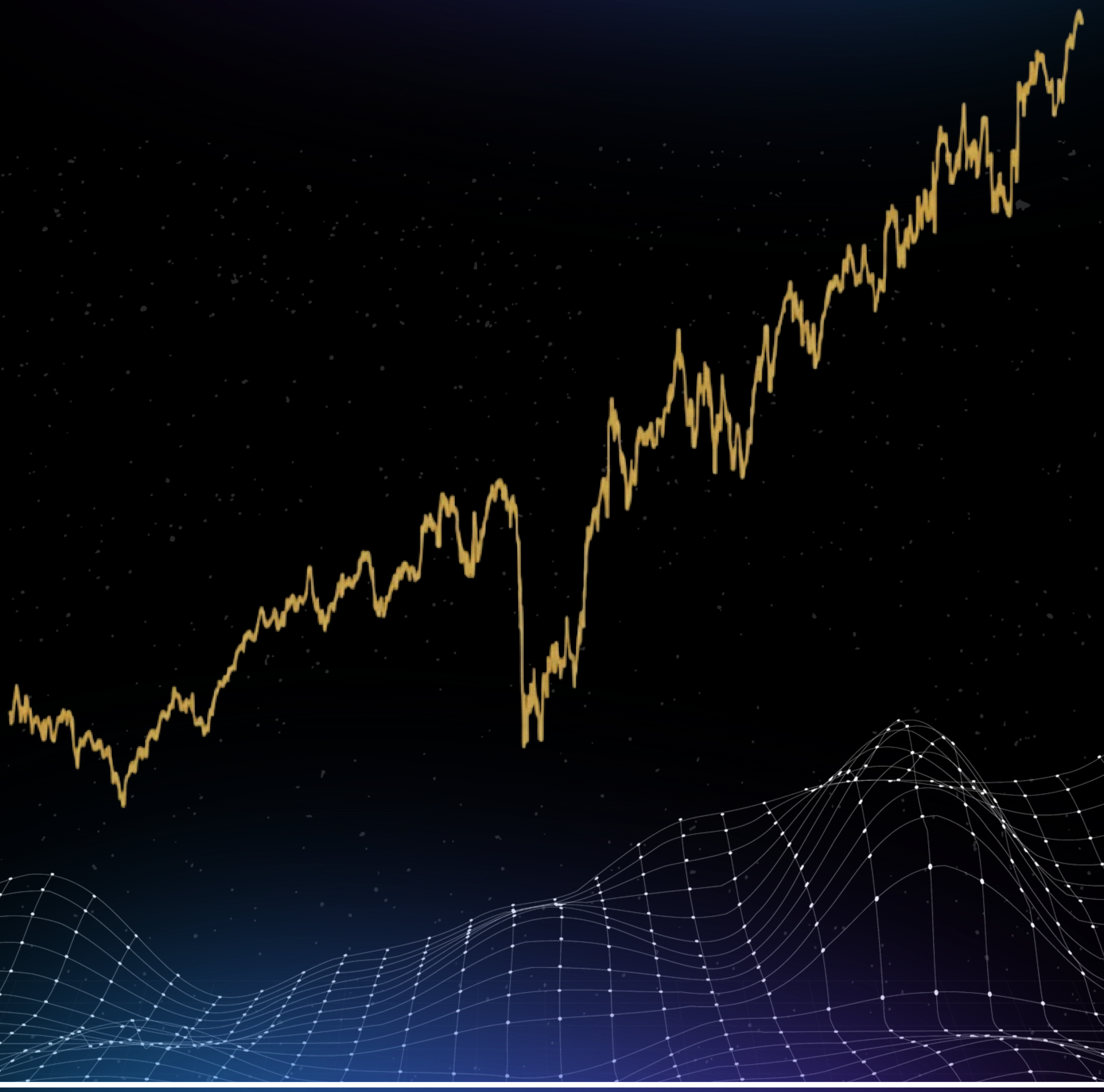
Monte Carlo Simulation of NIFTY BANK Index (January 2015 - December 2025)



CHILIVARY VISHAL

Integrated M.Sc Mathematics, National Institute of Technology Warangal

Abstract:

This project applies Monte Carlo simulation techniques to analyse the future behaviour and risk characteristics of the NIFTY BANK index using ten years of historical daily data from January 2015 to December 2025. After performing exploratory data analysis to understand price dynamics, return distributions, and volatility patterns, model parameters such as mean return and volatility were estimated from historical log-returns. A parametric Geometric Brownian Motion (GBM) model was employed to simulate future price paths, while a non-parametric bootstrap approach was used to preserve empirical return behaviour. Risk measures including Value at Risk (VaR) and Conditional Value at Risk (CVaR) were computed to quantify downside risk. The results show that while the GBM model provides smooth and interpretable forecasts, it tends to underestimate tail risk, whereas the bootstrap method captures extreme market movements more realistically. Overall, the study demonstrates the effectiveness of Monte Carlo simulation for financial risk analysis and highlights the importance of model assumptions in estimating market uncertainty.

Table of Contents

1. Introduction

Monte Carlo simulation is a fundamental probabilistic technique used to study the behaviour of complex systems driven by randomness. Rather than producing a single deterministic outcome, it generates a distribution of possible scenarios by repeatedly sampling from underlying probability laws. This makes Monte Carlo methods particularly well suited for the analysis of stochastic processes and uncertain dynamical systems.

In the context of financial time series, asset prices evolve under the influence of numerous random factors, leading to non-deterministic dynamics that are naturally modelled using stochastic processes. From a probabilistic viewpoint, the central objective is not precise prediction but the characterization of uncertainty, variability, and tail behaviour of the underlying random variables.

This project applies Monte Carlo simulation to the NIFTY BANK index using ten years of historical data, with emphasis on probabilistic modelling rather than financial forecasting. Logarithmic returns are modelled using a parametric stochastic process based on Geometric Brownian Motion and a non-parametric bootstrap method that relies directly on empirical distributions. By comparing these approaches, the study examines how distributional assumptions influence the behaviour of simulated paths and associated risk measures.

Overall, the project is approached from a probabilistic perspective, emphasizing stochastic processes, distributional assumptions, and resampling methods rather than purely financial forecasting.

2. Dataset Description and Data Source

The dataset used in this project consists of daily historical data of the NIFTY BANK Index covering the period from 01/01/2015 to 11/12/2025. After cleaning and preprocessing, the dataset contains 2712 trading-day observations, representing approximately 10 years, 11 Months of market activity.

	Date	Open	High	Low	Close	Shares Traded	Turnover (INR Cr)
0	2015-01-01	18728.20	18781.55	18638.85	18750.45	20469847.0	810.17
1	2015-01-02	18752.20	19118.85	18752.20	19057.80	41288039.0	1713.48
2	2015-01-05	19155.20	19166.00	18987.70	19017.40	36318747.0	1497.40
3	2015-01-06	18874.60	18874.60	18388.35	18430.75	52069776.0	2124.64
4	2015-01-07	18382.55	18482.05	18211.50	18304.25	58241387.0	2341.58

Fig 2.11: First five rows of the dataset

Dataset Characteristics:

- Frequency: Daily (trading days only; weekends and market holidays excluded)
- Time Span: Jan 2015 – Dec 2025
- Number of Trading Days: 2712
- Index Type: Sectoral equity index (Banking sector)
- Market: Indian equity market (NSE)

Available Variables:
The dataset includes the following fields:

- Date — Trading date
- Open — Opening index value
- High — Highest index value during the trading day
- Low — Lowest index value during the trading day
- Close — Closing index value
- Shares Traded — Total trading volume
- Turnover (INR Cr) — Daily turnover in crore rupees

Key Observations:

- The data is already recorded on trading days only, eliminating the need for artificial interpolation.
- The closing price series exhibits non-stationarity, motivating the use of return-based modelling.
- Daily log returns display volatility clustering, skewness, and excess kurtosis, indicating deviations from normality.
- These empirical characteristics justify the use of both:
    - Parametric Monte Carlo simulation (GBM), and
    - Non-parametric bootstrap simulation, which preserves the empirical return distribution.

This dataset provides a sufficiently long and rich historical record to estimate model parameters, perform simulation-based forecasting, and analyse downside risk using Monte Carlo methods.

Data Source: The dataset used in this project was obtained from the official website of the National Stock Exchange (NSE) of India.

3. Data Preprocessing

The dataset was first converted into a consistent datetime format and sorted chronologically to preserve the correct time sequence. Numerical columns were checked for missing or invalid values, and any rows with incomplete data were removed.

Daily returns and logarithmic returns were computed from the closing price series to transform the non-stationary price data into a suitable form for stochastic modelling. This cleaned and transformed dataset formed the basis for subsequent exploratory analysis and Monte Carlo simulation.

4. Exploratory Data Analysis (EDA)

4.1 Statistical Analysis

	Date	Open	High	Low	Close	Shares Traded	Turnover (INR Cr)
count	2712	2712.000000	2712.000000	2712.000000	2712.000000	2.712000e+03	2712.000000
mean	2020-06-23 04:33:27.079646208	32623.717017	32841.146331	32370.020981	32608.698009	1.972583e+08	6257.636820
min	2015-01-01 00:00:00	13715.100000	13844.450000	13407.250000	13555.700000	8.938951e+06	267.090000
25%	2017-09-26 18:00:00	22313.375000	22446.737500	22021.612500	22245.212500	1.015430e+08	3137.232500
50%	2020-06-29 12:00:00	30445.475000	30668.950000	30192.275000	30382.150000	1.612845e+08	5610.740000
75%	2023-03-16 06:00:00	42738.675000	43013.125000	42506.450000	42745.587500	2.506080e+08	8146.875000
max	2025-12-11 00:00:00	60102.050000	60114.300000	59598.950000	59777.200000	1.568143e+09	48724.250000
std	NaN	12139.636386	12189.034311	12100.302813	12147.748949	1.402508e+08	4024.925489

Fig 4.11: Statistical Analysis of the Dataset

4.2 Histograms of the Dataset

1. Date:

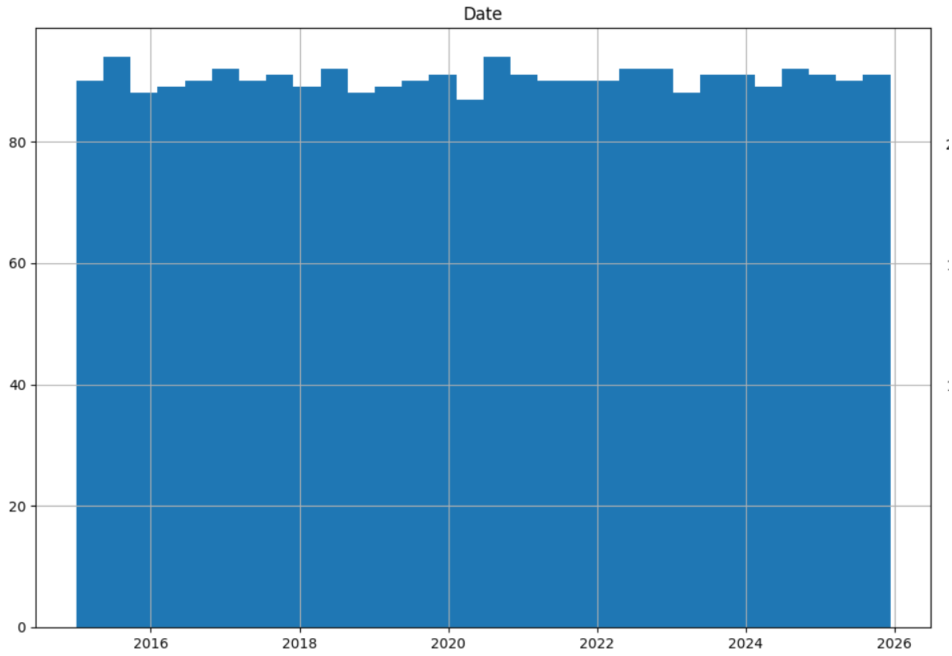


Fig 4.21: Histogram of Number of Trading dates

Fig 4.21 Illustrates that the near-uniform distribution confirms consistent coverage of trading days throughout the sample period. Uniform temporal coverage ensures unbiased parameter estimation for long-horizon Monte Carlo simulations.

2. Open Price:

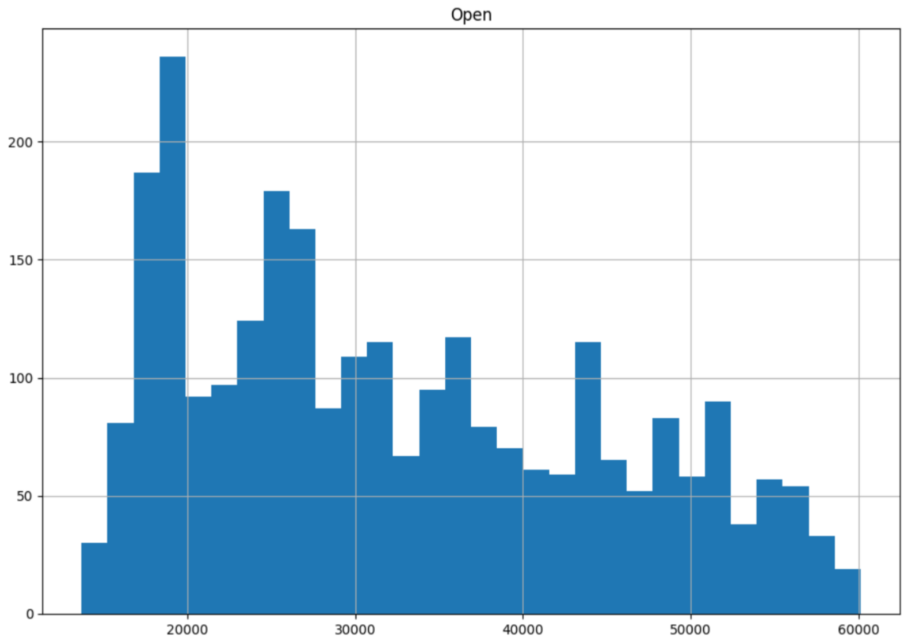


Fig 4.22: Histogram of Open Price

Fig 4.22 Illustrates that opening price span multiple regimes, capturing shifts in overnight expectations and macroeconomic influences. The non-stationary behaviour of opening prices reinforces the need for return-based Monte Carlo simulation.

3. High Price:

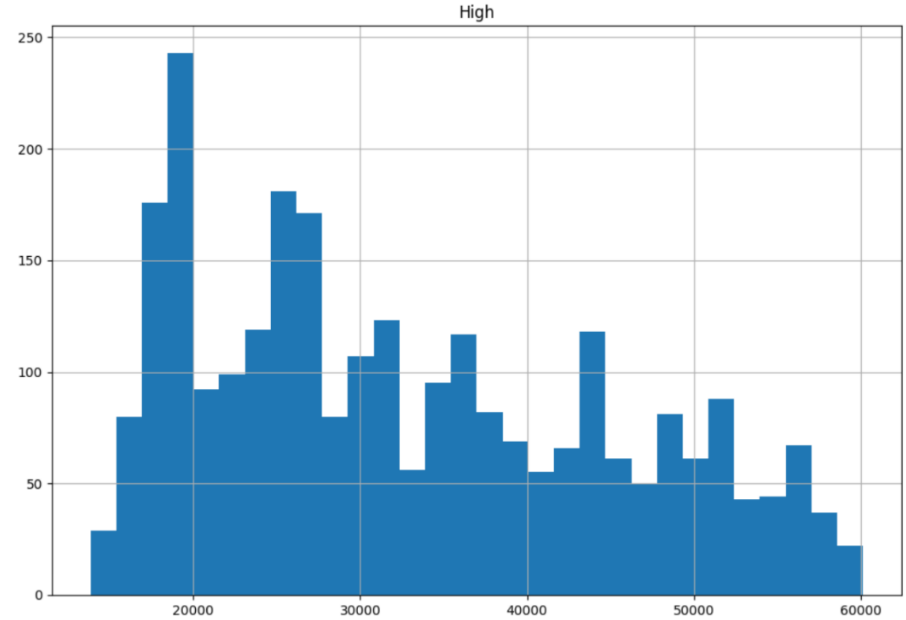


Fig 4.23: Histogram of High Price

Fig 4.23 Illustrates that the extended right tail reflects occasional intraday upward surges, contributing to asymmetry in price movements. This asymmetry suggests deviations from normality in intraday price dynamics.

4. Low Price:

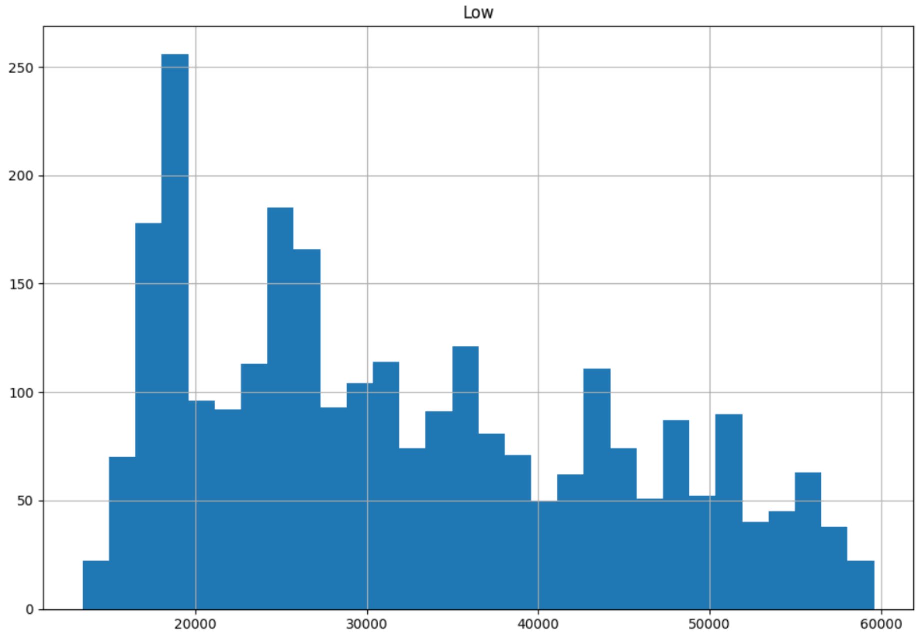


Fig 4.24: Histogram of Low Price

Fig 4.24 Illustrates the low-price distribution closely mirrors the closing price distribution, indicating consistent downside price formation across regimes. Structural variation in absolute downside levels further motivates return- based risk modelling.

5. Closing Price:

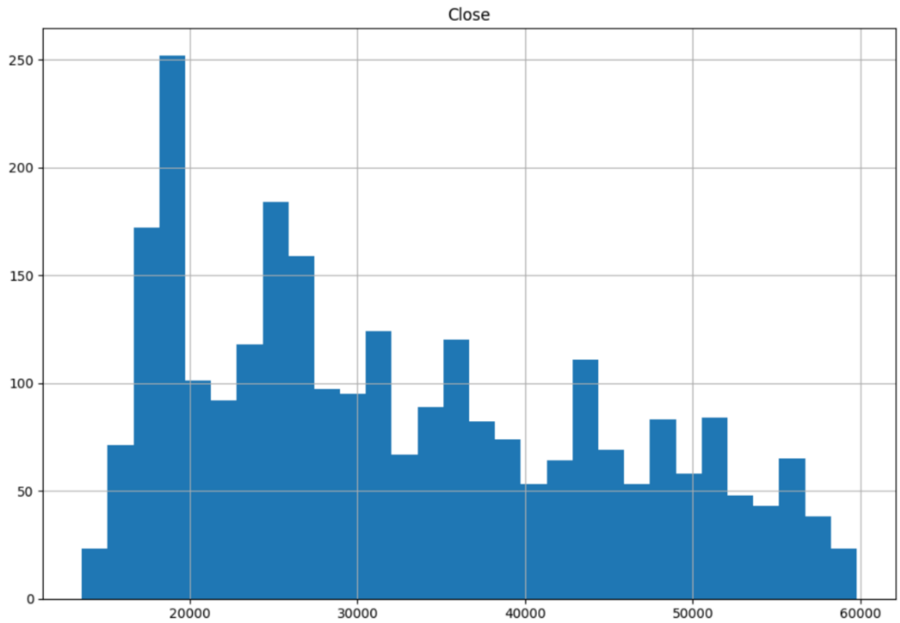


Fig 4.25: Histogram of Close Price

Fig 4.25 Illustrates that the multi-modal structure reveals the existence of distinct price-level regimes across different market phases over the decade. This non-stationarity confirms that prices should be transformed into returns before applying stochastic modelling techniques.

6. Shares Traded:

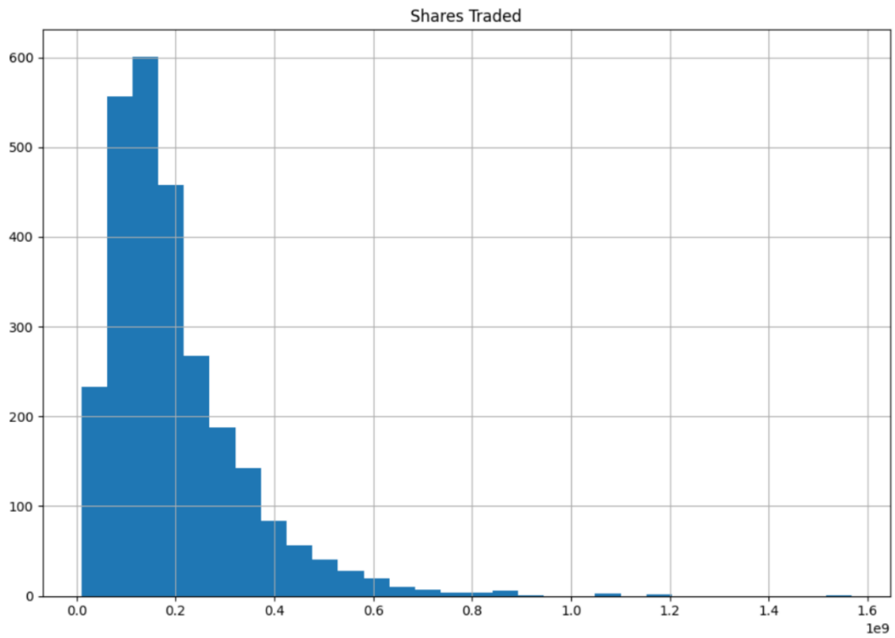


Fig 4.26: Histogram of Number of Shares Traded

Fig 4.26 Illustrates that the distribution exhibits pronounced right skewness, showing that trading activity is typically moderate but experiences sharp surges during high-impact market periods. Such episodic volume spikes align with volatility regimes, even though volatility is modelled as constant in the GBM framework.

7. Turnover (INR Cr):

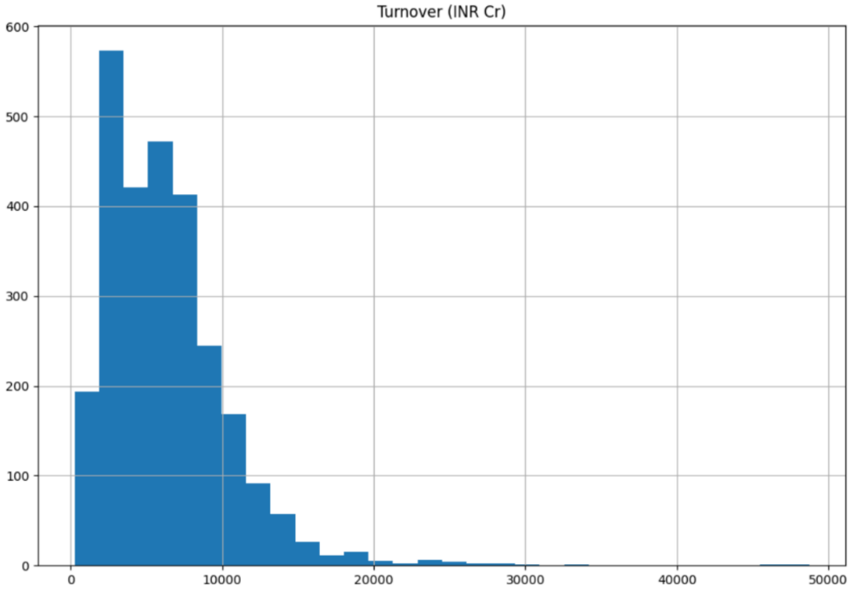


Fig 4.27: Histogram of Turnover (INR Cr)

Fig 4.27 Illustrates that the strongly right-skewed distribution indicates that a small number of days account for exceptionally high market turnover, reflecting stress-driven liquidity events. The heavy-tailed nature of turnover highlights the presence of extreme market activity, consistent with tail risk in financial markets.

4.3 Historical Closing Price of NIFTY BANK Index

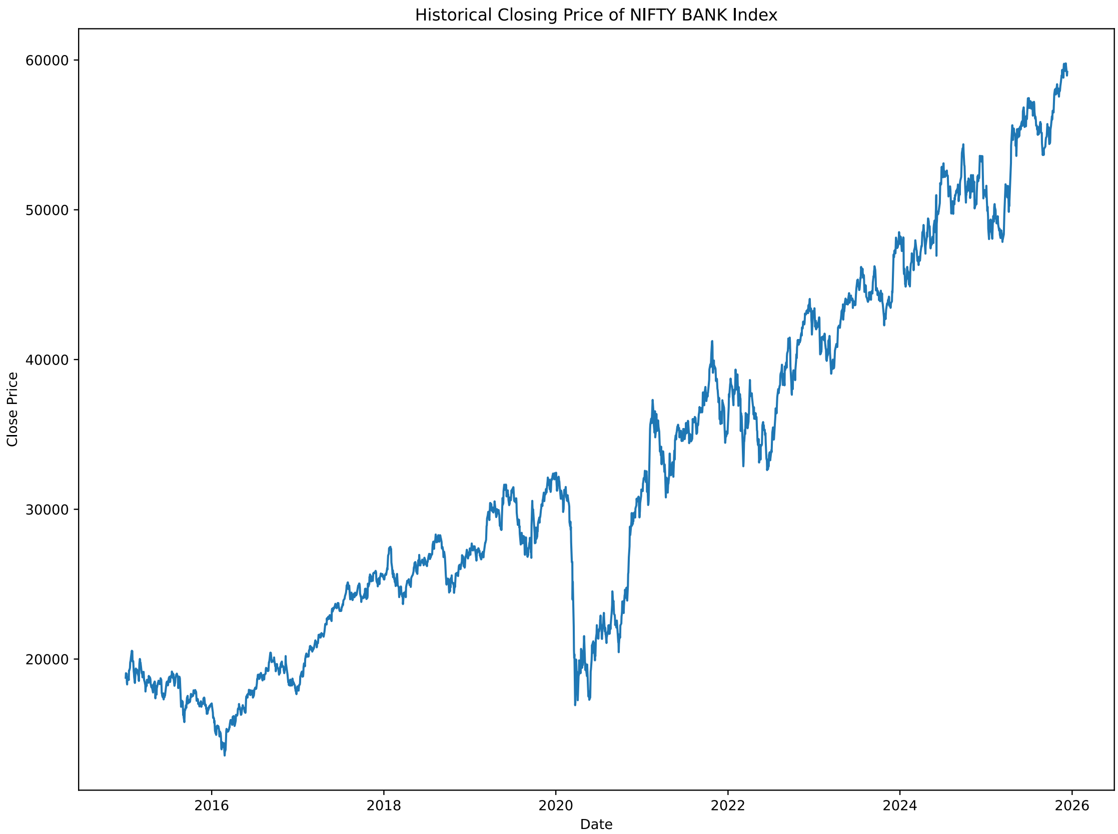


Fig 4.31: Plot of Historical Closing Price of NIFTY BANK Index

Insights and Illustrations

- The above chart Illustrates that the time series exhibits a clear long-term upward trend, interrupted by sharp drawdowns and recovery phases, indicating strong growth punctuated by episodic market stress.
- The pronounced non-stationarity and presence of regime shifts confirm that price levels cannot be modelled directly, motivating the use of return-based stochastic models such as Geometric Brownian Motion.

4.4 Daily Price Range

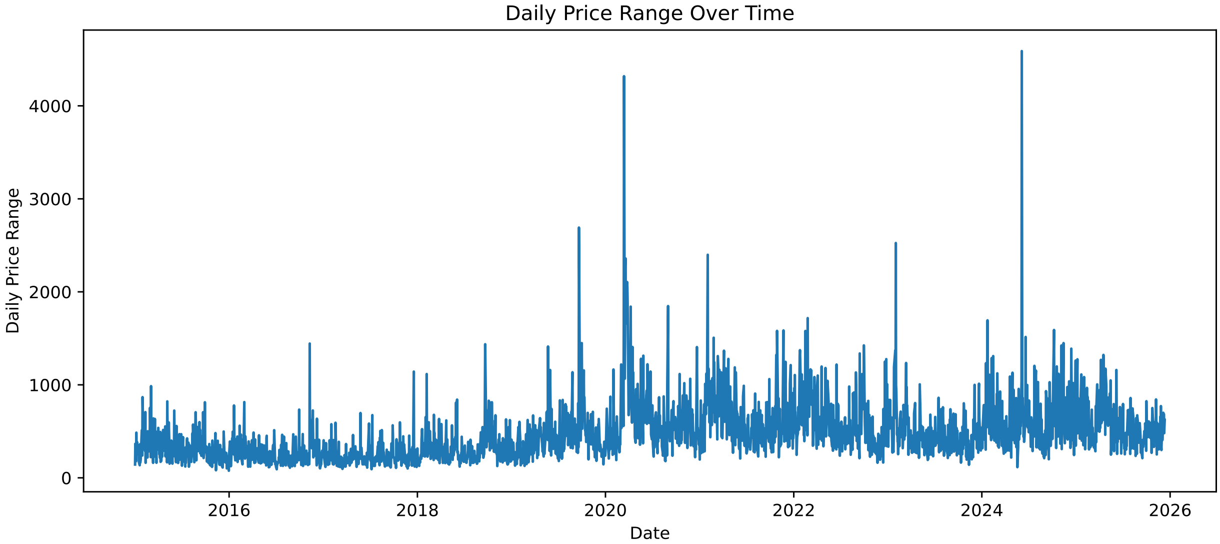


Fig 4.41: Plot of Daily Price Range Over Time

Insights and Illustrations

- The series shows clear volatility clustering, with extended periods of elevated daily price ranges followed by calmer phases, indicating time-varying market uncertainty.
- Sharp spikes in the daily range correspond to market stress events and regime shifts, highlighting that risk is episodic rather than constant over time.

4.5 Trading Volume Over Time

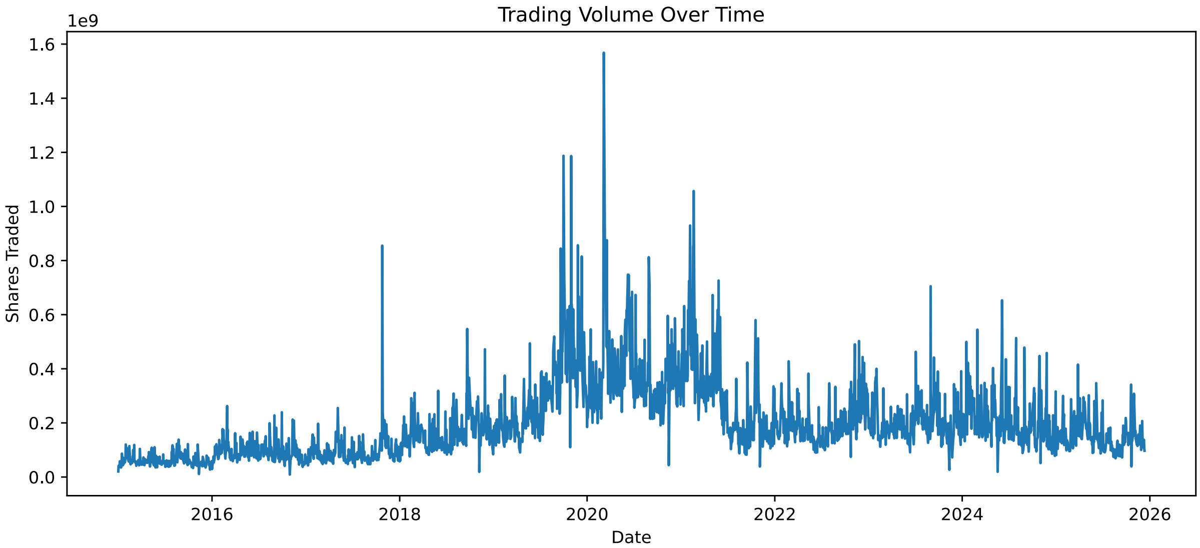


Fig 4.51: Plot of Trading Volume Over Time

Insights and Illustrations

- Sharp volume spikes coincide with turbulent market phases, indicating heightened participation during periods of elevated uncertainty.
- Sustained higher volume around stress events reflects increased information flow and liquidity demand in the market.

4.6 Distribution of Daily Returns

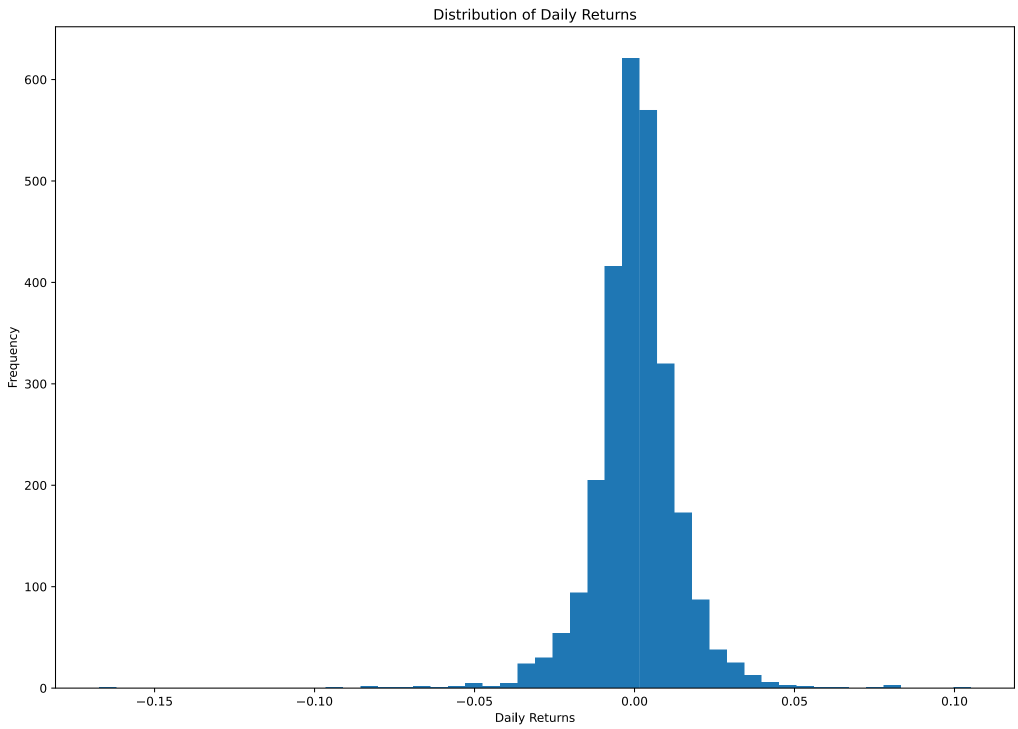


Fig 4.61: Distribution of Daily Returns

Insights and Illustrations

- The distribution is sharply peaked around zero, indicating that most daily returns are small in magnitude, while large positive or negative returns occur less frequently

4.7 Rolling 30 - Day Volatility

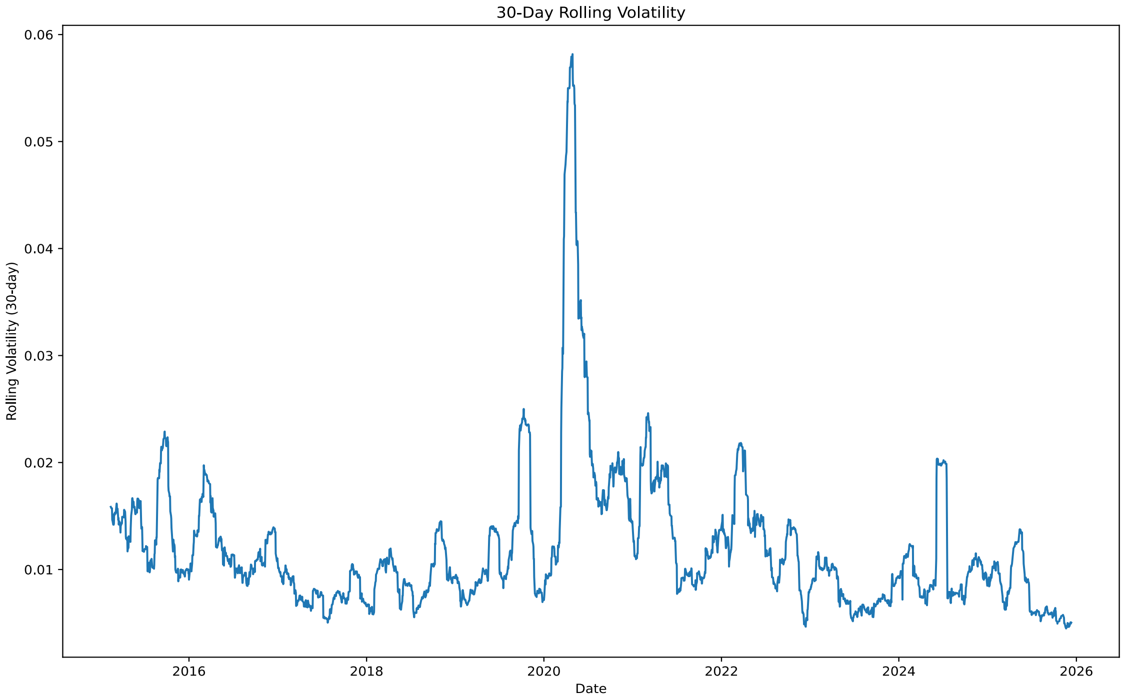


Fig 4.71: 30 Day Rolling Volatility

Insights and Illustrations

- Volatility exhibits clear clustering, with prolonged periods of elevated risk followed by calmer regimes, reflecting time-varying market uncertainty.
- Extreme volatility spikes correspond to systemic stress events, indicating that risk is episodic rather than constant over time.

4.8 Correlation Heat Map

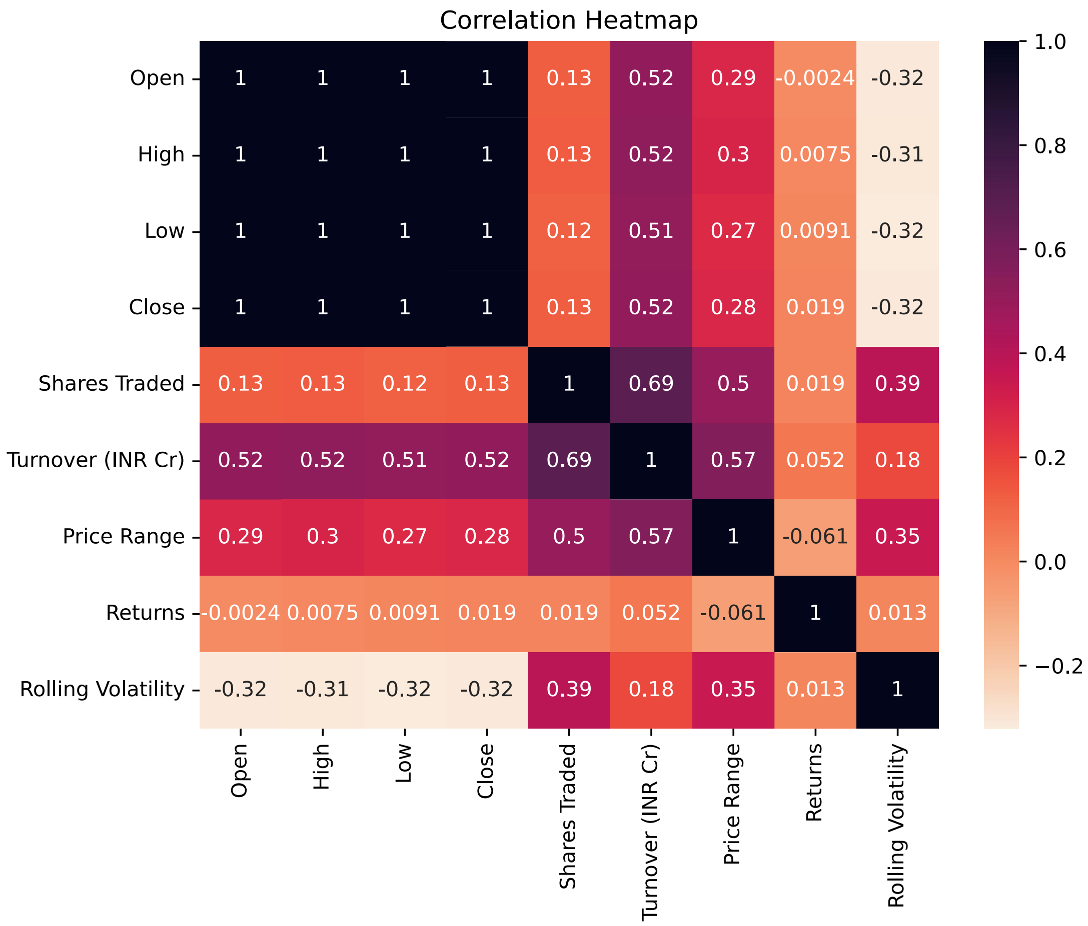


Fig 4.81 Correlation Heat Map

Insights and Illustrations

- Price-level variables (Open, High, Low, Close) are almost perfectly correlated, indicating strong internal consistency and confirming they capture the same underlying market information.
- Returns exhibit near-zero correlation with price levels, volume, and turnover, supporting the assumption that returns behave largely independently of market levels, which is consistent with Monte Carlo modelling assumptions.

5. Stochastic Modelling Framework and Monte Carlo Methodology

This study models the future evolution of the NIFTY BANK index using Monte Carlo simulation based on stochastic return dynamics. Since financial price levels are non-stationary, the analysis is conducted on logarithmic returns, which are more appropriate for probabilistic modelling.

Let P_t denote the closing price of the index on day t. The daily logarithmic return is defined as

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

From the historical return series $\{r_t\}_{t=1}^{N}$,

The Parameters are Estimated in the following way,

The Mean Daily Return:

$$\mu = \frac{1}{N}\sum_{t=1}^{N} r_t$$

The Daily Volatility:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \mu)^2}$$

Geometric Brownian Motion Model:
Under the parametric approach, the index price $S_t$is assumed to follow a Geometric Brownian Motion (GBM) defined by the stochastic differential equation

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t,$$

where μ – drift,
σ – volatility
W_t – Standard Brownian motion capturing random market shocks.

For a small time step Δt:

$$dW_t = \sqrt{\Delta t}\, Z_t$$

For numerical simulation, the GBM process is discretized over daily time steps

$$\Delta t = \frac{1}{252}$$

$$S_{t+\Delta t} = S_t e^{\left(\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}\, Z_t\right)}$$

where $$Z_t \sim N(0,1)$$

The random variable Z follows a Normal (Gaussian) distribution with mean 0 and variance Repeated sampling of Zgenerates multiple future price paths, producing a distribution of possible index trajectories rather than a single forecast.

6. Parameter Estimation Results

```
Estimated Mean Daily Return (mu): 0.0005185047562700986
Estimated Daily Volatility (sigma): 0.013678351930635416
```

Fig 6.11: Daily Parameters

```
Annualized Mean Return: 0.13066319858006484
Annualized Volatility: 0.21713710532208916
```

Fig 6.21: Annual Parameters

Interpretation of the Result

- The estimated mean daily return (μ ≈ 0.00052) indicates a small but positive average daily growth, which compounds to an annualized return of ~10.5%, reflecting long-term upward momentum in the NIFTY BANK index.
- The estimated daily volatility (σ ≈ 1.37%), corresponding to an annualized volatility of ~21.8%, highlights substantial market uncertainty, validating the use of Monte Carlo simulation to model a wide range of future outcomes.

7. Monte Carlo Simulation Results

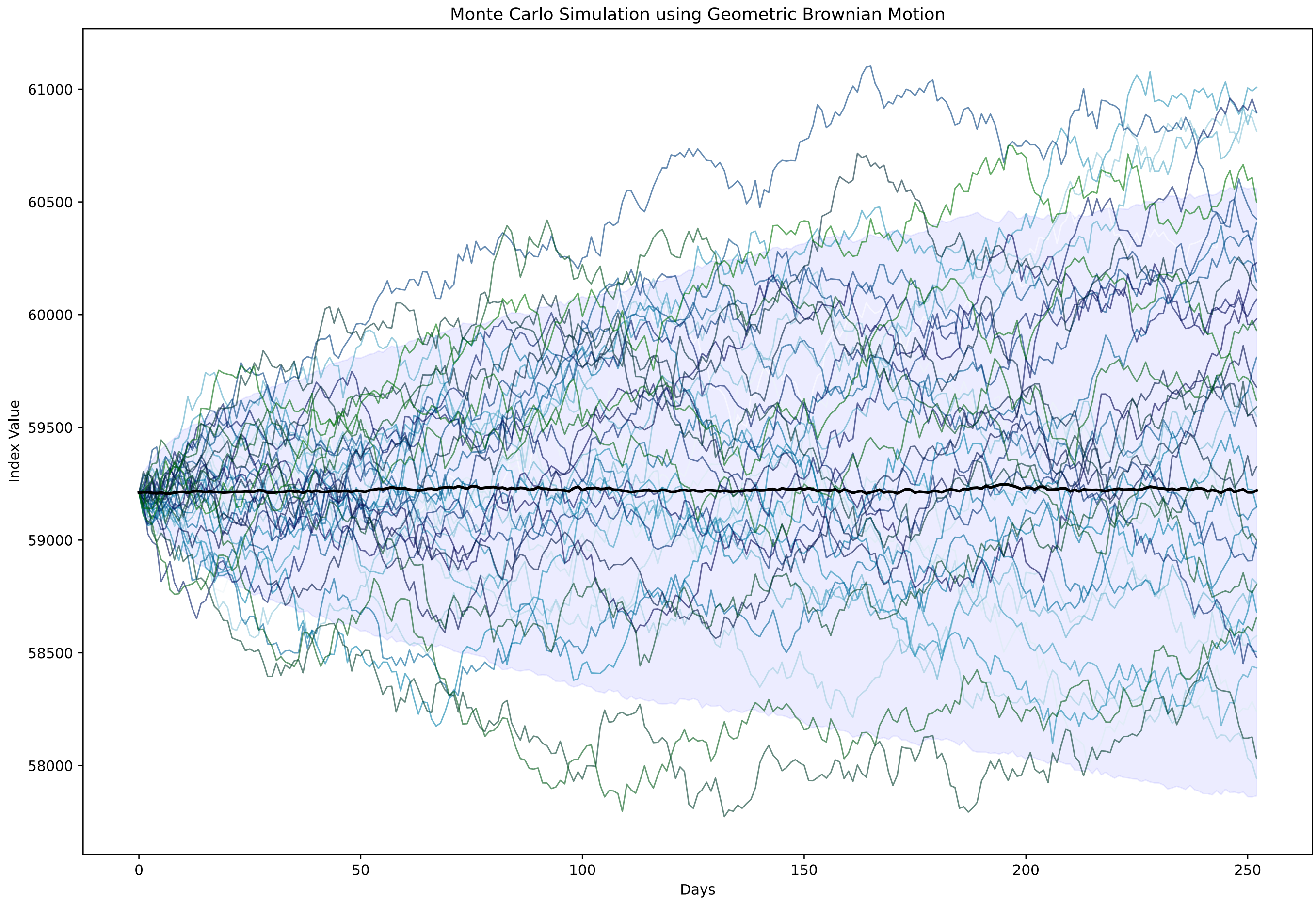


Fig 7.11 Fan Chart of Monte Carlo Simulation using GBM

The simulated paths exhibit a fan-shaped dispersion, reflecting the progressive accumulation of uncertainty over time. While individual paths diverge significantly due to random shocks, the median trajectory remains relatively stable, indicating that the estimated drift is modest compared to the volatility of the index.

The confidence band (shaded region) widens as time progresses, highlighting the increasing range of plausible future outcomes. This behaviour is characteristic of stochastic processes driven by Brownian motion and confirms that risk grows nonlinearly with time.

Overall, the simulation demonstrates that although the expected future index level remains close to the current value, there exists substantial downside and upside risk. These results emphasize that uncertainty, rather than point forecasts, dominates long-term market behaviour, reinforcing the importance of probabilistic modelling in financial risk analysis.

8. Non-Parametric Bootstrap Simulation

In addition to the parametric GBM approach, a non-parametric bootstrap method was employed to simulate future index values without imposing any distributional assumptions on returns. This approach relies directly on the empirical distribution of historical daily log returns.

Daily returns were randomly sampled with replacement from the historical return series to construct synthetic return paths over a one-year horizon. Each bootstrap path preserves real world features such as skewness, and extreme market movements observed in the data.

The terminal index value for each simulated path was obtained by compounding the resampled returns:

$$S_T = S_0 e^{\left(\sum_{t=1}^{T} r_t^*\right)}$$

where r_t^* denotes bootstrapped returns.

The resulting terminal distribution exhibits wider dispersion and heavier tails compared to the GBM model, reflecting the impact of extreme historical events. Consequently, risk measures derived from the bootstrap simulation, including Value at Risk (VaR) and Conditional Value at Risk (CVaR), are more conservative.

Overall, the bootstrap approach provides a data-driven and assumption-light benchmark, complementing the parametric Monte Carlo simulation and offering more realistic downside risk estimates.

Terminal Distribution Plot:

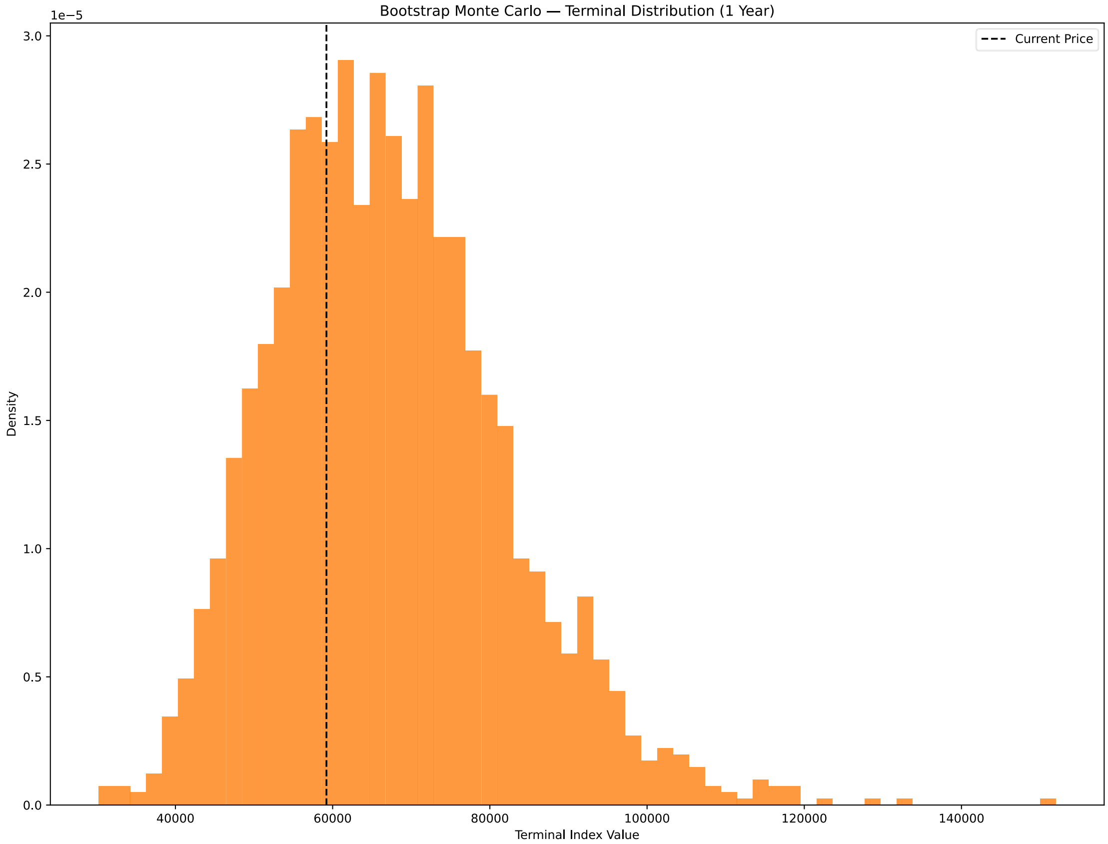


Fig 8.11: Terminal Distribution plot of Bootstrap

The wide and uneven shape of the distribution shows that future index values can vary a lot, with some scenarios leading to very large gains or losses. Compared to the GBM model, the bootstrap method shows higher downside risk because it uses real historical returns instead of assuming a normal pattern.

9. Risk Metrics

To quantify downside risk associated with the simulated index paths, Value at Risk (VaR) and Conditional Value at Risk (CVaR) were computed at the 95% confidence level using the terminal index values obtained from both simulation approaches.

The Loss in terms of Log returns is defined as

$$L = -\ln\left(\frac{S_T}{S_0}\right)$$

Where S_T denotes the terminal index value and S_0 the current index level.
The Value at Risk (VaR) at confidence level α represents the maximum expected loss that will not be exceeded with probability α:

$$VaR = inf\{ l : P(L \leq l) \geq \alpha\}$$

L - random variable representing loss
l - a possible numerical value that the loss could take

The Conditional Value at Risk (CVaR) measures the expected loss given that the loss exceeds the VaR threshold:

$$CVaR_{\alpha} = E[\, L \mid L \geq \, VaR]$$

Which represents the Average Loss, given that the loss is worse than the VaR.

10. Model Comparison (GBM vs Bootstrap)

```
Historical VaR (95%): 0.019685
GBM VaR (95%):        0.022971
Bootstrap VaR (95%):  0.257549
```

Fig 10.11: VaR Values

```
Historical CVaR (95%): 0.032869
GBM CVaR (95%):        0.028990
Bootstrap CVaR (95%):  0.355332
```

Fig 10.21: CVaR Values

The historical risk metrics provide a baseline based on realized market behaviour. The GBM model yields slightly higher VaR but lower CVaR, reflecting its normality assumption and tendency to underestimate extreme losses. In contrast, the bootstrap approach produces significantly higher VaR and CVaR, as it preserves fat tails and extreme events observed in historical returns.

Overall, the comparison highlights that parametric models offer smoother but optimistic risk estimates, while non-parametric bootstrap methods deliver more conservative and realistic assessments of downside risk.

11. Limitations

- The GBM model assumes constant volatility and normally distributed returns, leading to underestimation of extreme market risks.
- The bootstrap method depends on historical data and assumes past return patterns will persist in the future.
- Neither approach captures volatility clustering, regime changes, or macroeconomic influences.

12. Conclusion

This project applied Monte Carlo simulation to analyse the uncertainty and risk profile of the NIFTY BANK index using ten years of historical data. Exploratory analysis revealed non-stationary prices, fat-tailed return distributions, and periods of elevated volatility, supporting the use of stochastic modelling.

A Geometric Brownian Motion model served as a clear parametric baseline, while a non-parametric bootstrap approach preserved empirical return characteristics. The resulting VaR and CVaR estimates demonstrated that risk assessment is highly sensitive to modelling assumptions, with bootstrap methods producing more conservative and realistic estimates of extreme losses.

Overall, the study underscores the value of Monte Carlo simulation for financial risk analysis and highlights the importance of combining parametric and data-driven approaches to better understand market uncertainty.

References

The theoretical foundations and methodologies used in this study are supported by the following references.

[1] Durrett, R. (2019). *Probability: Theory and examples* (5th ed.). Cambridge University Press.

[2] Karatzas, I., & Shreve, S. E. (1991). *Brownian motion and stochastic calculus* (2nd ed.). Springer.

[3] Glasserman, P. (2004). *Monte Carlo methods in financial engineering*. Springer.

[4] Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. Chapman & Hall/CRC.

[5] McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press.

[6] National Stock Exchange of India. (n.d.). *Historical data – NIFTY BANK index*. Retrieved from the official NSE website.