Nifty 50 — Index Price Analysis (2015–2023) Report :

1. Dataset Description

1.1 Source:

- Kaggle dataset *"Nifty 50 (NSEI) Index Price Data (2015-2023)"* (daily index price history).

1.2 Typical Columns (daily OHLC time series):

- `Date` — trading date
- `Open` — opening index level
- `High` — intraday high
- `Low` — intraday low
- `Close` — closing index level
- `Adj Close` (if present) — adjusted close for corporate actions
- `Volume` (if present) — traded volume / turnover.
  Many public Nifty 50 historical files use these OHLC columns — see similar Kaggle daily datasets.

1.3 Coverage & context:

- Date range covered: 2015 — 2023 (daily trading sessions). The official NIFTY 50 index historical data is also available via Yahoo Finance and NSE Index historical data pages for reference.

1.4 Data quality note (typical):

- Expect trading-day gaps due to weekends/holidays (not "missing" in the financial sense).
- Check for duplicates or out-of-order dates; adjust timezone/format if needed.

2. Operations Performed (recommended EDA & analysis steps)

These mirror your insurance workflow but adapted for time-series finance data.

2.1 Data cleaning & exploration

- Read CSV and parse `Date` to `datetime`.
- Sort by `Date`.
- Check for duplicates and non-trading-day rows.
- Check columns & dtypes; ensure numeric conversion for OHLC and Volume.
- Fill/align `Adj Close` if present or use `Close` as main series.

2.2 Descriptive analytics & visuals

- Time-series line chart: `Close` over time (2015–2023).
- Year-wise returns bar chart (annual % returns).
- Histogram & boxplot of daily returns (distribution, skewness).
- Rolling statistics: 20/50/200-day moving averages & rolling volatility (20-day std).
- Drawdown plot (max drawdowns and recovery periods).
- Volume time-series and volume vs returns scatter.

2.3 Relationship / deeper analyses

- Daily log returns and distribution (normality checks).
- Autocorrelation (ACF/PACF) of returns.
- Seasonality checks: monthly average returns, day-of-week effects.
- Volatility clustering (GARCH-style patterns) and rolling volatility heatmap.
- Correlate index returns with major events (e.g., 2020 COVID shock) for event analysis.

3. Key Insights (example outcomes you'll likely get)

I'm writing these as expected, commonly observed results for Nifty 2015–2023 — exact numbers require running the code on the dataset (code included below).

3.1 Trend & returns

- The index shows a long-term upward trend from 2015—2023 with notable drawdowns around market-wide stress events (e.g., early-2020). (Historical index context: Yahoo Finance / NSE records).

3.2 Volatility & distribution

- Daily returns distribution is centered near zero with fat tails (occasional large moves); rolling volatility exhibits clustering (periods of high volatility).
- Rolling 20-day volatility highlights the highest turbulence windows (useful for risk-timing).

3.3 Seasonality & calendar effects

- Monthly average returns and day-of-week analyses commonly reveal slight seasonality (e.g., some months historically stronger). Verify on your file.

3.4 Moving averages / technical signals

- Crossovers (20/50/200-day MA) show regime shifts — can be used as simple trend filters or signal features for models.

3.5 Drawdowns & recovery

- Identify maximum drawdown periods and recovery lengths — critical for risk management and scenario design.

4. Recommendations (actionable analytics & business uses)

4.1 For quantitative research / ML models

- Create features: lagged returns, moving averages, rolling vol, volume changes, calendar features.
- Build models for short-term returns (classification: up/down) and forecasting (regression for next-day return). Validate with walk-forward cross-validation.

4.2 Risk management & stress testing

- Use empirical drawdowns to construct scenario-based VaR/stress tests.
- Monitor rolling volatility for dynamic risk limits.

4.3 Trading & strategy prototyping

- Test MA crossover and mean-reversion strategies using realistic transaction costs and slippage.
- Evaluate Sharpe, max drawdown, and CAGR over 2015–2023.

4.4 Further analyses

- Combine with constituent-level data or macroeconomic variables for richer factor models.
- Add news / event signals (earnings, policy changes) for event-driven studies.

5. Conclusion

The Kaggle Nifty 50 daily dataset (2015–2023) is a solid basis for time-series analysis, risk profiling, strategy prototyping, and predictive modeling. With the OHLC series and volume, you can compute returns, volatility, seasonality, and drawdowns and build both statistical and ML forecasting pipelines.