

Title: Mutual Fund Performance & Risk & Return Analysis (India) – Final Report

Objective:

This project aims to analyze and compare the historical performance and risk characteristics of selected Indian mutual funds using various fundamental financial metrics. The analysis includes calculating NAV trends, CAGR, Sharpe Ratio and Risk vs Return statistics over a 5-year period to evaluate and compare the funds risk-adjusted performance.

Business Value:

This project equips stakeholders with clear insights into mutual fund performance over time. It helps investors, analysts, and financial planners choose funds aligned with risk preferences and return goals.

Tools Used:

- PostgreSQL (for database and data merging)
- R and RStudio (for metric calculations, analysis, and visualization)
- Power BI (dashboard)
- Excel/CSV (original data source)

Data Source:

We collected mutual fund NAV data from the **Association of Mutual Funds in India (AMFI)** official website

Cleaned .csv files from 5 mutual funds:

- Axis Mid Cap
- SBI Large Cap
- ICICI Debt
- HDFC Balanced
- Nippon Small Cap

Methodology:

1. Data Collection:

- Selected top AMC names like SBI Mutual Fund, Axis Mutual Fund, HDFC Mutual Fund, etc.
- For each AMC, we chose equity schemes with long-term performance (e.g., Large Cap or Flexi Cap).

- Timeframe: **01 Jan 2020 to 31 Dec 2024** (max 5-year limit on AMFI).
- Data downloaded in Excel .xls format for each fund.

2. Data Cleaning (in Excel):

- Removed header rows and formatting issues.
- Kept only valid columns: **Date** and **NAV**.
- Standardized all column headers (e.g., renamed to Date, NAV).
- Added a new column **Fund_Name** manually to identify each fund.
- Added a new column **Fund_Name** manually to identify each fund.
- Saved all cleaned files in .csv format to use in PostgreSQL.

3. SQL Integration (PostgreSQL):

- CSV files were imported into PostgreSQL. Loaded each cleaned .csv file into a temporary table using PgAdmin's import tool.
- All individual NAV datasets were merged into a single mutual_fund_nav_all table. Inserted records from each file into a **final consolidated SQL table** named mutual_fund_data
- Fields: fund_name, nav, nav_date

4. Data Analysis in R


We used R for detailed metric calculation and visualization:

Metrics Calculated:

- **NAV Trend:** Net Asset Value recorded over time.
- **Daily Return:** % change in NAV from one day to the next.
- **CAGR (Compound Annual Growth Rate):**
Measures average yearly return, the annual growth rate of the NAV over the full time period.
$$\text{CAGR} = (\text{End_NAV} / \text{Start_NAV}) ^ { (1 / \text{Years}) } - 1$$
- **Sharpe Ratio:**
Risk-adjusted return calculated as
$$\text{Sharpe Ratio} = \text{Mean Daily Return} / \text{Standard Deviation of Return}$$

- **Standard Deviation (Volatility):** Indicates risk.
- **Risk vs Return:** Scatter plot comparing risk and return.
- **Annualized Return and Volatility:**
Based on 252 trading days:
Annual Return = Mean Daily Return × 252
Annual Volatility = SD of Daily Return × sqrt(252)
- **Summary Statistics:** Mean, Median, Min, Max NAVs for each fund.

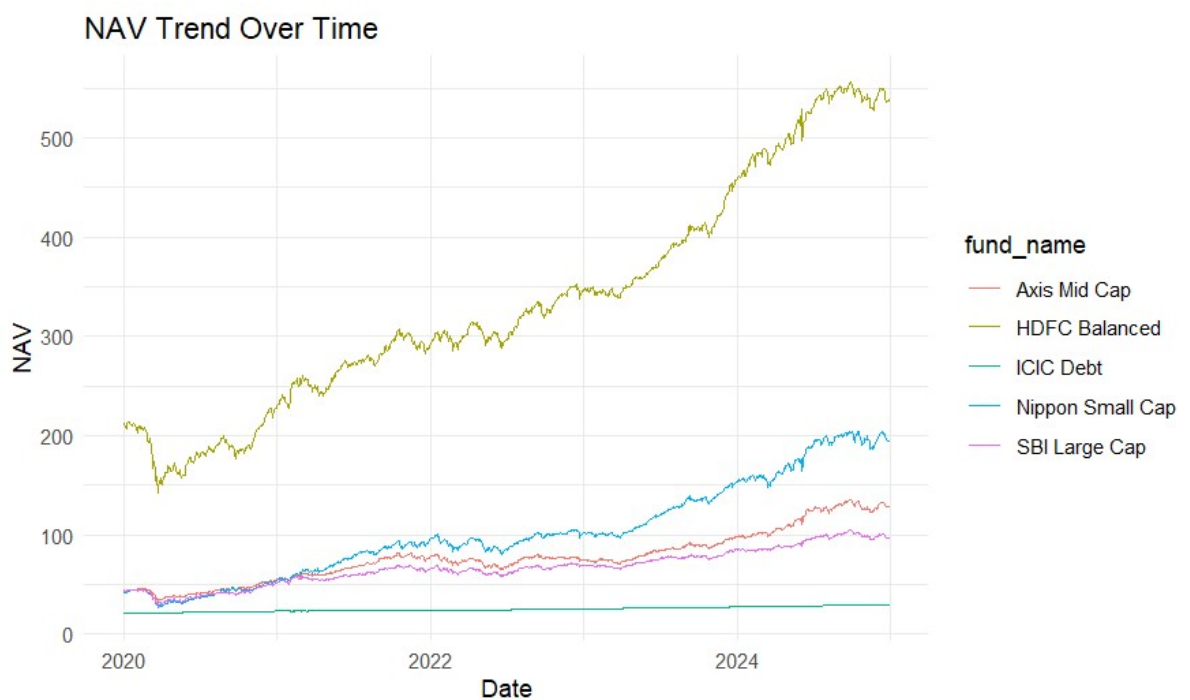
Simulated Data Cleaning in R:

- Checked for NA values.
-  Converted date and numeric columns to proper formats.

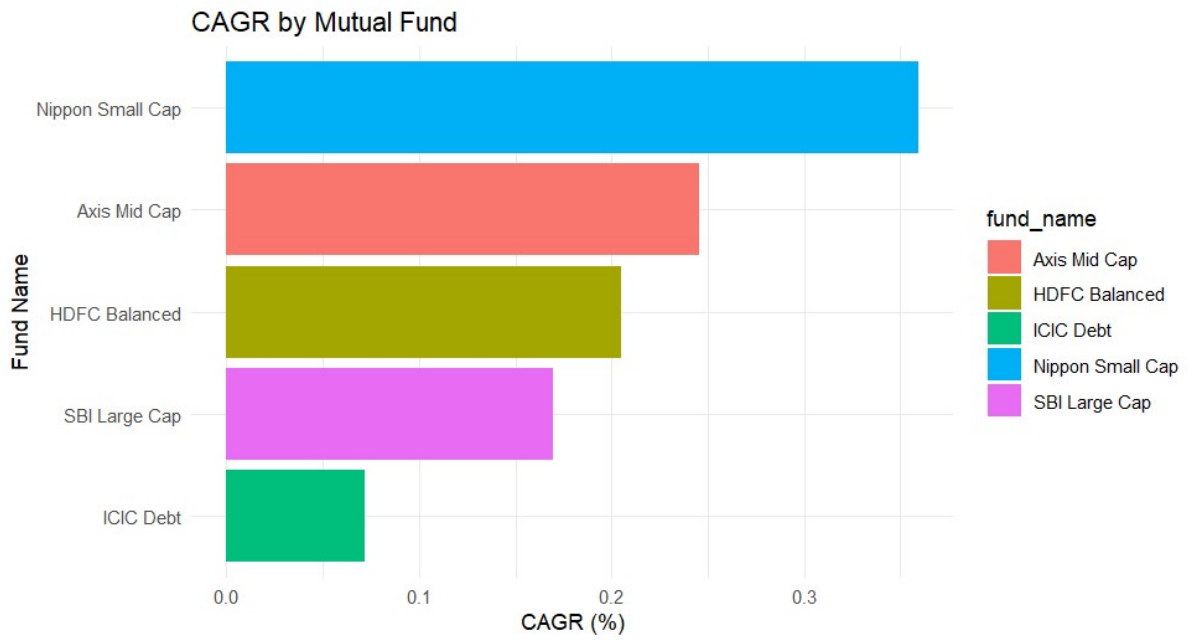
5. Visualizations Created:

All key metrics were visualized using ggplot2 and exported as .png:

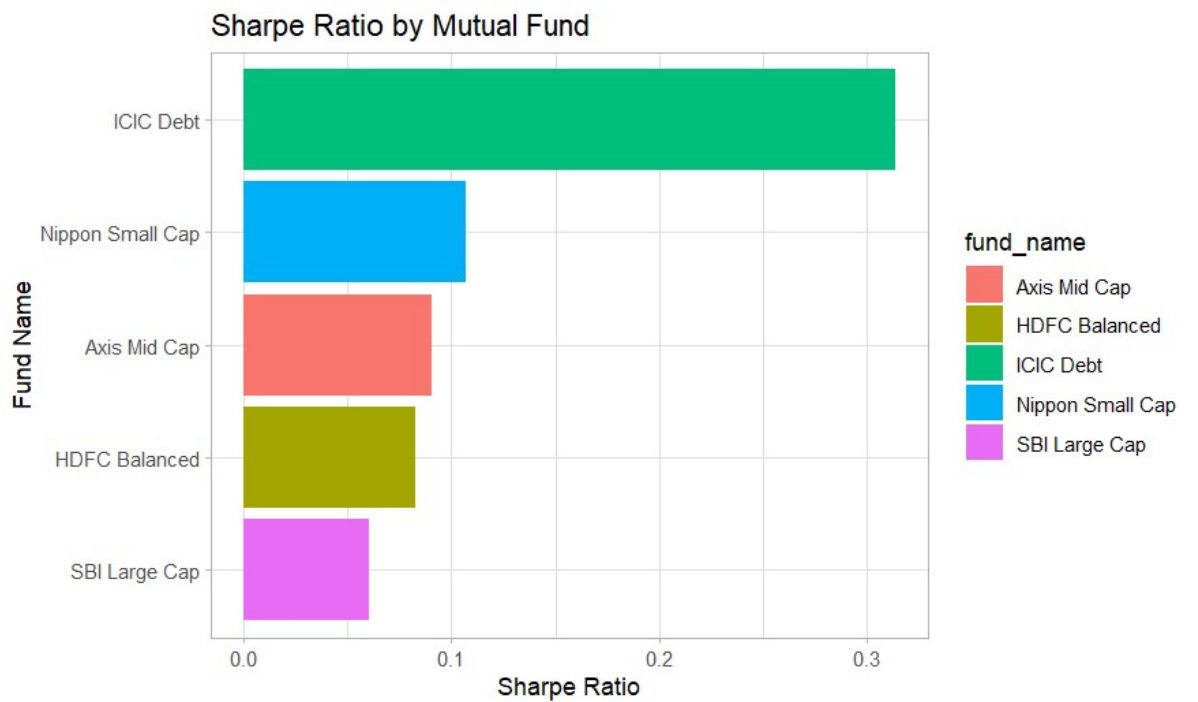
- NAV Trend over time per fund Line Chart (Fund vs Time)



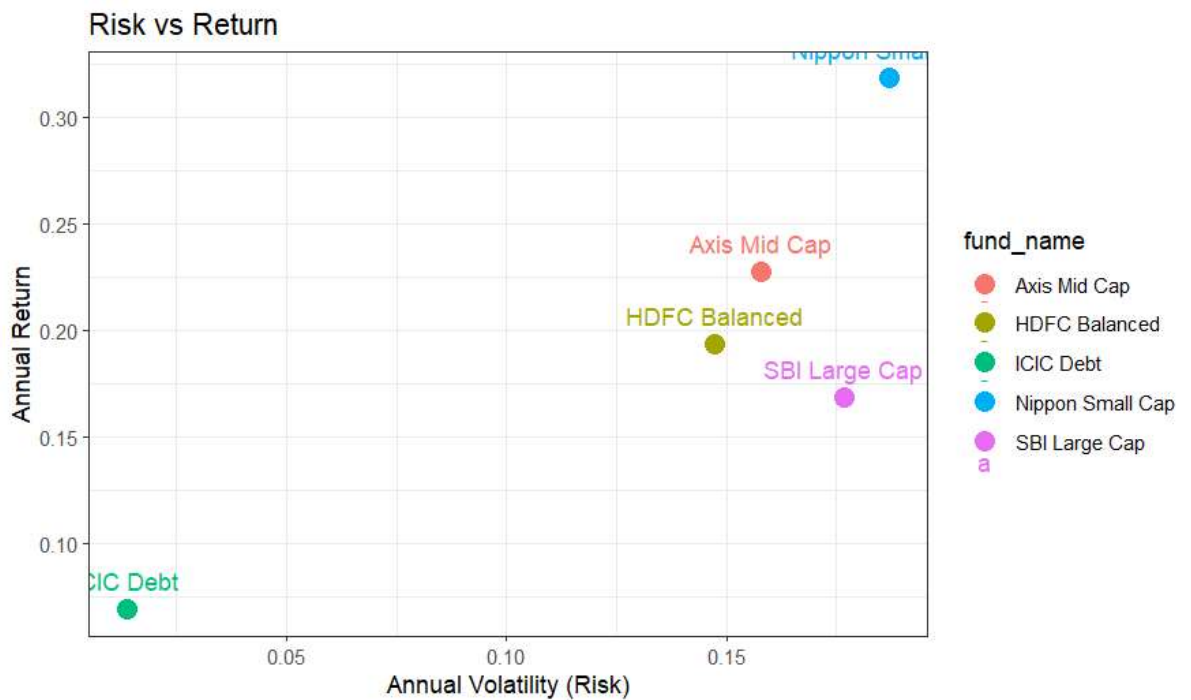
- CAGR comparison bar chart



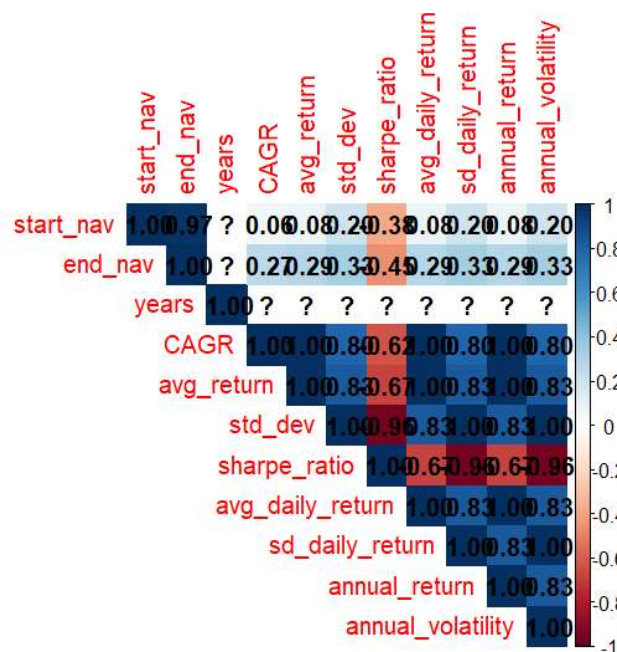
- Sharpe Ratio comparison



- Risk-Return scatter plot



- Relationship between Numeric Metrics

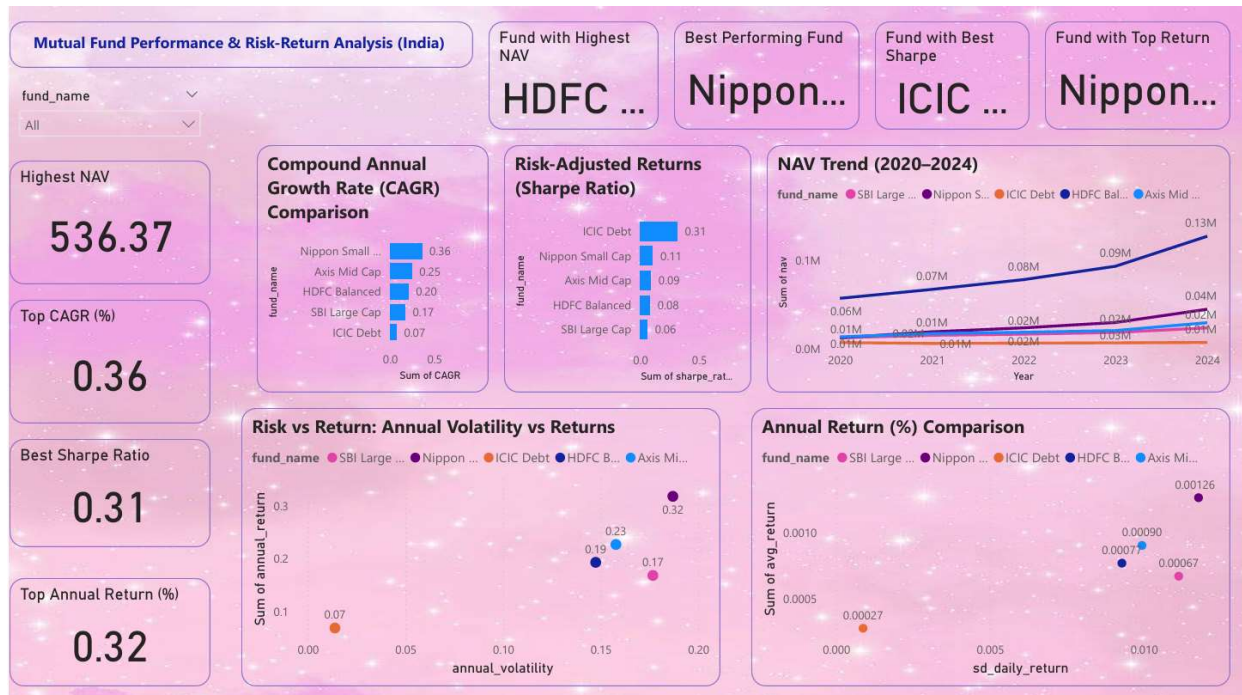


6. Power BI Dashboard

To visualize the key financial insights from our mutual fund analysis, an interactive Power BI dashboard was created using the final calculated metrics such as CAGR, Annual Returns, Sharpe Ratio, Expense Ratios, and NAV trends. The dashboard includes:

- **KPIs** highlighting the top-performing fund based on NAV, CAGR, Return, and Sharpe Ratio.
- **Charts** such as Risk vs Return (scatter plot), NAV Trend (line chart), and bar charts comparing CAGR and Sharpe Ratios.
- A **Slicer** for filtering data by fund name.

This dashboard enables stakeholders to quickly assess fund performance, volatility, and efficiency over the years 2020–2024



7. Exported Outputs:

- Final metrics table (final_metrics.csv)
- PNG visualizations
- Report
- Power BI dashboard

Suggestions & Key Insights:

- **ICICI Debt** offered lower returns with the least volatility — ideal for risk-averse investors.
- Funds like **Axis Mid Cap** and **Nippon Small Cap** showed high CAGR but also higher volatility.
- **Sharpe Ratio** revealed **HDFC Balanced** had the best risk-adjusted performance.

- **SBI Large Cap** maintained consistent returns with moderate risk.
- **Investors** should balance return expectations with volatility before selecting a mutual fund.

Conclusion:

This project demonstrates the complete process of financial analysis using Excel, SQL, Power BI, R, and data visualization. By analyzing CAGR, Sharpe Ratio, and volatility, we can guide investors toward informed decisions based on their risk tolerance and return expectations.