DOCUMENTATION

FINAL PROJECT REPORT: Mutual Fund Market Intelligence & AMC Performance Dashboard

Sector: Financial Services / Asset Management (Wealth Management)

Team Name: Section E - Group 3

- **Team Members:** Ved Pawar (Lead),Ranajeet Roy , Yatin Singh, Pratham Malhotra, Abhijeet Kumar, Pushkar Jain.
- **Institute:** Newton School of Technology

2. Executive Summary

This report details the transformation of "India's Ultimate Mutual Fund Dataset" into a strategic decision-making tool. By addressing a "messy" data environment of 16,000+ schemes, our team engineered a dashboard to identify market leaders by AUM and evaluate accessibility. Key findings highlight a high market concentration among legacy AMCs and a significant retail preference for liquid, low-barrier equity schemes.

3. Sector & Business Context

TheIndian Mutual Fund industryis characterized by high complexity. With thousands of schemes across Equity, Debt, and Hybrid categories, investors face "Analysis Paralysis." We chose this sector to provide transparency in how Asset Management Companies (AMCs) manage capital and how accessible these investments are to the common retail investor.

4. Problem Statement & Objective

FormalDefinition: "Which mutual fundsand Asset Management Companies (AMCs) should investors focus on based on assets managed, fund categories, accessibility, and overall market presence?"

- **Project Scope:** Comprehensive analysis of all AMCs listed in the provided CSV.
- **Success Criteria:** Development of an interactive Google Sheets dashboard that reduces fund comparison time by 80%.

5. Data Description

- **SourceCitation:** India's Ultimate Mutual Fund Dataset (`mutual_fund_data.csv`).
- **Data Structure:** 16 columns including `AMC`, `Scheme_Category`, `NAV`, `Average_AUM_Cr`, and `Scheme_Min_Amt`.
- **Data Size:** 16,319 records.
- **Limitations:** The dataset is a point-in-time snapshot and does not account for real-time daily NAV fluctuations.

6. Methodology (Data Source + Cleaning Steps)

AllengineeringwasconductedinGoogleSheets percapstone requirements:

- **Cleaning `Scheme_Min_Amt`:** Removed non-numeric strings (e.g., "Rs. 5000 and multiples...") to create a filterable numeric column.
- **Handling Nulls:** Identified and addressed missing values in `Average_AUM_Cr` to prevent skewed averages.
- **Standardization:** Converted `Launch_Date` into uniform Date objects for age analysis.
- **Feature Engineering:** Created "Fund Age Groups" and "NAV Buckets" for categorical segmentation.

- **Assumptions:** Assumed that schemes with null AUM were either recently launched or inactive during the reporting quarter.

7. KPI & Metric Framework

- **Total AUM (Cr):** Sum of assets to measure market power.
- **Average NAV:** Pricing level comparison across AMCs.
- **Scheme Count:** Variety of options offered by an AMC.
- **Avg. Min Investment:** Metric for retail accessibility.

8. Exploratory Data Analysis (EDA)

- **Trend:** A surge in Open-Ended Equity schemes in recent years.
- **Distribution:** `Scheme_Min_Amt` is heavily clustered around ₹500 - ₹5,000.
- **Comparison:** Legacy AMCs hold 3x more AUM on average compared to AMCs younger than 5 years.

9. Advanced Analysis

- **Segmentation:** Grouped AMCs into "Tier 1" and "Tier 2" based on total asset control.
- **Risk Profiling:** Analysis of NAV variance within the "Debt Scheme" vs "Equity Scheme" categories.

10. Dashboard Design

- **Objective:** To answer the core business question through interactive filtering.
- **View Structure:** KPI tiles, AUM Bar Charts, and Category Donut Charts.
- **Filters:** AMC Name, Scheme Type, Category, and Fund Age Group.

11. Analytical Insights (The "Meat")

1. **Market Concentration:** The top 5 AMCs (Aditya Birla, HDFC, Mirae, etc.) control over 40% of the total recorded AUM.
2. **Asset Class Preference:** Equity schemes dominate the market by volume, representing high retail interest.
3. **Liquidity:** Open-ended schemes comprise over 90% of the market, showing investor preference for liquidity.
4. **Accessibility:** 60% of schemes have entry points under ₹5,000, confirming a shift toward mass-market retail participation.
5. **Maturity Advantage:** "Legacy" funds (15+ years) consistently maintain higher AUM stability compared to new launches.

12. Recommendations (The "Value")

For Retail Investors:

- Focus on **Equity schemes with a high AUM-to-Scheme ratio**, as this indicates strong investor trust and fund house stability.
- Prefer funds managed by **legacy AMCs** that demonstrate consistent asset accumulation over time.
- Use AUM as a **risk-mitigation proxy**, especially when selecting long-term investment products.

For AMCs:

- Launch **Hybridmutualfundproducts** with **lower minimum investment thresholds (₹500)** to attract first-time and Gen-Z investors.
- Position hybrid schemes as a **balanced entry product**, combining growth potential with risk control.
- Leverage dashboard insights to identify **under-served investment segments** and design targeted offerings.

For Advisors:

- Use the **Fund Age filter** to recommend seasoned and time-tested schemes for conservative, low-risk clients.
- Combine fund maturity with **NAV stability and AUM size** to strengthen client trust and portfolio reliability.
- Avoid recommending newly launched schemes for risk-averse investors unless supported by strong AMC credentials.

13. Impact Estimation

Efficiency:

- Automated Pivot Tables reduce **manualresearchandcomparisontimefromhoursto seconds**.
- Interactive filters eliminate repetitive analysis tasks and enable faster decision-making.
- Standardized KPIs allow users to compare funds consistently across multiple dimensions.

Risk Reduction:

- Categorical filtering highlights **AUM and NAV outliers**, preventing unsuitable fund selection.
- Risk exposure is reduced by clearly differentiating between **Equity,Debt,andHybrid schemes**.
- Fund age and stability indicators help avoid investments in **volatile or unproven schemes**.

14. Limitations

Data Gaps:

- Approximately **50% of schemes lacked recent AUM data**, limiting precise market share estimation.
- Missing values may underrepresent the performance of newer or temporarily inactive schemes.

Dynamic Pricing:

- The dataset represents a **static snapshot** and does not capture real-time NAV fluctuations.
- Short-term market volatility and recent economic events are not reflected in the analysis.

15. Future Scope

Forecasting:

- Incorporating **historical NAV time-series data** would enable trend analysis and future growth predictions.
- Forecast models could support proactive investment planning and risk anticipation.

Integration:

- Linking **real-time APIs** would allow live NAV updates and dynamic dashboard refreshes.
- Integration with market indices could enhance benchmarking and performance comparison.

16. Conclusion

This project successfully transformed a raw CSV into a **professional market intelligence tool** through structured cleaning, analysis, and visualization.

We have demonstrated that while the mutual fund market is vast and complex, **data-driven filtering and segmentation** can significantly simplify fund selection.
The dashboard clearly highlights the **dominance of established AMCs**, retail accessibility trends, and risk differentiation across categories, enabling informed decision-making for multiple stakeholders.

17. Appendix

Data Dictionary:

- Detailed mapping of dataset headers such as **Scheme_Code, ISIN, AMC, NAV, and AUM**.

- Ensures transparency and reproducibility of analysis.

Regex Logic:

- Regular expressions were used to clean and standardize **Scheme_Min_Amt** values.
- This conversion enabled accurate numeric filtering and accessibility analysis.

18. Contribution Matrix (Who did what)

Team Member	Dataset	Cleaning	Analysis	Dashboard	Report	Overall Role
Ranajeet Roy	✓	✓	✓	✓		**Project Lead**
Yatin Singh	✓		✓	✓		**Strategy Lead (Value)**
Pratham Malhotra	✓	✓	✓			**Data Lead (Methodology)**
Abhijeet Kumar	✓	✓	✓	✓		**Analytics Lead (Meat)**
Ved Pawar	✓		✓	✓		**Dashboard Lead**

Pushkar Jain	✓			✓	✓	**Quality Lead (Report)**

Declaration: We confirm that these contributions are accurate and verifiable through Google Sheets version history.

Team Signature Block:
Ranajeet Roy | *Yatin Singh* | *Pratham Malhotra* | *Ved Pawar* | *Abhijeet Kumar* | *Pushkar Jain*