Morningstar European Funds

Team -8 Data mining project

G31327889 – Abishek Chiffon Muthu Raja

G32021287 – Nadadhur Srivallabh Siddharth

G41309434 – Pavani Chella

Problem statement

We've chosen "Morningstar European Funds" as our dataset for the project. The primary objective of this project is to conduct an extensive comparative analysis of Mutual Funds and ETFs in Europe and America. Morningstar specializes in compilation and analysis of funds, stocks and general market data. Morningstar ratings are generally considered to be high quality in the financial industry. The dataset is comprised of 57,603 Mutual Funds and 9,495 ETFs, which involves working with 132 independent variables. The project aims to Analyze the returns, volatility, and financial ratios (including price/earning, Treynor and Sharpe ratios, alpha, and beta) to assess the performance and risk associated with Mutual Funds and ETFs. The final goal is to build predictive models to forecast future fund performance based on historical data and identified trends.

Modeling methods :

Linear Regression: Used to predict a quantitative response. It can be applied if there is a linear relationship between the predictors and the fund performance.

Random Forest: An ensemble learning method that can be used for regression tasks to predict fund performance.

Gradient Boosting: Builds an additive model in a forward stage-wise fashion; it allows for the optimization of arbitrary differentiable loss functions.

Long Short-Term Memory Networks (LSTM) (Deep learning), ARIMA (AutoRegressive Integrated Moving Average) (Time Series), GARCH (Generalized Autoregressive Conditional Heteroskedasticity)

Econometric Models: Vector Autoregression (VAR)

SMART Questions

Specific:

1. What is the average ongoing cost of Mutual Funds with a 5-star rating?
2. How many funds have achieved a sustainability score above 80 and also have a top quartile performance rating?
3. How do funds with a high environmental score (above 70) compare in terms of ROA, ROE, and ROIC to funds with a low environmental score (below 30)?How do funds with a high environmental score (above 70) compare in terms of ROA, ROE, and ROIC to funds with a low environmental score (below 30)?

Measurable:

4. What is the average fund size for each fund category?
5. How has the average Price to Earnings (P/E) ratio changed over the past five years across different equity styles?
6. What is the correlation between the sustainability score and fund trailing returns over 1, 3, and 5 years?
7. Calculate the average fund return for each risk rating category during the last quarter.
8. What percentage of funds have a credit quality of 'AAA'?

Achievable:

9. Identify the top 10 funds in terms of 5-year trailing return within each risk rating category.

Among funds with a "Gold" analyst rating, which fund categories are most prevalent?

10. For funds with a high risk rating, what investment strategies are commonly employed?
11. Which fund category has shown the most improvement in average return over the past three years?

Relevant:

12. In sectors affected by COVID-19, how did the funds' performance change from Q1 2020 to Q2 2020?
13. What is the average ROE for funds in the "Consumer Defensive" sector, and how does it compare to the overall average?
14. How do funds with a low risk rating and high sustainability score perform in terms of 10-year trailing return?
15. In the context of the recent emphasis on ESG investing, how do funds with high sustainability scores compare in terms of management fees to those with low sustainability scores?

Time-Bound:

16. What were the top-performing funds in the last quarter, and have they consistently outperformed over the last three years?
17. Over the past five years, how has the asset allocation across stocks, bonds, and cash changed in funds within the “Balanced” category?

Predictive Analysis:

18. Predictive Modeling: Can we build a predictive model to estimate future returns or risk levels based on the available data?

Niche and Specific Questions:

19. Outliers and Extremes: Are there any funds that stand out as outliers in terms of returns, asset allocation, or financial ratios?
20. Size vs. Performance: Is there any correlation between the size of the fund (in terms of total net assets) and its performance?
21. Effect of Sectors on Returns: Does investing in specific sectors tend to yield higher returns or lower risk?

Github team repo :

https://github.com/abishekchiffon/Risk_prediction_Mutual_funds

Source of data :

European Funds dataset from Morningstar (kaggle.com)

Data has been scraped from the publicly available website of Morningstar UK website, https://www.morningstar.co.uk