

# 1. Mutual Fund Evaluation

## 1.1. Data

I collected all the mutual funds under "Sector equity" on Fidelity, on Jan 9 2026. The full dataset is "funds\_merged\_all.csv". It contains 536 funds.

Full list of features:

```
'Name', 'Morningstar Category', 'YTD (Daily)', '1 Yr', '3 Yr', '5 Yr',
'10 Yr', 'Life of Fund', 'Expense Ratio - Net', 'Expense Ratio - Gross',
'Morningstar- Overall', 'Morningstar- 3yrs', 'Morningstar- 5yrs',
'Morningstar- 10yrs', 'Morningstar Category Risk', 'Standard Deviation',
'3 Year Sharpe Ratio', 'Beta', 'R2', 'Manager Tenure', 'Inception Date',
'Assets (Millions)', 'Turnover Rates', 'Minimum Investment',
'Load (Y/N)', 'NTF', '1-year Rank', '3-year Rank', '5-year Rank',
'10-year Rank', 'Average Maturity', 'Average Duration', '7-Day Yield',
'30-Day Yield', 'Last Dividend', '1 Month', '3 Month', 'YTD', 'NAV ($)',
'NAV Change ($)', 'NAV Change (%)', '30-Day Yield (%)',
'7-Day Yield (%)'
```

See "Help Me Read This Table.pdf" for descriptions of the features.

We use the below features in our comparisons:

```
'Expense Ratio - Net', '3 Year Sharpe Ratio', 'Standard Deviation', '3 Yr', 'Beta',
'Manager Tenure', 'Inception Date', 'Assets (Millions)', 'Turnover Rates', 'Load (Y/N)',
'NTF', '3-year Rank'
```

### References for the selected features

'Expense Ratio - Net', 'Load (Y/N)', 'NTF': Barber, Brad M., Terrance Odean, and Lu Zheng. "Out of sight, out of mind: The effects of expenses on mutual fund flows." *The Journal of Business* 78.6 (2005): 2095-2120.

'3 Year Sharpe Ratio', 'Standard Deviation': Simons, Katerina. "Risk-adjusted performance of mutual funds." *New England Economic Review* 9 (1998): 33-48.

'3 Yr': Carhart, Mark M. "On persistence in mutual fund performance." *The Journal of Finance* 52, no. 1 (1997): 57-82.

'Beta': Fama, Eugene F., and Kenneth R. French. "Michael C. Jensen's empirical work." *Journal of Financial Economics* (2025): 104119.

'Manager Tenure': Chevalier, Judith, and Glenn Ellison. "Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance." *The Journal of Finance* 54, no. 3 (1999): 875-899.

'Inception Date': Evans, R.B., 2007. The incubation bias. Available at SSRN 972761.

'Assets (Millions)': Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik. "Does fund size erode mutual fund performance? The role of liquidity and organization." *American Economic Review* 94, no. 5 (2004): 1276-1302.

'Turnover Rates': Shen, Ke, Lin Tong, and Tong Yao. "Heterogeneous turnover-performance relations." *Journal of Banking & Finance* 124 (2021): 106054.

## 1.2. Comparison (To-Do)

Similar to property values comparison, we compare the mutual funds in pairs.

Specifically, the two funds to compare are from the same Morningstar category. Same as property value evaluations, we make sure in each pair, one fund does not "weakly dominate" the other, so that we do not have very straightforward comparisons.

- The difference in each feature should be at least  $0.2 * \text{std}$  to be considered different

The matched pairs are in "mutual\_funds\_pairs.csv". We have 11,951 pairs.

**For the below comparisons, we can for now just randomly sample 1,000 pairs and see the results.**

### 1.2.1. Comparison 1

In this comparison, we provide LLM/humans with the below variables:

```
'Expense Ratio - Net', '3 Year Sharpe Ratio', 'Standard Deviation', '3 Yr', 'Beta',
'Manager Tenure', 'Inception Date', 'Assets (Millions)', 'Turnover Rates', 'Load (Y/N)',
'NTF'
```

And we ask LLM which fund would it invest in and guess which one has a higher return in January 2026.

For now, we can use the '3-year Rank' in the data as ground truth.

### 1.2.2. Comparison 2

In this comparison, we provide LLM/humans with the below variables:

```
'Expense Ratio - Net', '3 Year Sharpe Ratio', 'Standard Deviation', '3 Yr', 'Beta',
'Manager Tenure', 'Inception Date', 'Assets (Millions)', 'Turnover Rates', 'Load (Y/N)',
'NTF'
```

And we ask LLM to guess which one has a higher increase in asset under management (AUM) and guess has a higher increase in AUM in January 2026.

For now, we can use the '3-year Rank' in the data as ground truth.