ETF Alternative Datasets for Systematic Trading, Investment Screening and Risk Measurement

COMPANY

ETFLogic brings a new level of transparency to the fast-growing world of ETFs. We help our clients grow their assets and their brand through our tools and analytics. The Insights web-platform provides actionable results tailored for ETF cost and risk measurement. We were founded in 2017 by experienced quants, traders and market-makers. Our clients are ETF issuers, buy-side traders, advisors, portfolio strategists, asset managers and corporate investor relations.

DATASET OVERVIEW

ETFs have experienced 22% annualized growth for much of the last decade. That growth is expected to drive ETF assets from $5 trillion today to $30 trillion by 2030. We combine our expertise in the creation, redemption and market-making processes and collation of ETF fund data from dozens of data sources to drive our signals and feature sets. Below you will find a brief description of our main alt-data and derivative data products. Typical use-cases range from systematic and quant-driven strategies, actionable signals for tactical trading to monitoring for pre-trade analysis, investment due-diligence and risk measurement.

FMD - Single Stock Flow Monitor Dataset - using daily ETF flows from creations and redemptions

Daily, single stock flow alpha signals for 3000+ US tickers due to ETF primary market activity. Incorporates creation, redemption and rebalancing activity. Use case: Systematic trading signal or individual stock flow monitoring. **History**: 5+ years of backfilled history. 1+ year of daily generated history.

J-FMD – Japan - Single Stock Flow Monitor Dataset

Our FMD signal also exists for the Japan market – where we produce daily signals on 1000+ Japanese single stock equities. **History** begins in April 2018.

S-FMD - Sector Flow Monitor Dataset

Sector scores in S-FMD are aggregated from the FMD data described above. Features are scored across 11 US sectors. Use case: Sector flow visualization and monitoring, sector rotation strategies, tactical risk management.

EFF - ETF Fundamental Financial Statement Features - based on the latest company financial statements

ETFLogic's ETF Fundamental Features (EFF) produces a set of 23 daily features for over 500 US-Listed ETFs. The ETF universe is constrained to US equity funds. The features are backwards-looking fundamental single-stock metrics that are averaged up to the ETF using our basket weights database. The fundamental metrics are balance-sheet, income statement and cash-flow ratios such as Debt-To-Equity, Price-to-Sales and Quick Ratios. Each feature in the dataset incorporates the latest underlying fundamentals as reported to the SEC and is normalized for different accounting standards. **History**: 5+ years

ETFLogic | www.etflogic.io
info@etflogic.io | +1.347.948.5155

EEE - ETF Earnings Consensus Estimates - based on latest company earnings consensus estimates

ETFLogic's ETF Earnings Estimates (EEE) produces earnings estimate features for over 500 US-Listed ETFs. These estimates are generated daily and incorporate the latest consensus estimates which, in turn, are generated from the most recent equity analyst P/E forecasts. The ETF universe is constrained to US equity funds. The features are forwards-looking PE ratios that are averaged up to the ETF using our basket weights database. . **History**: 5+ years

ETF Factor Exposures and Scores - to compare exposures across US ETFs for risk measurement

Daily risk factor exposures file for all US listed ETFs, comprised of close to 100 different factors from our proprietary factor database. Use cases: portfolio optimization, ranking and qualitative description of ETFs, monitoring style drift & visualization for "style box" analysis. . **History**: 5+ years

ETF Peer Group Taxonomy and Classification - for risk measurement and ETF comparisons

Labeling and taxonomy grouping of US listed ETFs based on their qualitative and quantitative characteristics. Use Cases: descriptive statistics, monitoring. Example: EWJ and HEWJ are both Japan-focused ETFs but would fall into different groups because one has a currency-hedged component.

ETFLogic Liquidity Metrics – advanced impact cost and implied liquidity estimates on ETFs

ETFLogic's platform metrics, TruCost and TruLiquidity, are also available in an easy-to-parse daily file. TruCost is a 5-star ranking that calculates the lower of ETF and basket costs and makes a comparison to peer ETFs. TruLiquidity considers basket implied liquidity using market impact cost (share volume, volatility, urgency), creation fees, redemption fees, regional taxes and stamp duties. We generate daily cost runs for US listed ETFs and provide highlights on recent ETF block trades versus fair-value, primary trading volumes, etc. Use case: pre-trade, due-diligence, market monitoring.

Microstructure Tickdata Features - ETF analytics on US Trades and Quotes data for investment screening and due-diligence

Trade and quote analytics across daily, weekly, monthly time ranges. Calculation examples are average bid-ask spread, average spread at time of day (morning, midday, afternoon), top-of-book average sizes, top-of-book exchanges, exchanges at top % of time, large trade monitoring, NAV vs Close metrics. Use case: ETF screening, monitoring, pre-trade and due-diligence reports

ETF ESG Ratings - screens funds for multiple warnings flags and ESG features

Not all ESG ETFs are created equally. We use Arabesque ESG ratings on global single stocks and percolate these ratings up to the global-equity ETFs. These ratings provide overall ESG scores and highlight individual warning flags for constituent stocks. The Arabesque methodology and ratings definitions can be found online:
https://arabesque.com/docs/sray/SRay_Methodology.pdf.

2019 DATA EXPANSION

- Global Daily Fund Holdings for ETFs (adding Europe and Asia)
- Mutual Funds Holdings

DATA DICTIONARY

A few of our daily datasets and their associated fields, field types and field descriptions are detailed below. Rows are timestamped at creation time. Updated versions of these data dictionaries may exist as we continue our support and buildout for global ETFs and single stocks.

ETFLogic Liquidity Metrics		
Field	Type	Description
ticker	string	The ETF Ticker
nav	float	The last closing NAV of the ETF
proxy_used	boolean	Whether the implied liquidity of the ETF is based on the actual basket impact or the impact of a proxy hedge
breakdown	boolean	For fund-of-fund ETFs where the basket component is an ETF, was that ETF further broken down before implied liquidity calculation
participation_rate	float	The participation rate when trading the ETF Basket or Proxy Hedge used in the Implied Liquidity Calculation. Defaulted to 10%
impliedNotional_10bps	long	The implied notional of the ETF basket if we are willing to pay up to 10bps. Costs include TCA-measured impact, creation fees, and individual country stamp duties and taxes
impliedNotional_20bps	long	Same as above, for 20bps impact. Used to overcome high stamp duties for international baskets.
impliedNotional_30bps	long	Same as above, for 30bps impact. Used to overcome high stamp duties for international baskets.
aum	long	The shares outstanding of the ETF multiplied by the NAV
tradability_score	float	A normalized trading score (over 100) versus peer group names.
adv	long	20-Day Average Trading Volume
spread	float	20-Day Average Bid-Offer Spread
assetclass	string	Asset Class of the ETF
geography	string	Geographic Exposure of the ETF Basket
legalstructure	string	ETF Legal Structure

Microstructure Tickdata Features		
Field	Type	Description
spread_bps	float	ETF secondary market, Average of 1-minute sample spread vs mid-quote
spread	float	ETF secondary market, Average of 1-minute sample spread
top_of_book_bid	float	ETF secondary market, 1-minute samples
top_of_book_ask	float	ETF secondary market, 1-minute samples
bid_offer_size-ratio	float	ETF secondary market, top_of_book_bid / top_of_book_ask
spread_bps_0930_1000	float	ETF secondary market, Spread in Basis Points, 1-minute samples between two times of the trading day
spread_bps_1000_1530	float	ETF secondary market, Spread in Basis Points, 1-minute samples between two times of the trading day
spread_bps_1530_1600	float	ETF secondary market, Spread in Basis Points, 1-minute samples between two times of the trading day
nav_vs_close	float	ETF vs Published Nav, Published NAV - Official Close Price
nav_vs_last_mid	float	ETF vs Published Nav, Published NAV - Calculated Last Quote Mid-Point
bid_from_inav_avg	float	ETF vs Basket Fair Value, 1-minute sampled averages of Best Bid-Fair Value
ask_from_inav_avg	float	ETF vs Basket Fair Value, 1-minute sampled averages of Best Offer-Fair Value

FMD - Flow Monitor Dataset		
Field	Type	Description

valid_asof	timestamp	Timestamp in EST timezone when features were built and updated to ETFLogic database. Please see Timing sheet.
ticker	string	
fmdscore1	float	Short term normalized flow score with less than 1-week horizon
fmdscore2	float	Medium term normalized flow score with 1 to 2-week horizon
fmdscore3	float	Longer term normalized flow score with 3-week horizon

EFF - ETF Fundamental Financial Statement Features

Field	Type	Description
ticker	string	
asset_turnover	float	Total Revenue TTM ETF / Total Asset Avg TTM ETF
current_ratio	float	Total Current Asset Avg TTM ETF / Total Current Liability Avg TTM ETF
debt_ratio	float	Total Liability Avg TTM ETF / Total Asset Avg TTM ETF
div_yield	float	Common Stock Dividend Paid TTM ETF / Market Value
ebitda_margin	float	EBITDA TTM ETF / Total Revenue TTM ETF
gross_margin	float	Gross Profit TTM ETF / Total Revenue TTM ETF
inventory_turnover_cogs	float	COGS TTM ETF / Inventory Average TTM ETF
inventory_turnover_sales	float	Total Revenue TTM ETF / Inventory Average TTM ETF
lt_debt_to_equity	float	Total LT Debt Avg TTM ETF / Total Shareholder Equity Avg TTM ETF
operating_cash_flow_margin	float	Cash Flow from Operations Activity TTM ETF / Total Revenue TTM ETF
operating_margin	float	Operating Income TTM ETF / Total Revenue TTM ETF
payout_ratio	float	Common Stock Dividend Paid TTM ETF / Net Income or Loss Available to Shareholders TTM ETF
price_to_book	float	Market Value / Total Shareholder Equity Avg TTM ETF
price_to_earning_basic	float	Market Value / Net Income or Loss Available to Shareholders Basic TTM ETF
price_to_earning_diluted	float	Market Value / Net Income or Loss Available to Shareholders Diluted TTM ETF
price_to_free_cash_flow	float	Market Value / Free Cash Flow TTM ETF
price_to_sales	float	Market Value / Total Revenue TTM ETF
profit_margin	float	Net Income or Loss Available to Shareholders / Total Revenue TTM ETF
quick_ratio	float	Total Liquid Assets Avg TTM ETF / Total Current Liability Avg TTM ETF
receivable_turnover	float	Total Revenue TTM ETF / Receivables Total Avg TTM ETF
res_dev_intensity	float	Research and Development Expense TTM ETF/ Total Revenue TTM ETF
return_on_asset	float	Net Income or Loss Available to Shareholders / Total Asset Avg TTM ETF
return_on_equity	float	Net Income or Loss Available to Shareholders / Total Shareholder Equity TTM ETF

EEE - ETF Earnings Consensus Estimates

Field	Type	Description
ticker	string	
price_to_earning_1q_high	float	Forward consensus PE, also available 2, 4 and 8 quarters out
price_to_earning_1q_low	float	Forward consensus PE, also available 2, 4 and 8 quarters out
price_to_earning_1q_mean	float	Forward consensus PE, also available 2, 4 and 8 quarters out
price_to_earning_1q_median	float	Forward consensus PE, also available 2, 4 and 8 quarters out