Similar ETPs

Summary

This function provides a list of Exchange Traded Products ("ETPs") that is generated by XTF, a third-party independent research provider. XTF uses a proprietary algorithm that compares all ETP products in pairs within each asset class and generates a list of similar ETPs. The "Similar ETPs" algorithm produces highly relevant and actionable results that allow the user to see a list of similar products and be able to compare them. XTF uses a proprietary algorithm to develop a list of potentially similar ETPs, updated weekly, based on a number of financial characteristics called investment objectives. XTF categorizes the ETPs according to their investment objectives as stated in each fund's prospectus. XTF's "Similar ETPs" algorithm utilizes the transparency in ETPs by looking primarily for ETPs with similar investment objectives and that have significant overlap of constituents. Similar ETPs may tend to be highly correlated to each other in terms of performance – and XTF does take correlation into account, but this algorithm places higher emphasis on transparency (exposure) and investment objectives that overlap rather than correlation because the latter can be arbitrary. XTF first ranks ETPs by the sum of fractional overlap of underlying constituents and investment objectives and then restricts those within well-defined investment objectives which tend to produce highly relevant results.

The process involves every ETP being compared to every other ETP in pairs to determine the level of similarity using the methods below in sequence.

Hard Restrictions

The following are all the restrictions on what subset of active ETPs can be similar to a given ETP or overall restrictions of the system. If any of these is violated, no similarity is possible.

- Comparison restricted to ETPs listed in the same country.
- Pairs of similar ETPs that have transparency (component weights or proxies) must have a minimum exposure overlap of 20% in order to be considered as similar. The percentage was arrived at by evaluating various thresholds and where the majority had significance. The exposure overlap of pairs of ETPs is stored in the SimilarETPsStaging_OverlapExposure table.
- When ETPs are being compared in pairs, if one ETP has transparency into the underlying holdings and the other does not they cannot be considered similar.
- For pairs of non-transparent ETPs (neither has exposure that can be determined) exposure overlap cannot be considered. They can be considered as similar using objectives and correlation only. All possible effort has been made to limit the number of ETPs of this type (usually leveraged ETPs or ETNs, ETNs or actively managed ETPs).

- The Exposure Type (long only, short only, long/short exposure) must be the same for a pair of ETPs to be similar. These are calculated and stored in the DailyWebsiteETPInfo table on a daily basis. The field is ShortExpID and the descriptions of the different types are in the XTFETPExplorerShortExposureType table.
- ETP Type: No restriction on ETNs, ETFs. ETMFs are excluded from the process altogether.
- Options Strategy ETPs
    - Are preferentially compared with each other but not restricted.
    - Volatility ETPs are in this group.
- Asset Class
    - ETPs are paired only with others in the same asset class.
        - Real Estate ETPs, although they have equity constituents, can only have other Real Estate ETPs that are similar.
        - Multi-Asset ETPs are paired with other MA
        - Hybrid are paired with Hybrid
    - Commodity ETPs have to have some component overlap. Price basis is not considered so futures and spot based components may overlap.
    - Currency ETPs:
        - If regional, then similar also must be regional and from the same region
        - If country-specific, must be paired with country-specific currency ETPs from the same country
        - If Emerging Markets, must be paired only with other emerging market currency ETPs
- Leveraging
    - Normal (unleveraged) ETPs have similar ETPs that are also normal.
    - Leveraged and Inverse ETPs are paired with ETPs that have the same leverage sign but not necessarily the same leverage factor (e.g. 2X and 3X can be similar as can -1X, -2X and -3X, but 2X cannot be paired with -3X).
- Sector (Equity ETPs only)
    - Sector ETPs can only be paired with others of the same sector *or with a multi-sector ETP*. Relaxing the sector restriction to include multi-sector has proven to be highly significant. Pairs of ETPs with broad exposure, but exclude specific exposure can be accurately compared. The latter does not include ETPs that have "sector not specified", only ETPs whose investment objective specifies them as "multi-sector."
- Investment Philosophy
    - There are no restrictions
- Geographical
    - US, Global, International (Ex-US), and Emerging Markets specific ETPs are paired only with like kind.
    - Region specific ETPs have similar ETPs only from the group that are specific to the same region

- Country specific ETPs are paired with ETPs that are specific to the same country or with ETPs specific to the region containing that country, but the country ones are preferred (This constraint is subordinate to any asset class related constraints above that may have geographical constraints as well. Currently only Currency ETPs are such.).

- Data Restrictions
  - **ETPs need to have been in existence for at least 15 days and need to be relatively liquid within 30 days of the date when the Similar ETPs stored procedure is run in order to have a similar ETP.** The basis of the Similar ETPs algorithm is the function that determines the correlation between pairs of ETPs (even though correlation has the lowest priority of all the metrics considered). The number of calendar days used to obtain the correlations is 30. It is required that at least half of the trading dates in this 30 day period (about half of 22 = 11) of good total return data (simultaneous dates for both ETPs being compared) is required or that pair of ETPs will not be selected by the algorithm. If a given base ETP is less than about 15 days old, it will not be selected when compared to any other ETP and therefore will not be selected by the algorithm at all.

Preferential Sorting

Once the subset has been determined by the constraints above, they are sorted in order of preference and the sorting algorithm is described here. There are 4 main factors used in the sorting algorithm, the details of which are described below: Overlap (component and exposure), Objective Similarity, Correlation, and Enhanced Strategy.

For each pair of ETFs, there will be a number associated with each of these factors. The sorting algorithm prioritizes some factors over others according to asset class. The main criteria is whether or not they have equity components (Equity, Real Estate, Multi-asset) or VIX futures (VIX ETFs), and whether or not we get those component weights daily in the NSCC file. We do get components for FI ETFs daily, but the FI ETFs tend to have so many bonds in each index, and the ETFs don't always have all the index's bonds, so the overlap of components tends to detract from the similarity – i.e. you could have 2 FI ETFs tracking the same index, with the same objectives, but with a small overlap of components because each is holding a different, yet valid, subset of the index's underlying bonds.

There are 2 levels of sorting used to determine the rank of similarity between any pair of ETFs.

1. Sort1: (Overlap components + Overlap exposure) * Correlation (3 month) * ES Factor
2. Sort2: Objective Similarity * ES Factor

For Equity, Real Estate, Multi-Asset, Hybrid, and Volatility ETFs, the sorting is done by Sort1 and then by Sort2. For the other asset classes -- FI, Commodity and Currency ETPs -- the sorting is done in the opposite manner.

Similarity Factors

1. Overlap (Aggregate Weight Differences)

a. Component Overlap

For each pair of ETPs, this is calculated by summing the absolute value of weight differences of each component in either ETP, including components that are only in one of the ETPs in the pair (components get a weight of zero if missing in an ETP).

Component Overlap (ETP 1 & ETP 2) =

$$100\% - 0.5 * \sum_{All\ Components\ in\ either\ ETF} abs(weight\ of\ comp.in\ ETF\ 1 - weight\ of\ comp.in\ ETF\ 2).$$

Or, written another way:

$$\text{Component Overlap}\ (ETF\ 1, ETF\ 2) = 100\% - 0.5 * \sum_{Components\ i=1}^{n} abs(w_{1,i} - w_{2,i}).$$

For pairs of ETPs with perfect overlap, each weight difference will be zero and therefore their sum will also be zero. So the second term in the equation above will be zero and the component overlap will be just the first term – 100%. An example of perfectly overlapping ETPs is comparing one ETP with itself, but otherwise it never occurs for equity or fixed income ETPs. An example of almost perfect overlap is the SPY-IVV pair, whose aggregate absolute weight difference is less than 1% for all 500 components, giving an overlap of > 99% (both track the S&P 500). If a pair of long-only ETPs have no overlap – no common components – their aggregate weight difference will be 200%. So the range of aggregate weight differences (for long-only ETPs) is 0% to 200%, with the lower values being the most similar, and the higher values the least:

$$0\% \leq \sum_{Components\ i=1}^{n} abs(w_{1,i} - w_{2,i})\ (long\ only\ ETFs) \leq 200\%$$

For pairs of ETPs where one or both of them is a long-short ETP, the aggregate weight difference can be outside of this range significantly.

The factor of one-half in front of the aggregate weight difference in the Component Overlap formula above is now evident; it renormalizes the weight difference aggregate to give a clean value of 0-100% for the overlap (only true for long-only ETPs).

b. Exposure Overlap

Exposure overlap is similar to the weight difference method above but instead of using an ETP's component weights, it uses an ETP's objective exposure (see the "Fund Exposure" tab for any ETP on xtf.com). XTF doesn't use all exposure categories and the ones it uses vary by asset class:

Asset Class	Exposure categories used
Equity	Sector, Industry
Fixed Income	Maturity, Credit Grade, (Sub-) Asset Class
Commodity	Hard Asset, Commodity
Currency	Currency
Real Estate	Sector, Industry
Multi-Asset	Sector, Industry, Maturity, Credit Grade, Hard Asset, Commodity
Hybrid	Sector, Industry, Maturity, Credit Grade
Volatility	None (not calculated for Volatility ETPs)

2. Objective Similarity

ETP objectives like geography, market capitalization (for equity ETPs) and credit grade (for Fixed Income), are weighted to give a definite order of priority. All ETPs that have the same geography as the target ETP will get a +4 in the geography objective otherwise they will get a -4 (see exceptions below). The other fields work the same way but have different weights. All of the fields for the given asset class of the ETP are summed together to generate a single value for this field. The weights establish the priority of one field over another, but it is the sum of the weighted field values that gives the sorting field value.

Here is the list of fields and their weights:

a. All ETPs
   i. Geography (weight = 4)
   ii. Investment Philosophy (weight = 1)
b. All non-Fixed Income ETPs (except commodity ETPs)
   i. Style (weight = 3)
   ii. Capitalization (weight = 2)
c. Fixed Income ETPs
   i. Maturity (weight = 4)
   ii. Debt Issuer (weight = 3)
   iii. Credit Grade (weight = 2)

3. Correlation

We calculate ***correlations*** between daily total returns over the previous 3 months of every pair of ETPs.

4. Enhanced Strategy

This is a special objective type overlap for enhanced strategy ETPs. We calculate and Enhanced Strategy factor (ES Factor) for each pair of ETPs. The algorithm is simple: It treats pairs of ETPs with the same enhanced strategy preferentially, is neutral to pairs of ETFs where neither member is enhanced strategy, and penalizes pairs of ETPs with different enhanced strategies, including if one is enhanced and the other is not. Here are the quantitative values for each pair of ETPs (the value of ϵ is determined empirically):

- Same enhanced strategy type: ES Factor = $1 + \epsilon$,
- Neither is enhanced strategy: ES Factor = 1,
- Different enhanced strategy type: ES Factor = $1 - \epsilon$.

Timing/Frequency

The similar ETPs stored procedure is run every Sunday and takes about 1.5 hours to complete. The process could be run daily, but running the comparison on a weekly basis has proven to be statistically sufficient.