

Morningstar ETF scraper test

Overview

To create a scraper than can capture data on Exchange Traded Funds from the Morningstar website and record the results.

Main site: <http://portfolios.morningstar.com>

I don't believe you need an account but it is free to make one if you do.

Technology

We recommend using [headless Chrome](#) to allow the scraper to be run in parallel and using SQLite to store the results.

Pages to be scraped

Holdings data, e.g. ASX:STW: <http://portfolios.morningstar.com/fund/holdings?t=XASX:STW>

Data points:

- Company Ticker (e.g. ASX:CBA), note this can be found in the latest analyst report URL
- Company Name
- % Portfolio Weight
- Shares Owned
- Sector (record title tag)
- Style (record class)
- First Bought
- Country
- YTD Return

Deliverable

Database of results (ideally SQLite) and project code.

Index list

See attached CSV which contains a list of ETFs to be scraped. This is in exchange:ticker format, which might need to be modified to work with the format used by Morningstar.

Also worth noting the typically US equities/ ETFs do not need the exchange symbol, just the ticker.

For some reason different sites use different exchange formats, the index we have given you only has ASX, LSE and ARCA. In Morningstar these are XASX, XLSE and ARCX.

Bonus Task

Scrape the Fee and Expense data from

<http://financials.morningstar.com/etfund/operations.html?t=STW®ion=aus&culture=en-US>