**Goal:**

Within 2 weeks, collect and use relevant text data (such as news, company announcements, policy releases, earning call transcripts, etc.) for predictions of stock returns.

**Details:**

* Coverage: current FTSE China A50 index constituents.
* Text data collection: web scraping few days of live text data is expected to show that you have the skills to do so; however, to conduct meaningful research and backtests, obtaining historical data from whatever sources you find useful is necessary; preprocessed historical data such as sentiment scores is accepted, but then we’d have higher expectation on its results than using raw original text as input.
* Market data: we assume you could find access to daily close prices of these stocks (or higher frequency if needed, but please start with daily data). If indeed have trouble to obtain, let us know.

**Output:**

By the end of the project, please write up a concise but self-explanatory report presenting the processes and findings of the study.

As you can tell, this is a toy project mainly to assess your abilities to divide and conquer a problem, i.e. identify problems, draft plans, make assumptions, access and process data, build models, test, revise based on feedbacks from results, communicate and present, etc. We totally understand the limitations given scarce resources -- do as much as you can and that’s enough.

When clarification is needed at any stage, let us know – do not blindly proceed without making sure what to do. Time extension is possible upon reasonable request.

Thanks and good luck.