Final Project, Financial Markets Analytics 2023

Inspired by the investment strategies defined as "successful" in the Book QEPM (starting on p. 166 et seq.), develop a screening model using as input the data contained in the "euro.xls" file provided in the e-learning space.
The file includes several key indicators for companies:

```
% 1 PE_RATIO
% 2 FIVE_YR_AVG_PRICE_EARNINGS
% 3 T12M_DIL_PE_CONT_OPS
% 4 10_YEAR_MOVING_AVERAGE_PE
% 5 PX_TO_TANG_BV_PER_SH
% 6 CURRENT_EV_TO_12M_SALES
% 7 CURRENT_EV_TO_T12M_EBITDA
% 8 FIVE_YEAR_AVG_EV_TO_T12_EBITDA % 9 T12M_DIL_EPS_CONT_OPS
%10 TRAIL_12M_EBITDA_PER_SHARE
%11 TRAIL_12M_SALES_PER_SH
%12 NET_DEBT_PER_SHARE
%13 TANG_BOOK_VAL_PER_SH
%14 NORMALIZED_ACCRUALS_CF_METHOD
%15 EBITDA_MARGIN
%16 EBITDA_MARGIN_3YR_AVG
%17 RSI_14D
%18 PX_LAST
%19 MOV_AVG_50D
%20 MOV_AVG_20D
%21 MOV_AVG_10D
%22 MOV_AVG_5D
%23 MOV_AVG_40D
%24 MOV_AVG_30D
%25 3MTH_IMPVOL_90.0%MNY_DF
%26 VOLATILITY_90D
%27 VOLATILITY_30D
%28 VOLATILITY_180D
%29 CAP_EXPEND_TO_SALES
%30 PX_VOLUME
%31 T12M_DVD_PAYOUT_RATIO
%32 EQY_DPS_NET_5YR_GROWTH
%33 EQY_REC_CONS
%34 BEST_EPS
%35 WACC_COST_EQUITY
%36 NORMALIZED_ROE
%37 5YR_AVG_RETURN_ON_EQUITY
%38 CUR_MKT_CAP
%39 NORMALIZED_ACCRUALS_BS_METHOD
%40 PX_TO_BOOK_RATIO
```

The data refer to a time period of 111 months, while the number of securities is 797.

1. Analyze the achievable strategies based on the univariate ranking against the available factors, calculating appropriate absolute and relative performance measures to the Benchmark and the IR to establish a ranking on the different strategies.
2. Always consider transaction costs at a level set by you and freely choose the number of securities to include in the portfolios of the different strategies.
3. Construct a multivariate strategy (z-score) by considering the different factors jointly. Try to base it on the IRs obtained in step 1 and an equal-weighted solution.
5. Try to solve the problem for which the correlation between univariate strategies is not considered in a simple z-score. Propose some solutions, also just theoretical.

6. You can consider sequential screening if you want (bonus point).

Execution details

Tips and rules:

This work is expected to be carried out by groups, preferably of 2-3 people. The expected outputs are codes in working R/phyton, preferably on Notebook, where the work done, choices, data used and results obtained are summarized through appropriate graphs and tables with annualized performance statistics.

1. You are free to make as many assumptions and hypotheses as you see fit to complete your task; it is essential, however, that these are made explicitly clear;
2. You may also modify one or more subpoints of the 'exercise if you have "creative" and more interesting insights.
3. Write the paper as if you were handing your boss a work report, so: be clear, complete, concise and convincing in justifying your choices;
4. Count towards the final exam result: the work grade ranges from 0 to 30 points and weighs 50% towards the final grade and then the oral exam. The oral exam can also be taken in a different exam call than the group presentation. The project grade is also related to the group size; the more participants, the higher my expectation of the outcome.
5. Positive elements for evaluation: critical elements for a good review will be considered, clarity of presentation, ability to answer each of the subpoints of the exercise, neatness and clarity of the code (insert comments on the various pieces of the code to increase readability).

Challenging tasks

Try using techniques you are familiar with, such as Deep Learning, Support Vector Machine, Random Forest, or ensemble, on a portion of the data sample sufficient to give statistical robustness, which will be used as a rolling window. The explanatory variables for each asset (returns being the variable to be explained) will always be the foundational or momentum factors from the previous exercise.

As a bibliographical point of reference related to the articles:

(a) L. Takeuchi and Y.Y. Lee, "Applying Deep Learning to Enhance Momentum Trading Strategies in Stocks" (2013).

(b) R. Huerta, F. J. Corbacho, C. Elkan, "Nonlinear support vector machines can systematically identify stocks with high and low future returns." Algorithmic Finance 2(1): 45-58 (2013).