2017 Morgan Stanley Prize for Excellence Project

Introduction

Morgan Stanley has a leading securities market making business globally, and data analysis is an essential component in such a business. Predicting future returns is, and always will be, one of the greatest challenges any securities markets participant faces. In this age of readily available, high-quality time series data, a key component of predictive analysis is the identification of which variables (out of a vast universe of possibilities) are actually relevant to the prediction one is attempting to make.

With this in mind, we have designed our 2017 Prize for Excellence project as an open ended exercise in prediction, where the participant will use a large repository of time series data, in any way he or she chooses, to predict daily returns for 3 benchmark series in *either* the interest rates, foreign exchange, or equity markets.

The deadline for submitting all reports and code is December 2, 2016.

Data Description

FRED (Federal Reserve Economic Data) is an electronic data repository hosted by the Federal Reserve Bank of St. Louis. FRED provides hundreds of thousands of economic time series to the open public. Participants are welcome to use any available time series, but are encouraged to provide quantitative and/or qualitative justification.

For this competition, participants are expected to write a program which interfaces with and retrieves data programmatically from one of the FRED API (Application Programming Interface). *The accepted languages are* Python, R or Matlab, each of which has working packages that utilize the FRED API. For further details specific to each language, please refer to:

- Python: https://github.com/zachwill/fred
- Matlab: http://www.mathworks.com/help/datafeed/fred-1.html
- R: https://cran.r-project.org/web/packages/quantmod/index.html

Participants may use other packages to connect to FRED provided they are cleared with the Project Committee ahead of time.

Prediction

Valid submissions are required to generate predictions for 10 consecutive business days of daily returns after December 2, 2016 (note that *which* 10 days is not being disclosed) for all 3 benchmarks in *one* of the 3 asset classes below (e.g. an equity project would predict Dow Transportation, S&P 500, and VIX returns… no mixing and matching between asset classes).

- Interest Rates (Treasury Constant Maturity): 2y, 5y, 10y
- Foreign Exchange (FX): JPY/USD, USD/EUR, USD/GBP
- Equities: Dow Transportation, S&P500, VIX

There is no limitation on the number or type of time series to be used in the model. However, it is highly encouraged to be creative rather than just use brute force.

Assumptions and Rules

1. Please make sure to include in the project submission everything necessary to evaluate performance (code *must* run to be considered a valid submission).
2. Please do *not* include any data in the submission. It is required to utilize the FRED API to query for the necessary data programmatically.
3. Comment all code, and clearly mark sections such as data retrieval and model fitting.
4. The judges' primary interface to submitted code will be via a main function, provided by the participant, with two input parameters:
    a. Prediction start date
    b. Prediction end date

    Training must only use data prior to December 2, 2016. At such time as the judges run the submitted code, it must not attempt to use data in the prediction period to fit the model.
5. The main function should return two structures:
    a. The predicted daily returns of the three benchmark series in tabular form
    b. A structure containing all relevant fitted parameters used in the model.
6. Accepted programming languages are Python, R, Matlab.
7. Please clearly state packages/libraries used in the code, and avoid using any commercial packages/libraries. The use of *any* 3rd party packages/libraries, regardless of the programming language, must be cleared with the Project Committee ahead of time.

Evaluation Guidelines

1. A portion of the score assigned to each project will be derived from a quantitative assessment of prediction quality. The predictions will be evaluated on an undisclosed prediction period after December 2, 2016. Submissions must be able to predict any 10 consecutive business days of daily returns after this time.
2. The remainder of the score will be determined by design, innovation, and exposition. Please provide a carefully and clearly documented model description in the report, as well as any testing results. Better diversification, better economic rationale, and innovative approaches to data analysis will impact the final score positively.

Attachments

Item	Description	Folder
PrizeForExcellence2017.pdf	Problem description file	./problem

Final Submissions / Questions

Final submissions (as well as any questions) should be e-mailed to MSSM.Recruiting@morganstanley.com by **December 2, 2016.**

Note: Finalists will be invited into our headquarters in NYC to present findings on **January 5, 2017**. If you are selected as a finalist, you will be notified no less than 1 week prior to that date and travel arrangements will be made at that time. Please keep this date available on your calendar. The first place winner will be determined and announced on this date.