Investment Strategy - Final Project

**What Does The Model Do?**

The model takes as its input a set of 19 different types of data covering financial, macro-economic and behavioural features across 14 different equity markets. This data (commonly known as ‘factors’ in finance or ‘features’ in machine learning) is thought to be important in predicting which markets will outperform over the next investment period. If we can identify what features are important for predicting outperforming markets then we can generate profits by investing exclusively in those markets that are predicted to perform the best and ignoring the others.

**What Data Did I Use?**

The data comes from many different commercial sources – Bloomberg, MSCI, OECD – and is processed to create market features.

**What Model Did I Use?**

The model is a random forest classification model. Its job is to classify what features are the most important in predicting better performing markets.

**How Did I Optimise the Hyperparameters**

The random forest’s main hyperparameters are set in the in-sample (or training/validation phase) using Bayesian optimisation and then used in the out-of-sample or testing phase.

**Model Performance**

To assess the model’s performance we look at the three most important features that the random forest identifies as predicting excess returns and then use these to select markets out-of-sample.

As noted in the graph below, these features that the random forest identifies (collectively called ‘Strategy’ below) do indeed generate significant outperformance over the testing period. The strategy significantly beats the benchmark of the average market performance (called ‘Equally Weighted BM’ below) over the period and also outperforms on a risk-adjusted basis as noted by the Sharpe ratio calculation.

