A draft of the final project

You are a quantitative investment analyst. There is a mandate for you to come up with a long only product with tracking error of 3%.   Your job is to present a trading strategy given the universe in your homework. The benchmark is also the index associated with your universe. In your work, you need to present your idea based on the following steps.

1) Your alpha model.
2) Construct your portfolio given a Tracking error constraint.
3) Your portfolio should not have any beta bet.
4) Other related constraints would be suggested.

Your Universe

 You can assume that the universe is the top 30 names by market capitalization of your index (SP500 or NASDAQ). That is, the data you already have for homework 1. But you can also use a larger or different universe with my consent.

Alpha Model

Explain what you would do to construct an alpha model. You need to be able to explain the idea behind your methodology. When in doubt ask me. You should be able to present the pros and cons of your ideas. For example, would this idea work all the time? If not, when might it work and when might it not work? Also you need to see if you can explain your alpha idea by the well-known factors (like Fama-French factors and Carhart's momentum.)

Portfolio Construction

Given your alpha model, construct a portfolio using the data with the given tracking error and other constraints.

Analysis of Your Results

Present the performance of your strategy. For example, you should know what the IR of your strategy is. As the class progresses, we would learn more on performance evaluation and attribution.

Sample for topics you can use for the alpha.

- Fundamental Factors: Finding fundamental factors (e.g., Book to Price) that can help explain future returns.
- Statistical factors and their predictive/explanatory power: By following Connor and Korajczyk (1986, 1988), you can employ the principal component technique to extract potential statistical factors that can predict or explain stock returns.
- Higher moments and their predictive/explanatory power: There is an area of research showing that higher moments, e.g., skewness and kurtosis, could be helpful in term of explaining/ predicting stock returns.
- Economic factors and their predictive/explanatory power: By following Chen, Roll and Ross, we can try to uncover economic factors that could be explaining/predicting stock returns. I want to emphasize the use of other factors in combination with the factors cited in the paper.
- Time series or other advanced econometrics technique and their application in predicting stock returns: By choosing a time-series of your choice, you can explore a technique that could be helpful to predict stock returns. Since I have seen many projects using ARCH/GARCH/EGARCH, I would love to see other techniques for this project.
- Stock return distribution and its implication in testing/predicting stock returns: The distribution of stock returns is not normally distributed. How could we use this fact to help us predict stock returns and/or test an empirical model?
- Timing: Can you time the market? Can we find a way to predict the up-down cycle of market returns?
- Behavioral Reason: We may not be able to cover this topic in depth in this class. However, it is a very fruitful area of research. For example, investors might over-react or under-react to specific shocks. Or they might behave differently regarding profit and loss. Can we use them to predict stock returns?
- Other exogenous factors that can be used to predict stock returns.
- Other econometrics techniques to help predict future stock returns.
- Sector Selection: Instead of picking stocks, you can try to tactically select sectors. Can you use any economic or any other variables to predict the sector returns?
- Data Mining: How to use data mining technique to come up with an idea for an alpha model?
- Machine Learning: Would you be able to utilize the most recent development in machine learning to build an alpha model?

Additional Guideline on Your Final Projects

What you should have in your project.

1) You need to have the following sections.
    a. Introduction: State briefly what are your motivation, methodology, and findings. Think about it as an executive summary.
    b. Literature Review: Explain the paper(s) you are using for your research.
    c. Methodology: If your methodology is new (not covered in class), you need to explicitly explain it. Although some methodology may be introduced in class, it may not hurt to talk a bit about it.
    d. Data: Where did you get your data? How did you construct your factor/moments/other relevant metric for your work? What is the data coverage? You should also put a table with summary statistics of your data (number of observations, mean, median, min, max, std, etc.)
    e. Empirical session:
        i. For those groups working on fundamental factors, you should try to show why your factors should be selected. You should also try Fama-French three factors analysis. That is, how it looks when you include Fama-French three factors? For other groups, it would not hurt to show what make your alpha model interesting. Most of your references should already have a section doing so. You can follow them.
        ii. After you have come up with an alpha model, you need to perform potentially rolling window backtest. To be more specific, you use an estimation window of, say, 3 years, to predict the next 1-day or 1-month return (alpha). Form a portfolio according to your model. Then, compute the portfolio return. Then, you roll to the next estimation window and repeat the prediction process and portfolio construction until you exhaust your data. From those portfolio returns you construct, you then generate
            1. IR
            2. IC
            3. Tracking Error
        iii. Explain your results. For example, your findings might be completely different from your references. You need to come up with a reasonable explanation.
    f. Summary: Summarize your work and can indicate if you want to extend your work, what could be done?
    g. Appendix: Put all of your programming source codes here. In addition, you should submit all of your source codes and data with your project. In the zipped file, you need to have the following subdirectories. Keep in mind that if I unzipped your file, I should be able to run your programs and generate those important results.
        i. Data: Contain your raw data
        ii. Programs : Contain all of your programs

Note that for each section, you need to have figure(s), and/or table(s) as needed.