**Project 1: Constructing Big data Alpha Model**

Due: Week 8 (or week 9 depending on the class pace).

Group assignment with 3-4 members (ensure your group has at least one member good at programming and one knowledgeable on financial trading).

**Description**

The project is designed to integrate what you learned in class to construct a big data alpha model. You will identify a portfolio of stocks (of your own interest); retrieve various data that you think will help you predict the future price of these stocks; process the big data and construct factors accordingly; implement the big data alpha model (follow the R example codes in class) using R (or other computer language of your group’s choice); examine and report its performance.

**What to do**

Specifically, your project should at least consist of these steps:

1. Choose a portfolio of securities
   * Can be stocks, cryptocurrencies, ETF etc.)
   * You decide on portfolio size, minimum two
2. For every stock, retrieve at least a couple of years’ data
   * retrieve the fundamental data you deem relevant
     + e.g. Financial statement data (quarterly at least), macroeconomics data (monthly or quarterly)
   * retrieval the daily trading data
     + Practice with charting tools in Quantmod
     + Extract the technical analysis indicators
   * Retrieve data from social media (e.g. twitter or reddit or other social media sources)
     + Analyze the data to calculate media attention level, sentiment etc.
     + Hopefully every team can think of at least one thing innovative in big data analytics, e.g. Trump factor, informal news (reddit), create TA for sentiments, geographic location analytics, follower-followee social network, analyze co-attention and so on

3. Identify good factors from the data

* + Evaluate which factors predict the next month’s return best

1. Execute the big data alpha model and report the performance

**Deliverables**

* Write a concise word report (no more than 10 pages) that document what you have done , submit your code separately
* Submit to elearning
* I will choose 1-2 best groups to present in class