The class has been divided into ten groups. Each group has been assigned a stock for a publicly traded company. Analysis will be performed on the returns of the stock, and a model created to understand the extent to which the stock price performance can be predicted by the performance of the overall stock market, as well as by the size and valuation of the company. Each group must perform all of the below tasks, and present the results in a slide presentation.

All below items, with figures, statistics, plots and diagrams, should be included in the slide presentation. I expect explanations and comments to be on the slide, do not simply include statistics and comment verbally. You need to clearly explain, verbally and on the slide what you have done, why you have done it, and how you come to your conclusion. Pretend I do not know statistics. Your presentation neatness and clarity will be considered. Did you just copy and paste regression output onto a slide? Or did you take time and make your slides look professional? Only the figures you are discussing in the moment need to be on the slide, and they should be clearly visible from the back of the room. When I review your slides at the end of the term, you do not want there to be any question in my mind that you know what you are doing. Your project grade will depend on the performance and presentation of all tasks, as well as overall presentation clarity, writing quality, neatness and organization. Presentation order will be selected at random, and *your group will be penalized if all members are not present when it is your turn to present*. Often students like to add additional calculations I do not ask for. I do not mind this, and it can make the presentations interesting, but your grade will depend on how well you follow instructions and deliver what I ask for. A small part of each member's project grade will be based on their assessing members' respective contributions, as well as the assessment from the members. Refer to the syllabus for the specifics of the project grading.

I intend to look at your presentations carefully, and ask specific questions. All members of your group should be prepared to answer questions on any part of the presentation.

Your group can be found in Bright Space. The name of your group is the ticker symbol for the stock you will be analyzing.

The second column of the datasheet is titled RF. RF stands for risk-free. It is the rate one could earn by holding 1-month US Treasury bills. So, in evaluating what a stock earns, one is interested in what it earns above the "risk-free" rate, which one could earn without taking the risk of holding a stock. The data for the project therefore consist of "excess" returns, or the additional monthly return that can be earned by taking risk to invest in a stock or portfolio of stocks, as opposed to holding 1-month Treasury bills. You do not need to subtract RF from the monthly returns, it has already been done for you.

To begin, we use as independent or explanatory variable the excess return of the market. The series in the next column called Mkt-RF is the excess return on a portfolio consisting of all listed stocks on US stock exchanges.

(1) Produce histograms and boxplots for the excess returns of your stock and the market portfolio, and comment on the distributions. (At this point I am very interested in the quality of your histograms. They should be actual histograms, not bar charts, with all of the characteristics of histograms we have learned. Histogram quality also pertains to bin selection, which is a judgment. Do your histograms depict the data well? Are all the data clumped up into just a few bins? Or unnecessarily spread out with gaps? It is not simply performing a task and checking it off the list. Your job is to depict the data in a way that best captures the nature of the distribution. Do the job well, as an employer would expect of you.) Comment on the distributions. Your comments need not be overly detailed, just highlight the basics; symmetry?, how tight is the distribution?, outliers?, and anything that stands out.

(2) Produce a scatterplot with your stock returns on the vertical axis and the market returns on the horizontal axis. Include trendline, regression equation and R^2. Explain and interpret your plot, equation and coefficient of determination.

Eugene Fama and Kenneth French developed a model that explains variation in stock returns based on three factors. These factors account for differences in return on the stocks of small companies versus large, difference in return on value stocks versus growth stocks, and the return on the market as a whole. I would like for you to estimate this model to explain the variation of returns for your stock.

The three factors (or independent variables) are:
MKT: the series you have already been using, which is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks.
SMB (*Small-Minus-Big*): return on a portfolio of small capitalization stocks minus the return on a portfolio of large capitalization stocks
HML (*High-Minus-Low*): return on a portfolio of high book-to-market ratio stocks (also called value stocks) minus the return of low book-to-market ratio stocks (called growth stocks)
These three variables are known as the *Fama-French* factors. These additional two data series, SMB and HML you will also find in your data file.

(3) Estimate the linear factor model $R_t^{stock} = \alpha + \beta_{MKT}\, R_t^{MKT} + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \epsilon_t$. Include 99% confidence intervals for the coefficient estimates.
(4) Interpret your model. What do the coefficients mean? Be specific.
(5) Discuss the goodness of fit of the model. Which measure do we use? Why? What does it mean? How is it calculated? How much of the variation in your stock can be explained by your model. Refrain from blanket declarations of "good" or "not good". These are not useful, as such proclamations depend on contextual knowledge you do not have.
(6) Discuss the overall significance of the model. Does it have explanatory power for your stock returns? Explain. Be clear and thorough. How are you assessing overall model significance? How do you come to your conclusion on whether the model has overall significance?
(7) Discuss the significance of the individual coefficients. Which ones have significant explanatory power and which ones do not. Reconcile your confidence intervals with your p-values.
(8) Test the null hypothesis that $\beta_{MKT} \leq 1$ against the alternative that $\beta_{MKT} > 1$ at 1% significance. (Be very careful to pay attention to what is being asked here).
(9) Calculate the pairwise correlations amongst MKT, SMB, and HML and discuss whether there could a problem of multicollinearity in the regression. You may do these individually or produce a correlogram. Be specific in your comments what we are looking for here, what the potential harm is and how conclusion is reached.
(10) Test the null hypothesis that $\alpha \leq 0$ against the alternative that $\alpha > 0$ at 5% significance. Again, be careful to pay attention to what is being asked for. Comment on the results of this test in light of the Fama-French Three Factor Model. Is there evidence that the stock provides significantly positive returns after accounting for size, valuation and the market?
(11) Assess the validity of the model. Do you have confidence in its validity, why or why not? Do you have reason to doubt the model's assumptions? Be clear and detailed in your explanation here. Discuss your assessment using a standardized residual plot.
(12) Download the file titled Group Project Peer Evaluations. Follow instructions and upload in the designated place in Brightspace. Do this individually. I alone will see the assessments.

Our presentations will take place in the classroom on April 27 and May 4. Presentation order will be chosen at random. Therefore all groups must be submit their presentations on BrightSpace and be prepared to present at the beginning of class on April 27. I will take role on both of these days, and your attendance will count as your participation grade. If any member of your group is not present when it is time for your group to present your group will be penalized.